M | MICHIGAN ROSS

Working Paper

Analyst Information Discovery and Interpretation Roles: A Topic Modeling Approach

Allen Huang
Hong Kong University of Science & Technology
Department of Accounting

Reuven Lehavy
Stephen M. Ross School of Business
University of Michigan

Amy Zang
Hong Kong University of Science & Technology
Department of Accounting

Rong Zheng
Hong Kong University of Science & Technology
Business Statistics and Operations Management

Ross School of Business Working Paper
Working Paper No. 1229
November 2016

Analyst Information Discovery and Interpretation Roles: A Topic Modeling Approach

Allen Huang
Department of Accounting
Hong Kong University of Science and Technology
allen.huang@ust.hk

Reuven Lehavy
Ross School of Business
University of Michigan
rlehavy@umich.edu

Amy Zang
Department of Accounting
Hong Kong University of Science and Technology
amy.zang@ust.hk

Rong Zheng
Department of Information Systems, Business Statistics and Operations Management
Hong Kong University of Science and Technology
rzheng@ust.hk

November 2016

Allen Huang, Amy Zang, and Rong Zheng would like to express thanks for financial support provided by HKUST. Reuven Lehavy would like to express thanks for financial support from the Harry Jones Endowment Fund. We would also like to thank the Editor, the reviewers, Zahn Bozanic, Lian Fen Lee, Weining Zhang, Bin Zhu, seminar participants at Boston College, Chinese University of Hong Kong, HKUST, IESE Business School, Nanyang Business School, Shanghai University of Finance and Economics, Singapore Management University, Southern Methodist University, Tel Aviv University, Tsinghua University, University of British Columbia, University of Melbourne, and University of Technology Sydney, as well as the participants at the 2014 Workshop on Internet and BigData Finance (WIBF), the 2014 China Summer Workshop on Information Management (CSWIM), the 2014 MIT Asia Conference in Accounting, and the 2016 New York University Accounting Summer Camp for helpful comments. An Internet Appendix is available at http://tinyurl.com/hofgz4t.

Analyst Information Discovery and Interpretation Roles: A Topic Modeling Approach

Abstract

This study examines analyst information intermediary roles using a textual analysis of analyst reports and corporate disclosures. We employ a topic modeling methodology from computational linguistic research to compare the thematic content of a large sample of analyst reports issued promptly after earnings conference calls with the content of the calls themselves. We show that analysts discuss exclusive topics beyond those from conference calls and interpret topics from conference calls. In addition, we find that investors place a greater value on new information in analyst reports when managers face greater incentives to withhold value-relevant information. Analyst interpretation is particularly valuable when the processing costs of conference call information increase. Finally, we document that investors react to analyst report content that simply confirms managers' conference call discussions. Overall, our study shows that analysts play the information intermediary roles by discovering information beyond corporate disclosures and by clarifying and confirming corporate disclosures.

1. Introduction

Financial analysts play an important information intermediary role in capital markets. The culmination of their efforts are the research reports distributed to investors, which contain several quantitative summary measures, including earnings forecasts, stock recommendations, and target prices, as well as a textual discussion about the company. This textual discussion covers a wide range of topics, such as the company's current and future financial performance, recent corporate events, business strategies, management effectiveness, competitive landscape, and macroeconomic environment. Extant literature generally suggests that these analyst outputs provide value to capital market participants (e.g., Bradley et al. 2014; Huang et al. 2014; Li et al. 2015). To advance the literature, several review papers (Ramnath et al. 2008; Beyer et al. 2010; and Bradshaw 2011) call for additional research to better understand the sources of analyst value.

This study investigates how financial analysts serve their information intermediary role by conducting a large-scale comparison of the textual content of analyst research reports to that of closely preceding corporate disclosures. Specifically, we employ a topic modeling method to compare the thematic content of a large sample of analyst reports issued on the day of and the day following quarterly earnings conference calls (hereafter, prompt reports) to that of managers' narratives in these conference calls. Quarterly earnings announcements and their related conference calls are arguably the most important corporate disclosures. Accordingly, an overwhelming number of sell-side analyst research reports are issued immediately following these corporate events, because only timely reactions to these events can offer the analyst clients an informational advantage in trading.[1] The textual comparison allows us to investigate the following questions: (1) What type of information do analysts provide in prompt reports? (2) Do analyst discussions of new topics and of conference call topics provide incremental value to investors? And (3) under what conditions do analyst reports provide more value to investors?

[1] Indeed, we find that 46.5% of analyst revision reports are prompt reports. An All-Star analyst from a large brokerage house we interviewed commented that "*the market is efficient but impatient. Analysts need to feed the market with prompt reaction to management's thinking and outlook guidance. In addition, competition forces everyone to issue reports quickly; otherwise their reports may not be read by clients.*"

As information intermediaries, analysts can provide value to investors in two ways: First, through their private research efforts, they collect and generate information that is otherwise not readily available to investors; second, they could facilitate investors' understanding of the existing public information by analyzing and clarifying it and by offering their own opinions on issues raised through public disclosures. Following the literature (e.g., Ivković and Jegadeesh 2004; Chen et al. 2010), we term these efforts the analyst information discovery role and the analyst information interpretation role, respectively.

Several sources of potential value can arise from the analyst information discovery role: Analysts conduct their own private research and channel checks, for example, by visiting stores and warehouses, investigating supply chains, and surveying customers;[2] they have *private* interactions with not only CEOs and CFOs, but also division-level managers from operating regions and product lines (Soltes 2014); they package information collected from multiples sources, such as other information intermediaries, peer firms in the industry, independent research agencies, and government agencies, and undertake original analysis by "connecting the dots"; and they generate new information signals, such as firms' valuations, earnings forecasts, and long-term growth rates, using their high level of financial expertise. In our setting, discovery reflects analysts' private efforts to generate new topics that are otherwise not readily available in the conference call, but the sources of the information can include a variety of public and private channels.

We consider analysts as serving an information interpretation role when they discuss topics that have already been discussed in the recent corporate disclosure. Similar to the media's providing value through information dissemination or rebroadcasting, as shown by recent research (e.g., Miller 2006; Bushee et al. 2010; Drake et al. 2014), analysts might be able to provide value by discussing these conference call topics. First, by interpreting only the relevant topics in corporate disclosures, analysts attract and direct investors' limited attention to what they view as being important.[3] Second, analysts can clarify managers'

---

[2] For example, in a Morgan Stanley report issued on August 12, 2011, immediately after J.C. Penney's conference call, an analyst alludes to a consumer survey: "*The top reason consumers say they shop JCP is due to 'low prices, great discounts' (as per our most recent consumer survey).*"

[3] This point is supported by the All-Star analyst we interviewed, who mentioned that "*some topics discussed during call are ignored by analysts. Only those 'valuable' topics are picked up and interpreted.*"

disclosure by using their own language, offering their opinions on issues raised by managers, and quantitatively assessing management's subjective statements. Third, perceived as independent agents, analysts can enhance the reliability of statements from managers, who may suffer from agency problems. Taken together, we posit that the analyst information interpretation role helps investors understand corporate disclosures better by lowering processing costs and enhancing information quality. Whether and when investors consider these information roles as being useful are empirical questions our study attempts to address.

A few studies compare the relative value of the analyst information roles based on the market reaction to analyst earnings forecasts and stock recommendations (Ivković and Jegadeesh 2004; Chen et al. 2010; and Livnat and Zhang 2012). These studies infer analyst information roles from the timing of the analyst revisions relative to corporate disclosures. That is, they assume that analyst forecast revisions following (preceding) public announcements are more likely to reflect their information interpretation (discovery) role. We extend this stream of literature by introducing a new textual technique to construct explicit measures of analyst information roles that have been traditionally inferred from the quantitative research outputs of analysts. Specifically, we partition the discussion in analyst reports into a discussion of topics already covered in the immediately preceding calls and a discussion of new topics. The former likely provides an interpretation of the information already contained in the calls, based on which we assess the analyst information interpretation role; the latter likely provides information beyond what managers had released publicly, based on which we assess the analyst information discovery role. To extract economically meaningful topics from a large sample of analyst reports and conference calls, we exploit a topic modeling approach called *Latent Dirichlet Allocation* (LDA), an advanced textual analysis technique that uncovers underlying topics in a large set of documents based on the statistical correlations among words in these documents (Blei et al. 2003).

Our empirical measures of the analyst information roles are based on a comparison of the thematic content of 159,210 prompt analyst reports (denoted as AR) to that of manager narratives in a sample of

17,750 earnings conference calls (denoted as CC).[4] We first employ LDA to extract topics from AR and CC and then conduct a battery of validity tests to verify the effectiveness of LDA in identifying economically interpretable topics. When we compare the thematic content of AR and CC, we find that analysts spend an average of 31% of their discussion on exclusive topics that receive little or no mention by managers, and thus 69% of their discussion focuses on conference topics. This suggests that both analyst information discovery and interpretation roles are substantial.

Next, we find that investor reactions to both information roles are economically significant and incremental to their reaction to the conference call information and earnings news. To better understand the sources of analyst value, we predict and find that investor reaction to the analyst information discovery role is more pronounced when managers face greater incentives to withhold value-relevant information (i.e., when firms have greater proprietary cost, face higher litigation risk, or experience bad performance) and that investor reaction to the analyst interpretation role is greater when the conference call information has higher processing costs (i.e., when the call contains a greater amount of uncertain or qualitative statements, or does not deliver bad news).

We shed further light on the interplay between analyst reports and corporate disclosures by documenting that analysts respond to investor demand and exert more effort to serve in either information role depending on the features of the corporate disclosures. Specifically, analysts increase the amount of information discovery when managers are more likely to withhold information, and analysts use a greater amount of their own language to clarify management disclosures (as opposed to merely repeating managers' words) when the cost of processing management disclosure is higher. An additional analysis shows that investors value analysts' efforts. That is, the value of each information role increases with the length of the discussion; the interpretation's value increases further when analysts use their own language to discuss conference call topics.

[4] Conference call narratives include manager discussions in both the presentation and the question and answer (Q&A) parts of the conference call. Analyses based on only the presentation part of the conference calls yield similar results.

Finally, we demonstrate that analysts sometimes engage in confirming what managers say in the conference calls. Within the 69% of the discussion in prompt reports that is devoted to interpreting conference call topics, 23% does not entail a different vocabulary from that used by managers. Although such confirming discussions are unlikely to contain new insights or provide greater clarity than managers' original discussions, we find that the investor reaction to them are significant, albeit having a considerably smaller economic magnitude than that of analysts' discovery or interpretation using their own language. This evidence suggests that even when analysts selectively repeat management disclosures, they provide confirming value by identifying and disseminating the useful conference call topics to their clients (i.e., confirming these topics' importance and relevance) and by enhancing the credibility of managers' statements (i.e., confirming these topics' validity).

Our study provides several contributions to the literature. First, we provide new insight into the sources of analyst value as an information intermediaries and extend our understanding on the interplay between analyst research and corporate disclosures. In particular, we document the value of analyst discovery and interpretation roles immediately after corporate disclosure events and identify the economic conditions under which each role provides value to investors. These economic conditions relate to the features of earnings conference calls, including managers' tendency to withhold information during the calls and the processing costs of the calls' information. Second, our study introduces a textual measurement of information content to the literature, which is based on comparing the discussions of economically meaningful topics in analyst reports and management disclosures. Finally, our study contributes to the emerging area of textual analysis by introducing the topic modeling approach to the accounting and finance literature and validating the approach for financial documents (see a recent review by Loughran and McDonald 2016). Much of this research focuses on the textual characteristics (e.g., readability and tone) of corporate financial disclosures (e.g., Management Discussion and Analysis in 10-K, and S-1). Our topic modeling methodology provides another avenue through which researchers can expand their analyses of the textual content of corporate financial disclosures from "how texts are being said" to "what is being said" in these disclosures.

2. Topic modeling and Latent Dirichlet Allocation (LDA)

We obtain our empirical measures of analyst information intermediary roles by comparing the textual narratives in AR to those in CC at the topic level. To identify topics, we use LDA, which is developed by Blei et al. (2003) and has become a widely used topic modeling algorithm. LDA uses a statistical generative model to imitate the process of how a human writes a document. Specifically, LDA assumes that each word in a document is generated in two steps. First, assuming that each document has its own topic distribution, a topic is randomly drawn based on the document's topic distribution. Next, assuming that each topic has its own word distribution, a word is randomly drawn from the word distribution of the topic selected in the previous step. Repeating these two steps word by word generates a document. The LDA algorithm discovers the topic distribution for each document and the word distribution of each topic iteratively, by fitting this two-step generative model to the observed words in the documents until it finds the best set of variables that describe the topic and word distributions. Essentially, LDA reduces the extraordinary dimensionality of linguistic data from words to topics, based on word co-occurrences in the same document, similar to cluster analysis or principal component analysis applied to quantitative data. Appendix I and Internet Appendix I provide a detailed discussion of the intuition and technical features of LDA, respectively.

LDA offers several advantages over manual coding. First, it is capable of processing a massive collection of documents that would be too costly to code manually. Second, LDA provides a reliable and replicable classification of topics. Neither of these features can be attained with manual coding, which relies on human coders' subjective judgment. Third, LDA does not require researchers to pre-specify rules or keywords for the underlying taxonomy of categories. Topics and their *probabilistic* relations with keywords are discovered by LDA from fitting the assumed statistical model to an entire textual corpus. In contrast, manual coding or dictionary methods require researchers to pre-specify a *deterministic* set of rules or keywords to categorize topics. It is close to impossible to determine a priori the topics across all documents, the keywords that identify each topic for an entire textual corpus, or the probabilistic relation between keywords and topics.

To allow the LDA algorithm to fully identify the topic structure, we use all available earnings conference calls (18,607 transcripts obtained from Thomson Reuter's StreetEvents Database) and analyst reports (476,633 analyst reports obtained from Thomson Reuter's Investext Database) for S&P 500 firms during the 2003-2012 period.[5] As described in detail in Appendix II, prior to applying the LDA algorithm, we conduct several preprocessing steps to clean and parse the textual data. We conduct the LDA analysis for each industry separately, because many topics are industry-specific. The total number of topics for each industry is set at 60 based on the analysis of the Perplexity Score (discussed in Appendix II). The LDA outputs clusters of words in each topic, as well as the words' probabilistic relation with each topic. In mathematical form, it comprises a matrix of word probabilities in each topic. Using this matrix, we assign each sentence in our documents to the most likely topic, by summing up the probabilities of its words in each topic and assigning the sentence to the topic with the highest probability.

Validation of LDA outputs

We provide several validation tests for the LDA topic outputs. First, following the procedure in Quinn et al. (2010), Atkins et al. (2012), and Bao and Datta (2012), we manually read the high-probability words in key topics and their respective sentences, to provide a short and intuitive label for each topic. These labels are intended to validate that LDA is able to discern the underlying economic content of the topics.[6] Table 1 presents the twenty most frequent words in each of the top ten topics in the Capital Goods and Health Care Equipment and Services industries.[7]

[Insert Table 1 here]

Overall, the results in Table 1 validate the effectiveness of the LDA algorithm in identifying distinct, economically meaningful topics in conference calls and analyst reports. For example, the frequent

[5] We begin with 2003, as the Thomson Reuter StreetEvents database's coverage of conference calls prior to 2003 is incomplete. There are only 270 conference calls in 2001 and 1,379 conference calls in 2002 for S&P 500 firms in the database. For comparison, for the 2003-2012 period, the database contains between 1,900 to 1,950 conference calls for S&P 500 firms.

[6] These subjective labels have no bearing on the empirical analyses, because the analyses treat each topic as a distinct cluster of words regardless of the label.

[7] These two industries are among the five largest industries in our sample. Internet Appendix Table IA1 reports the keywords for the remaining three industries, including Energy, Software and Services, and Materials.

appearance of semantically related words "multiple," "target-price," "valuation," "EPS," and "price-to-earnings" in a topic in the Capital Goods industry suggests that this topic is related to "valuation model." Similarly, the frequent appearance of the words "drug," "trial," "announce," "clinical," and "phase," in a topic in the Health Care Equipment & Services industry suggests that this topic relates to drug trials. We also find that LDA effectively uncovers general topics related to a firm's financial performance, as well as industry-specific topics, such as offshore drilling in the energy industry, enterprise software and IT services in the software industry, and steel production in the materials industry (see Internet Appendix Table IA1). Finally, our results verify that the LDA algorithm recognizes the polysemy or contextual nature of words by assigning the same word to multiple topics. The word "price," for example, is related to both "valuation" and "raw materials and input price" in the Capital Goods industry, reflecting the contextual nature of the word.

In our second validation test, we compare the temporal variation in the amount of discussion dedicated to key topics with important industry and economy-wide events.[8] Specifically, Figure 1 depicts the proportion of key topics in earnings conference calls and analyst reports for the banking and telecommunication industries from 2003 to 2012 and the performance of their respective sector indices (Financial Sector SPDR – XLF and iShares US Telecommunications – IYZ, respectively). We select these two industries based on the turmoil in the banking industry and the technology evolution in the telecommunication industry during our sample period.

[Insert Figure 1 here]

Panel A of Figure 1 presents visual evidence of a reliable relation between the temporal variation in the distribution of key topics and economic performance in the banking industry. From 2003 to 2006, for example, management and analyst discussions are devoted primarily to the topics of "Growth" (mostly in loans and deposits) and "Mortgage Origination." The discussion of these topics declines substantially in 2007, however, with the advent of the financial crisis, while that of "Real Estate Loans" and

[8] A similar validation technique is used in Quinn et al. (2010), who find that the proportion of key political topics in the *Congressional Record* tracks exogenous events, such as the September 11 attacks and the Iraq War.

"Deteriorating Performance and Losses" increases. Not surprisingly, after the approval of the Troubled Asset Relief Program (TARP) in October 2008, we see an increase in discussions of the topic "Equity Issuance and TARP." Panel B of Figure 1 depicts the relation between technological developments and topic discussions for the telecommunications industry. Here, we see that landline-related topic discussions (e.g., DSL technology) decrease during our sample period, while topics labeled as "Smartphone Business" and "Wireless Subscribers" increase.

In the third validation, we compare the LDA's topic assignment to that of a human coder for a small sample of conference calls and the associated prompt analyst reports from Food, Beverage, and Tobacco industry. We first randomly select a conference call from each of the three companies in this industry (i.e., Campbell Soup, Coca-Cola, and Altria Group) and obtain the conference call transcripts and associated prompt analyst reports. Next, we invited an expert, who is an *Institutional Investor* All-Star analyst covering this industry, to label the intuition of the LDA-generated topics based on their keywords.[9] Lastly, we provided the topic intuition generated by the All-Star analyst to a human coder (a graduate student in an accounting master degree program) and asked him to label the topics of each sentence in the call transcripts and analyst reports.[10] The LDA topic assignment is consistent with manual assignment in 69%, 66%, and 60% of the sentences for Campbell Soup, Coca-Cola, and Altria Group, respectively. These consistency rates are much higher than 5%, which is the consistency rate between a random assignment of topics and manual coding.[11]

Taken together, we interpret the evidence from the validation tests as supporting the effectiveness of LDA to identify and quantify economically meaningful topics in earnings conference calls and analyst reports.

---

[9] The analyst considered the LDA topics "*quite comprehensive and meaningful*" and pointed out that "*the key challenge (of topic classification) is a wide coverage of the topics and a flexibility of topics used in different situations.*"
[10] The coder was given only the intuition of the 60 topics but not the keywords from the LDA outputs to avoid mimicking the LDA results. We asked the coder to assign each sentence to up to three topics due to the challenging nature of manual coding.
[11] If LDA randomly assigns one of the 60 topics to each sentence, the probability that this topic happens to be one of the three topics selected by the human coder is 3/60 = 5%.

3. Sample selection and descriptive statistics

The sample involved in the regression analyses is comprised of quarterly earnings conference call transcripts and analyst reports issued on the day of or the day following these conference calls for S&P 500 firms from 2003 to 2012.[12]

Table 2 describes our sample selection criteria. As shown in Panel A, we start from 18,607 earnings conference call transcripts available in the StreetEvents database. To verify that these are earnings conference calls, we match them with earnings announcement dates from I/B/E/S. This matching reduces our sample to 18,236 conference calls that occurred during days [0, +1] relative to the I/B/E/S earnings announcement dates. Next, we require each conference call to be accompanied by at least one analyst report, which yields a final sample of 17,750 earnings conference calls with matched analyst reports.[13]

[Insert Table 2 here]

As reported in Panel B of Table 2, the initial sample of analyst reports includes all reports issued for S&P 500 firms during the 2003-2012 period (476,633 reports) that we use to perform our LDA analysis. We then exclude reports not issued on the day of or the day following an earnings conference call. We also exclude reports issued on the day of a call but prior to its start time. Our final sample is comprised of 159,210 analyst reports. Prompt analyst reports constitute 33% of the entire population of analyst reports (or 46.5% if we only consider revision reports), an overwhelming percentage considering that they are concentrated in only eight days of a year. These statistics reinforce the importance of understanding the analyst information intermediary roles immediately following corporate disclosure events.

Over the entire sample period, an average of nine analyst reports are issued in the two-day window after the calls. Since our focus is on the information role of analysts in aggregate, we combine all analyst reports issued during this two-day window and denote it as AR. To examine the difference between the

---

[12] Our sample firms constitute, on average, about 72% of the total U.S. market capitalization, or 77% of the total U.S. firms covered by analysts. We acknowledge that our findings based on S&P 500 firms might not directly apply to smaller firms that receive less analyst coverage.

[13] Thomson Reuters Streetevent Database provides tickers of firms hosting the conference calls. We manually match the conference calls to Compustat's S&P 500 list using these tickers. For analyst reports, we extract firms' tickers from analyst reports and match the reports to Compustat's S&P 500 list using tickers.

respective topic proportions in the analyst and manager narratives, we conduct a Pearson's chi-square test for the homogeneity of the distribution of topics discussed in each AR and CC pair (see Internet Appendix Table IA2).[14] The homogeneity between the topic distributions in these documents is rejected at the 10% level for 91% of conference calls. That is, in 91% of the AR-CC pairs, managers and analysts devote different proportions of narratives to each topic. In contrast, the topic distributions of analyst questions (CCQ) and manager answers (CCA) in the Q&A session are significantly different at the 10% level for only 0.17% of the conference calls. This finding is consistent with intuition and provides further validation for LDA topic measures. Finally, to reduce noise, we include in the empirical analyses topics if their length exceeds 2% of the document's entire length.[15] On average, CC (AR) in our sample contain 14 (12) such topics; furthermore, the combined length of these topics accounts for over 80% (85%) of the entire discussion in the CC (AR).

4. Empirical measures, tests and results

The evidence in Section 3 suggests that topic distributions in CC and AR are different. Based on this evidence, we operationalize the analyst information discovery role as cases when analysts discuss topics that receive little or no mention by managers during the CC and their information interpretation role as cases when they discuss CC topics in their prompt reports. Internet Appendix II provides two illustrative examples for each role using excerpts from conference call transcripts and analyst reports. In our sample, analysts spend an average of 31% (69%) of their discussion on discovery (interpretation). The value of analysts' discovery and interpretation roles, however, depends on whether analyst efforts, combined with their high level of financial expertise and in-depth knowledge of the firm and industry, result in valuable information beyond the conference calls. The following empirical analyses investigate whether and under what circumstances analyst information discovery and interpretation provide value to investors.

---

[14] The topic distribution of a document can be expressed as a topic vector, in which element k is the percentage of the sentences dedicated to the discussion of topic k. Pearson's chi-square test tests the null that the topic vector of CC equals the topic vector of AR.

[15] As a robustness check, we rerun our empirical tests with topics defined as those receiving no less than 1% or 3% of the discussion, or as the top ten topics based on the proportion of discussion, and find similar results.

4.1. *Do investors value analyst information discovery and interpretation roles?*

We assess the value analyst information discovery and interpretation provided to investors by estimating the following regression:

$$CAR[0,1] = \alpha_1 Tone\_Discovery + \beta_1 Tone\_Interpret + \gamma_1 Tone\_CC + Controls + \varepsilon, \quad (1)$$

where the market reaction, $CAR[0,1]$, is the cumulative market-adjusted return during [0, 1] relative to the earnings announcement dates.[16] Because the market return is directional, we follow Huang et al. (2014) and Davis et al. (2015) and use the tone of the narratives (i.e., the percentage of positive sentences less the percentage of negative sentences) contained in AR and CC to explain CAR.[17] $Tone\_Discovery$ is the favorableness of analyst opinions contained in the new topics discussed in analyst reports, whereas $Tone\_Interpret$ is the favorableness of analyst opinions contained in the topics that appear in CC and are discussed in analyst reports. Because previous research indicates that managers' tone is sticky (e.g., Davis et al. 2015), $Tone\_CC$ is measured by subtracting the tone of the company's previous earnings conference call from the tone of the current one. Our control variables are: earnings surprises (EPS_Surp), a dummy variable indicating whether a firm's earnings miss the most recent analyst consensus forecast ($Miss$), and their interaction term ($EPS\_Surp * Miss$) to capture the nonlinear relation between earnings surprise and market returns; recent news is captured by the abnormal returns during the ten trading days prior to the report date ($Prior\_CAR$); firm characteristics that impact its information environment (Lang and Lundholm 1993), including firm size ($Size$), book-to-market ratio ($BtoM$), and number of analyst reports being considered ($\#Analysts$); and year fixed-effects.[18] Detailed variable definitions are in Appendix III. Standard errors are estimated with a two-way cluster at the firm and year levels.

[Insert Tables 3 and 4 here]

---

[16] This return window encompasses the earnings announcements, conference calls, and analyst reports in our sample. We obtain similar results using return windows of [-1, 1] and [-1, 2] relative to the earnings announcement dates.

[17] We follow the procedures described in Huang et al. (2014) to classify each sentence as positive, negative, or neutral with the naïve Bayes algorithm.

[18] In a sensitivity test reported in Internet Appendix Table IA3, we also include other research outputs contained in the analyst reports as control variables, including the revisions of earnings forecasts, stock recommendations, and target prices. The estimated coefficients on $Tone\_Discovery$ and $Tone\_Interpret$ remain significant and positive.

The summary statistics reported in Table 3 show that both mean values of $Tone\_Discovery$ and $Tone\_Interpret$ are positive (0.188 and 0.217, respectively), consistent with the overall analyst optimism documented in the literature. The average tone of earnings conference calls is 0.276, which is significantly more positive than analysts' tone, suggesting that managers in general are more optimistic than analysts.

The results of estimating Eq. (1) are reported in Table 4. We find positive and significant (at the 0.01 level) coefficients on both $Tone\_Discovery$ and $Tone\_Interpret$, after controlling for managers' disclosure, earnings surprises, and other variables that can explain market reactions, consistent with both roles providing incremental value to the market. A one standard deviation increase in $Tone\_Interpret$ increases the two-day market adjusted return by 1.09%; a one standard deviation increase in $Tone\_Discovery$ increases it by 0.65%, indicating that both roles trigger economically significant market reactions. F-tests show that the coefficient on $Tone\_Interpret$ is significantly greater than that of $Tone\_Discovery$. Overall, results in Table 4 suggest that investors value both analyst information discovery and interpretation roles and place a greater weight on their interpretation role immediately after earnings conference calls. [19, 20]

[19] The estimated coefficients on EPS_Surp and $EPS\_Surp * Miss$ imply an earnings response coefficient (ERC) of 2.828 (0.325) for firms beating (missing) analyst forecasts. These values are consistent with prior literature (e.g., Lopez and Rees 2002), but are likely too low to be considered as a reasonable price-to-earnings ratio for our sample. This may be due to the fact that unexpected earnings do not have the same degree of permanence as current earnings (Ohlson 1991) or because of non-linearities in the return-earnings relation (Freeman and Tse, 1992). To examine the impact of non-linearities in the return-earnings relation on ERC, we regress the market reaction on the earnings surprise in four different regions: large positive (EPS_Surp is larger than 0.005), small positive (EPS_Surp is between 0 and 0.005), small negative (EPS_Surp is between -0.005 and 0), and large negative (EPS_Surp is smaller than -0.005). In untabulated results, we find that the ERCs of the aforementioned four groups are 3.070, 9.410, 7.650, and 1.041 respectively, in line with results in prior studies (see, Freeman and Tse 1992). We note that the ERCs for small positive and small negative surprises are significantly higher than the ones reported in Table 4. Finally, including the interactions of EPS_Surp with these four indicator variables in all of our regressions of market reactions (i.e., Tables 4, 5, 7 and 8) yields similar results.

[20] To examine whether market reaction to analyst information roles depends on the consistency in $Tone\_Discovery$ and $Tone\_Interpret$, we include in regression model (1) the interaction terms of $Tone\_Discovery * Diff\_D\_I$ and $Tone\_Interpret * Diff\_D\_I$, where $Diff\_D\_I$ is the absolute difference between $Tone\_Discovery$ and $Tone\_Interpret$. The results, tabulated in Internet Appendix Table IA4, show significant (at the 0.01 level) and negative coefficients on $Tone\_Discovery * Diff\_D\_I$, consistent with the intuition that analysts' discovery becomes more useful when its tone is more consistent with the tone of analyst interpretation.

4.2. What determines investors' value of analyst information discovery and interpretation?

To understand the economic determinants of the value of analyst information discovery and interpretation, we estimate the following regression:

$$
\begin{aligned}
CAR[0,1] = {} & \alpha_1 Tone\_Discovery + \sum \alpha_i Tone\_Discovery * Determinants_i \\
& + \beta_1 Tone\_Interpret + \sum \beta_j Tone\_Interpret * Determinants_j \\
& + \gamma_1 Tone\_CC + Controls + \varepsilon.
\end{aligned} \tag{2}
$$

In Eq. (2), coefficient estimates on the interaction terms demonstrate whether discovery or interpretation triggers additional market reaction under various economic conditions. We conjecture that investors would place a greater value on the analyst information discovery role when managers withhold value-relevant information from investors. Prior literature on voluntary disclosure identifies several situations in which managers are more likely to withhold information, including firms with high proprietary costs, high litigation risk, and bad news, all of which we examine in Eq. (2). For the analyst information interpretation role, we posit that investors would place a greater value on this role when the processing cost of conference call information is higher. Next, we discuss our measures of proprietary cost, litigation risk, bad news, and processing cost.

Proprietary cost

Managers may choose to withhold proprietary information if disclosing it hurts firms' competitive advantage. Numerous studies on the proprietary cost of disclosure find that such costs represent a significant consequence that prevents managers from being forthcoming (see reviews in Verrecchia 2001; Dye 2001; Healy and Palepu 2001). Managers, for example, may withhold information on research and development related to an innovative product or a new drug. In this case, analysts may exert more efforts in private research, such as communicating with the company's employees, researching the company's patent filing, investigating the company's suppliers, and attending company-hosted or industry conferences to collect value-relevant information that they can provide to investors. We follow Li et al. (2013) and measure the proprietary cost of disclosure (denoted as $Competition$) as the percentage of competition references (i.e., occurrence of words related to competition) in the firm's previous conference call.[21] Li et al. (2013) argue that this measure reflects managers' perceptions of competition and thus does not rely on industry boundaries or comprehensive identification of all sources of competition (e.g., competition from private firms, foreign firms, and potential new entrants).

Litigation risk

Another factor that previous research identifies as affecting disclosure is the litigation risk faced by a firm (Healy and Palepu 2001; Johnson et al. 2001). Rogers and Van Buskirk (2009), for example, find that, despite the protection of the Safe Harbor provision of the 1995 Private Securities Litigation Reform Act, firms that have been subjects of disclosure-related shareholder lawsuits are more wary about providing information to investors. Consistent with the results in these studies, Hollander et al. (2010) find that managers are less likely to answer participant questions during earnings conference calls when litigation risk is high. We follow Hollander et al. (2010) and Field et al. (2005) and measure litigation risk using the standard deviation of monthly returns over the one year prior to the conference call (denoted as $LitigRisk$).

Bad news

Theoretical models generally predict that disclosure increases with firm performance (e.g., Dye 1986; Verrecchia 1983). When a manager has bad news to deliver, he may choose to withhold relevant information, such as the true explanations for the bad performance, because such information may decrease his human capital and reputation (Verrecchia 2001). Empirical studies generally support this theory (e.g., Lang and Lundholm 1993; Miller 2002; Chen et al. 2011). It is also possible that when there is bad news, managers are forced to focus on past performance and cannot disclose other relevant information. This is suggested in the survey evidence in Graham et al. (2005) that "if the company fails to

[21] Following Li et al. (2013), we consider a number of competition references: "competition," "competitor," "competitive," "compete," and "competing." We include words with an "s" appended and remove phrases that contain negation, such as "less competitive," and "few competitors." We also scale the number of counts by the total number of words in the document. Although Li et al. (2013) construct their measure using the MD&A section of 10-K filings, we capture managers' perceptions of competition from CC. We examine 100 randomly selected competition references from our sample and find that they highly resemble the examples provided in Appendix A of Li et al. (2013). We use the competition measure based on the firm's previous earnings conference call to mitigate endogeneity concern.

meet the guided number…the focus shifts to talking about why the company was unable to meet the consensus estimate" as opposed to talking about the firm's future prospects. For these reasons, we expect investors to place a greater value on the analyst information discovery role when firms deliver bad news during their conference calls. We measure firm news using two variables: an indicator variable of whether a firm's earnings have missed the analyst consensus forecast (denoted as $Miss$) and the earnings surprise (denoted as EPS_Surp).

Processing cost

Previous research shows that earnings conference calls may entail high information processing costs if managers' statements are unstructured, ambiguous, subjective, or qualitative (Frankel et al. 1999; Brochet et al. 2016). Prior research also documents that the demand for analyst research increases when investors' understanding of corporate disclosures requires high processing costs (Lehavy et al. 2011). Accordingly, we expect that investors find the analyst information interpretation role more valuable when the information disclosed during the conference call is more difficult to process.

We use five measures to evaluate the processing cost of conference call information. The first two are based on the notion that ambiguous language imposes higher processing costs (Epstein and Schneider 2008). Ambiguous language normally contains uncertain words and qualitative and subjective statements. We follow Loughran and McDonald (2013) and measure the percentage of uncertain words contained in a CC (denoted as $Uncertain$).[22] Specifically, when managers use words such as "may," "assume," "possibly," and "approximately," it is more difficult for investors to judge the quality of the information. Consistent with this argument, Loughran and McDonald find that having a greater number of uncertain words in Form S-1 filings increases the volatility in the valuation of the IPO. Compared to quantitative information, qualitative and subjective language is harder to process because of the lack of precision, reliability, and objective benchmarks (Huang et al. 2014). We follow Huang et al. (2014) and measure the extent to which qualitative vocabulary is used to discuss firm performance in the CC (denoted as

[22] The complete list of uncertain words is available at http://www3.nd.edu/~mcdonald/Word_Lists.html.

$Qualitative$) as one minus the percentage of sentences that contain "$" or "%." The third measure is based on the intuition that the complexity of the disclosure might increase with the complexity of firms' operations. Following Frankel et al. (2006), we measure the complexity of firm operations using the number of firm segments ($\#Segments$). The last two measures concern firm performance. Hutton et al. (2003), among others, argue that investors are naturally skeptical about good news from managers, because managers benefit from good news but have no incentives to exaggerate bad news. Hutton et al. (2003) show that bad news from managers is always informative regardless of the inclusion of supplementary statements, but good news from managers is informative only when accompanied by supplementary statements. Their finding suggests that compared to the situation of bad news, investors rely more on analysts' interpretation of good news from managers, because the information is more ambiguous and less credible. We use $Miss$ and EPS_Surp to measure firm performance.

Analyst characteristics

We consider whether analyst forecast experience ($Expr$) and their All-Star status ($Star$) influence the value of their information discovery and interpretation roles. $Expr$ is the average forecasting experience, in terms of the number of years appearing in I/B/E/S, of analysts issuing reports immediately after the conference call; $Star$ is the percentage of Institutional Investor All-Stars among the analysts issuing prompt reports. Prior research yields mixed results regarding the relation between star status and forecast accuracy (Stickel 1992; Emery and Li 2009), and between experience and forecast performance (Mikhail et al. 1997; Clement 1999); and Huang et al. 2014 find no association between star status and market's reaction to the textual opinion of analysts. Accordingly, we do not have an a priori belief as to how these characteristics may affect the value of either role.

The descriptive statistics in Table 3 show that the mean of $Competition$ is 0.071 words per one hundred words in a CC (or four competition-related words per CC), comparable to the sample mean of 0.058 in Li et al. (2013). The mean value of $Miss$ indicates that 22.2% of our sample conference calls contain earnings that have missed the consensus forecast. The mean value for $Uncertain$ is 0.836 words per one hundred words in the CC, which corresponds to an average of around 72 uncertain words in a CC. As a benchmark, the mean value for $Uncertain$ reported in Loughran and McDonald (2013) for their sample of S-1 filings is 1.41 words per one hundred words. Our mean value for $Qualitative$ indicates that, on average, 80.7% of the sentences in our CC are qualitative. The median number of business segments for our sample firms ($\#Segments$) is two (the natural log of which is 0.693). The mean forecasting experience of our sample analysts is eight years, and average percentage of stars in them is 22.4%.

[Insert Table 5 here]

The results of estimating Eq. (2) are reported in Table 5, Panel A. We find that the coefficients on the interaction terms $Tone\_Discovery * Competition$, $Tone\_Discovery * LitigRisk$ and $Tone\_Discovery * Miss$ are positive and significant (at least at the 10% level), supporting our prediction that investors place a greater value on the analyst information discovery role when managers have greater incentives to withhold relevant information during conference calls – that is, when firms face higher proprietary cost or litigation risk, or deliver bad news in the earnings conference calls. Moreover, consistent with our prediction that investors put a greater value on the analyst information interpretation role when the processing cost of the disclosure is higher, we find that the coefficients on the interaction terms $Tone\_Interpret * Uncertain$ and $Tone\_Interpret * Qualitative$ are positive and significant (at the 5% level) and that the coefficient on $Tone\_Interpret * Miss$ is negative and significant (at the 5% level). That is, the analyst information interpretation role provides more value when managers' statements are more uncertain and qualitative, and less value when managers deliver bad news in the conference call. The coefficient on the interaction term between $Tone\_Interpret$ and $\#Segments$, however, is insignificant, probably because the number of segments is a noisy measure of operations complexity for S&P 500 firms (more than half of sample firms have either one or two segments). The coefficient estimates on the interaction terms between tones and EPS_Surp are insignificant, suggesting that the impact of earnings performance on analyst value is non-linear and driven by the occurrence of bad news.

We find negative (positive) and significant coefficients on the interaction term between $Tone\_Discovery$ ($Tone\_Interpret$) and $Expr$, suggesting that less experienced analysts trigger greater market reaction with discovery, while more experienced analysts trigger greater market reaction with interpretation. This result is consistent with the finding in Soltes (2014) that less experienced analysts have more private interaction with management, which is an important source of analyst discovery. Soltes (2014) argues that less experienced analysts are not as familiar with the economics and institutional features of the industries and firms they cover (consistent with their interpretation being less informative), and thus they compensate for this deficiency in experience by creating opportunities to gain additional information about the firms through private interaction with managers. Consistent with the mixed evidence in the literature, we do not find that investors react to $Star$ analysts' discovery or interpretation roles differently.

In Panel B of Table 5, we estimate the regression model of Eq. (1) for subsamples where managers' incentive to withhold information or the processing cost of the conference calls is particularly high or low. We examine investors' relative value of each role in these subsamples by comparing the magnitude of the estimated coefficients on $Tone\_Discovery$ and $Tone\_Interpret$. The F-test results are reported in Panel B of Table 5. In the subsample with $Competition$ in the top (bottom) decile, the coefficient of $Tone\_Discovery$ is not statistically different from (significantly smaller than) that of $Tone\_Interpret$; a similar pattern is observed in the subsample with $LitigRisk$ in the top and bottom deciles. The F-tests show a statistically greater coefficient on $Tone\_Intepret$ than that on $Tone\_Discovery$ in the subsamples with $Uncertain$ in both the top and bottom deciles, and in the subsample with $Qualitative$ in the top decile. The coefficients on $Tone\_Interpret$ and $Tone\_Discovery$ are not statistically different from each other for the subsample with $Qualitative$ in the bottom decile. Combined, these statistics suggest that investors put a greater value on the analyst information interpretation role immediately following earnings conference calls when managers' incentive to withhold information is low, and when processing cost is high. The value of analyst discovery becomes as important as that of analyst interpretation when managers' have strong incentives to withhold information or when the amount of qualitative statements in CC is very low.

4.3. Analysts' response to investors' information demands

To further investigate the interplay between the information in analyst research reports and the closely preceding corporate disclosure, we examine the relation between analyst efforts in each information role and the economic determinants identified in the previous section. We measure the effort spent on the information discovery role as the proportion of discussion in analyst research reports devoted to topics that receive little or no mention by managers in the CC (denoted as $Discovery$). Because investors value analyst information discovery more when managers are more likely to withhold information, that is, when firms face higher proprietary cost, higher litigation risk, or deliver bad news in the conference call, we predict that $Discovery$ increases with these economic determinants.

Because investors value analyst information interpretation more when conference calls contain a greater amount of uncertain and qualitative/subjective language, we posit that in these situations, analysts expend more effort to clarify managers' statements. That is, analysts likely transform management's original statements from the CC into a more meaningful narrative, which should manifest itself as different word usage from that of managers ($NewLanguage$). To measure this construct, we calculate the average difference between the word vectors of AR and CC for the CC topics that are also discussed in analyst reports (i.e., the average of one minus the cosine similarity between these vectors).[23]

As additional control variables, we include the number of analyst questions during the Q&A session ($\#Questions$, measured as the natural log of one plus the number of questions raised by analysts in the Q&A session). Because analysts likely request managers to clarify some statements during the Q&A

[23] Word vector of topic k in a document, $(w_{1k}, w_{2k}, \ldots w_{Nk})$ contains the frequency of all N word in the discussion of topic k in the document (N is the total number of unique words in the corpus). Cosine similarity is computed as the dot product of the two vectors normalized by their vector length and captures the textual similarity between two vectors of an inner product space using the cosine angle between them. Two vectors with the same orientation (i.e., two exact same or proportional topic vectors) have a cosine similarity of one; two orthogonal vectors have a similarity of zero. Cosine similarity is widely used in textual analysis research to compare narratives (see the review of Loughran and McDonald 2016). Internet Appendix II provides two illustrative examples from excerpts of conference calls transcripts and analyst reports, with high and low levels of $NewLanguage$, respectively.

session, analysts' questions might reduce the need for further clarification in prompt reports. In addition, we control for the magnitude of the earnings news using the absolute value of the earnings surprise (ABS_EPS_Surp). Finally, we control for the length of the combined prompt analyst reports (AR_Length), because Brown and Tucker (2011) find that measures based on cosine similarity are positively correlated with document length.

Descriptive statistics of the variables included in this test are reported in Table 3. The average $NewLanguage$ level of 0.54 is consistent with the existence of analysts' interpretation using their own language (this variable is bounded within [0, 1]). In our sample, analysts ask 26 questions, on average, during the Q&A session of the call (mean value of $\#Questions$ is 3). The mean (median) length of the combined prompt analyst reports (AR_Length) is 360 (320) sentences, reflected across an average of nine reports ($\#Analysts$).

[Insert Table 6 here]

Table 6 reports the regression results for the cross-sectional determinants of analyst efforts. The dependent variables in columns (1) and (2) are $Discovery$ and $NewLanguage$, respectively. The positive and significant (at least at the 5% level) coefficients on the proprietary cost measure ($Competition$), litigation risk ($LitigRisk$), and bad performance ($Miss$) (reported in column 1) are consistent with our prediction that analysts increase their efforts in information discovery when managers have greater incentives to withhold relevant information during conference calls. Column (2) of Table 6 shows that $Uncertain$ and $Qualitative$ are positive and significant at the 1% level, which also supports our prediction that analysts increase their interpretation efforts when the conference call is more difficult to process.[24] One interesting finding of note is that in column (2), the results in the regression of

[24] An alternative explanation of this result is that the processing cost of CC reflects a difficult-to-understand business environment, and such an environment naturally demands a larger vocabulary to describe, which results in dissimilar language among *any* information preparers who attempt to describe it. To investigate the validity of this explanation, we conduct a placebo test related to language differences among analysts. In this test, we randomly divided AR into two groups and re-run the $NewLanguage$ regression by replacing its dependent variable to the language difference between the two groups of analyst reports. The results are reported in Internet Appendix Table IA5. This table shows that the higher processing costs of CC do not explain the language differences among analysts, which is inconsistent with the alternative explanation.

$NewLanguage$ yield a significantly negative coefficient for $\#Questions$, suggesting that analysts embark on their information roles during the Q&A session of the earnings conference calls by asking questions; this involvement, in turn, preempts some efforts on the interpretation they exhibit in their prompt reports. It is also consistent with evidence shown by Matsumoto et al. (2011) that the information content of earnings conference calls increases with analyst involvement. Finally, the coefficients on $Expr$ and $Star$ are insignificant except for the one on $Expr$ in the $Discovery$ regression. This is probably because the effort allocation between discovery and interpretation is mostly driven by investors' information demand and less by analyst traits. That is, a star analyst can provide more discovery for one firm but more interpretation for another, depending on the firm's characteristics or the corporate disclosure.[25]

Overall, the findings in Table 6 indicate that analysts' efforts spent on information discovery and interpretation reflect their prompt responses to information demands from investors, which ultimately are driven by the characteristics of the firms and managerial disclosures.

4.4. Do investors value analysts' efforts?

Having established that investors respond to analysts' information interpretation and discovery and that analysts, in turn, respond to the demands for this information, we next examine the effect of analyst efforts on the value of the different types of discussions in AR and CC using the following model:

$$\begin{aligned} CAR[0,1] = {} & \alpha_1 Tone\_Discovery + \alpha_2 Tone\_Discovery * Discovery + \alpha_3 Discovery \\ & + \beta_1 Tone\_Interpret + \beta_2 Tone\_Interpret * (1 - Discovery) \\ & + \beta_3 Tone\_Interpret * (1 - Discovery) * NewLanguage + \gamma_1 Tone\_CC \\ & + \gamma_2 Tone\_CC * NewLanguage + \gamma_3 NewLanguage + \gamma_4 Miss \\ & * NewLanguage + \gamma_5 EPS\_Surp * NewLanguage + Controls. \end{aligned} \tag{3}$$

Eq. (3) expands the regression model of Eq. (1) by including the interaction terms between the tone of the discussion and the measures of analysts' efforts (i.e., the proportion of discovery and interpretation in AR, $Discovery$ and $(1 - Discovery)$, respectively, and the extent of new language used by analysts to interpret the CC topics, $NewLanguage$). If investors value analyst efforts spent on each role, we

[25] We repeat the analysis in Table 6 at the *individual* analyst report level and tabulate the results in Internet Appendix Table IA6. At the analyst level, analyst experience and star status are statistically insignificant.

expect the coefficients on these interaction terms to have positive signs. We also include the interaction term, $Tone\_CC * NewLanguage$, to examine whether the manner of analysts' interpretation affects the market reaction to managers' discussion. We do not provide a predicted sign for the interaction term, because the extent of using dissimilar language by analysts can have opposing effects on how the market reacts to managers' disclosures: On the one hand, when analysts use the same or similar language as managers, they provide a confirming value by enhancing management statements' trustworthiness (i.e., the market reacts more to management disclosures when analysts use a more similar language), which suggests a negative predicted sign; on the other hand, analyst new language provides clarification and helps investors understand managers' discussions (i.e., the market reacts more to management disclosures when analysts use a more different language), which suggests a positive predicted sign. Finally, we include the interaction terms $Miss * NewLanguage$ and $EPS\_Surp * NewLanguage$ to examine whether the market reacts to earnings news more intensely because analysts adopt new language to interpret corporate disclosure.

[Insert Table 7 here]

The results of estimating Eq. (3) are reported in Table 7. We find that value of each role increases with its proportion of the analyst report, as evidenced by the positive and significant (at the 1% level) coefficients on $Tone\_Discovery * Discovery$ and $Tone\_Interpret * (1 - Discovery)$. We also find that the coefficient on $Tone\_Interpret\ * (1 - Discovery) * NewLanguage$ is positive and significant, indicating that investors additionally value analysts' interpretation, given its length when analysts use more of their own language to discuss the CC topics. The estimated coefficient on $Tone\_CC * NewLanguage$ is significant and positive, suggesting that, on average, the clarification effect of analysts' interpretation on managerial disclosure dominates its confirming effect. We do not find, however, that market reaction to earnings news intensifies with the level of $NewLanguage$, probably because analysts use new language to clarify the qualitative information released by managers but not the quantitative signals, such as earnings surprises. Overall, the evidence in Table 7 indicates that investors value analysts' efforts expended on their information discovery and interpretation roles and that the market finds managers' disclosures more informative when analysts spend more efforts to clarify these disclosures.

4.5. Does analyst confirmation provide value to investors?

Prior studies on media (e.g., Miller 2006; Drake et al. 2014) distinguish the media's role of creating new information from its role of disseminating information.[26] In this section, we investigate the possibility that analyst research reports provide value to investors without discovering new information or interpreting corporate disclosure using new languages.

To do so, we design a test that identifies the parts of the discussion in analyst reports that simply provide confirmation to managers' discussions. Empirically, we employ the Pearson's chi-square test for each CC topic in each CC-AR pair to test whether the words used by managers and analysts to discuss a given topic are statistically different. We classify the analyst interpretation of a CC topic as using similar language when the difference in the distribution of words used to discuss this topic by analysts and managers is not statistically significant.[27] Defined in this way, we find that the interpretation using similar language constitutes 23% of a prompt report, on average, whereas interpretation using new language constitutes 46% (the remaining 31% is in discovery).

We employ the following regression model to investigate whether investors consider all three types of information valuable:

$$CAR[0,1] = \alpha_1 Tone\_Discovery + \beta_1 Tone\_NewLanguage + \beta_2 Tone\_SimilarLanguage + \gamma_1 Tone\_CC + Controls; \quad (4)$$

[Insert Table 8 here]

The results of estimating the above regression are reported in Table 8. We find positive and significant coefficients for all tone variables, i.e., $Tone\_Discovery$, $Tone\_NewLanguage$, and

---

[26] In a similar vein, Clement et al. (2003) show that management earnings forecasts that confirm market expectation provide value to investors and reduce the uncertainty about future earnings.

[27] That is, if the Pearson's chi-square test fails to reject the homogeneity between AR and CC with respect to their word distributions in this topic at the 10% level.

$Tone\_SimilarLanguage$, consistent with the usefulness of all types of analyst discussions. As one might expect, the magnitude of the positive and significant (at the 1% level) coefficient on $Tone\_SimilarLanguage$ is significantly smaller than that of $Tone\_Discovery$ and $Tone\_NewLanguage$ (both F-tests significant at the 1% levels). More importantly, the positive and significant coefficient on $Tone\_SimilarLanguage$ indicates the confirming value provided by analysts. That is, by selectively repeating CC topics, analysts attract and direct investors' limited attention to what is important from managers – confirming these topics' usefulness. Moreover, repeating CC topics likely enhances the reliability of the statements of managers, who may suffer from agency problems, thus confirming these statements' validity.[28]

5. Conclusion

An overwhelming proportion of analyst reports are issued immediately following important corporate disclosure events. Despite the vast literature on analysts, we know surprisingly little about how analysts serve the information intermediary roles in this narrow window. We fill the gap in the literature by examining the information content of analyst textual reports, in comparison to information in the preceding corporate disclosure, and whether their efforts, as well as their value, are driven by the characteristics of corporate disclosures.

We use algorithmic analyses of the topics discussed in the textual data of the conference calls and analyst reports to develop novel measures of analysts' information roles. We find that, on average, 31% of an analyst prompt report discusses exclusive topics not referred to in the conference call, which we consider as analysts serving an information discovery role. In the remaining 69% of the discussion in a prompt report, analysts discuss conference call topics, which we consider as the information interpretation role. We show that both discovery and interpretation trigger economically significant market reactions

[28] Consistent with this intuition, when asked whether analysts sometimes simply confirm what managers say, the All-Star analyst we interviewed replied: "*My experience is that sometimes analysts selectively pick up managers' comments to repeat. In many cases it is because he/she believes certain topics are more interesting to the market or have meaningful impact on earnings. In the extreme case of 'parroting', analysts use it to show investors that their thinking is in line with management.*"

beyond the associated earnings news and conference call discussions, suggesting that analyst information roles provide value. To understand the sources of their value, we show that investors rely on analyst information discovery more when managers have stronger incentives to withhold information during the conference calls and rely on analyst information interpretation more when the processing cost of the conference call information is higher. Analyst effort to discuss new topics and their effort to use their own language to clarify managers' statements suggest they offer prompt responses to investors' information demands. Finally, we show that within the 69% of the discussion in prompt reports where analysts interpret conference call topics, 23% of such discussion does not entail a different vocabulary than that used by managers in the conference call. That is, analysts sometimes confirm managers' statements. Interestingly, we find that investors value such confirmation, albeit to a lesser extent than analyst discovery or interpretation using the analyst's own language. This finding is consistent with analysts providing confirming value to the topics' relevance and validity by selectively repeating management statements.

Our study advances the literature by contributing to the understanding of the different information roles that analysts play, as well as the interplay between their information roles and corporate disclosures. We also make a contribution by explicitly quantifying the thematic content of analyst research reports and contrasting it with managers' discussions during earnings conference calls. Our study provides insight into how to use topic modeling to significantly expand the application of textual analysis to incorporate financial disclosures beyond an understanding of "*how* texts are being said" to a broader understanding of "*what* is being said" in these texts.

Finally, topic modeling has the potential to be used in a variety of research settings as a way to reduce large amounts of textual data into a manageable and conceptually interpretable set of topics. These topics can be used to address a variety of questions, including the cross-sectional and temporal variation in topic discussions in regulatory filings (e.g., Dyer et al. 2016), the characteristics (e.g., breadth) of firm disclosure on social media or in management guidance, the information content of corporate filings issued by firms relative to that of reports issued by information intermediaries (e.g., analysts, credit rating agencies, or business press), or detection of financial misreporting (e.g., Brown et al. 2016).

References

Atkins DC, Rubin TN, Steyvers M, Doeden MA, Baucom BR, Christensen A (2012) Topic models: A novel method for modeling couple and family text data. *J. of Family Psych.* 26(5):816-827.

Ball C, Hoberg G, Maksimovic V (2013) Disclosure and firm separation: A Text-based examination. Working paper, Maryland University.

Bao Y, Datta A (2014) Simultaneously discovering and quantifying risk types from textual risk disclosures. *Management Sci.* 60(6):1371-1391.

Beyer A, Cohen DA, Lys TZ, Walther BR (2010) The financial reporting environment: Review of the recent literature. *J. Accounting Econom.* 50(2-3):296-343.

Bhojraj S, Lee CMC, Oler DK (2003) What's my line? A comparison of industry classification schemes for capital market research. *J. Accounting Res.* 41(5):745-774.

Blei DM, Ng AY, Jordan MI (2003) Latent dirichlet allocation. *J. Machine Learn. Res*. 3(Jan):993–1022.

Boni L, Womack KL (2006) Analysts, industries, and price momentum. *J. Financial Quant. Anal.* 41(1):85-109.

Bradley D, Clarke J, Lee S, Ornthanalai C (2014) Are analysts' recommendations informative? Intraday evidence on the impact of time stamp delays. *J. Finance* 69(2):645-673.

Bradshaw MT (2011) Analysts' forecasts: what do we know after decades of work? Working paper. Boston College. http://ssrn.com/abstract=1880339.

Brochet F, Naranjo PL, Yu G (2016) The capital market consequences of language barriers in the conference calls of non-U.S. firms. *The Accounting Rev*. 91(4): 1023-1049.

Brown NC, Crowley RM, Elliott WB (2016) What are you saying? Using topic to detect financial misreporting. Working paper, University of Delaware.

Brown SV, Tucker JW (2011) Large-sample evidence on firms' year-over-year MD&A modifications. *J. Accounting Res.* 49(2):309-346.

Bushee BJ, Core JE, Guay W, Hamm SJ (2010) The role of the business press as an information intermediary. *J. Accounting Res.* 48(1):1-19.

Chen X, Cheng Q, Lo K (2010) On the relationship between analyst reports and corporate disclosures: Exploring the roles of information discovery and interpretation. *J. Accounting Econom.* 49(3):206-226.

Chen S, Matsumoto D, Rajgopal S (2011) Is silence golden? An empirical analysis of firms that stop giving quarterly earnings guidance. *J. Accounting Econom.* 51(1):134-150.

Clement BM (1999) Analyst forecast accuracy: Do ability, resources, and portfolio complexity matter? *J. Accounting Econom.* 27(3):285-303.

Clement M, Frankel R, Miller J (2003) Confirming management earnings forecasts, earnings uncertainty, and stock returns. *J. of Accounting Res*. 41(4): 653-679.

Davis AK, Ge W, Matsumoto D, Zhang JL (2015) The effect of manager-specific optimism on the tone of earnings conference calls. *Rev. of Accounting Studies*, 20(2):639-673.

Drake SM, Guest MN, Twedt, JB (2014) The media and mispricing: The role of the business press in the pricing of accounting information. *The Accounting Rev.* 89(5):1673-1701.

Dye RA (1986) Proprietary and nonproprietary disclosures. *J. of Business* 59(2):331-366.

Dye RA (2001) An evaluation of "essays on disclosure" and the disclosure literature in accounting. *J. Accounting Econom.* 32(1-3):181-235.

Dyer T, Lang M, Stice-Lawrence L (2016) The evolution of 10-K textual disclosure: evidence from latent dirichlet allocation. Working paper, University of North Carolina at Chapel Hill.

Epstein LG, Schneider M (2008) Ambiguity, information quality, and asset pricing. *J. Finance* 63(1):197-228.

Emery RD, Li X (2009) Are the Wall Street analyst rankings popularity contests? *J. Financial Quant. Anal.* 44(2):411-437.

Field L, Lowry M, Shu S (2005) Does disclosure deter or trigger litigation? *J. Accounting Econom.* 39(3):487-507.

Frankel R, Johnson M, Skinner DJ (1999) An empirical examination of conference calls as a voluntary disclosure medium. *J. Accounting Res.* 37(1):133-150.

Frankel R, Kothari SP, Weber J (2006) Determinants of the informativeness of analyst research. *J. Accounting Econom.* 41(1-2):29-54.

Freeman RN, Tse SY (1992) A nonlinear model of security price responses to unexpected earnings. *J. Accounting Res.* 30(2):185-209.

Graham JR, Harvey CR, Rajgopal S (2005) The economic implications of corporate financial reporting. *J. Accounting Econom.* 40(1):3-73.

Healy PM, Palepu KG (2001) Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical disclosure literature. *J. Accounting Econom.* 31(1-3):405-440.

Heylighen F, Dewaele JM (2002) Variation in the contextuality of language: An empirical measure. *Foundations of Sci*. 7(3):293-340

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

Huang A, Zang A, Zheng R (2014) Evidence on the information content of text in analyst reports. *The Accounting Rev*. 89(6):2151-2180.

Hutton PA, Miller SG, Skinner JD (2003) The role of supplementary statements with management earnings forecasts. *J. Accounting Res.* 41(5):867-890.

Ivković Z, Jegadeesh N (2004) The timing and value of forecast and recommendation revisions. *J. Financial Econom.* 73(3):433-463.

Johnson MF, Kasznik R, Nelson KK (2001) The impact of securities litigation reform on the disclosure of Forward-Looking information by high technology firms. *J. Accounting Res.* 39(2):297-327.

Kadan O, Madureira L, Wang R, Zach T (2012) Analysts' industry expertise*, J. Accounting Econom.* (54):95-120.

Kaplan S, Vakili K (2015) The double-edged sword of recombination in breakthrough innovation. *Strategic Management J.* 36(10):1435-1457.

Lang M, Lundholm R (1993) Cross-sectional determinants of analyst ratings of corporate disclosures. *J. Accounting Res*. 31(2):246-271.

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

Lehavy R, Li F, Merkley K (2011) The effect of annual report readability on analyst following and the properties of their earnings forecasts. *The Accounting Rev.* 86(3):1087-1115.

Lopez T, Rees L (2002) The effect of beating and missing analysts' forecasts on the information content of unexpected earnings. *J. of Accounting, Auditing & Finance* 17(2):155-184.

Levelt, WJM (1989) *Speaking: from intention to articulation*. Cambridge, Massachusetts: The MIT Press.

Li F, Lundholm R, Minnis M (2013) A measure of competition based on 10-K filings. *J. Accounting Res.* 51(2):399-436.

Li X, Ramesh K, Shen M, Wu S (2015) Do Analyst Stock Recommendations Piggyback on Recent Corporate News? An Analysis of Regular-Hour and After-Hours Revisions. *J. Accounting Res.* 53(4):821-861.

Livnat J, Zhang Y (2012) Information interpretation or information discovery: Which role of analysts do investors value more? *Rev. Financial Stud.* 17(3):612-641.

Loh RK, Stulz RM (2011) When are analyst recommendation changes influential? *Rev. Financial Stud.* 24(2):593-627.

Loughran T, McDonald B (2013) IPO first-day returns, offer price revisions, volatility, and form S-1 language. *J. Financial Econom.* 109(2):307-326.

Loughran T, McDonald B (2016) Textual Analysis in Accounting and Finance: A Survey. *J. Accounting Res.* Forthcoming.

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

Miller GS (2002) Earnings performance and discretionary disclosure. *J. Accounting Res.* 40(1):173-204.

Miller GS (2006) The press as a watchdog for accounting fraud. *J. Accounting Res.* 44(5):1001-1033.

Mikhail BM, Walther RB, Willis HR (1997) Do security analysts improve their performance with experience? *J. Accounting Res.* 35(1):131-157.

Ohlson J (1991) The theory of value and earnings, and an introduction to the Ball-Brown analysis. *Contemporary Accounting Res.* 8(1):1-19.

Porter MF (1980) An algorithm for suffix stripping. *Program*, 14(3):130-137

Quinn KM, Monroe BL, Colaresi M, Crespin MH, Radev DR (2010) How to analyze political attention with minimal assumptions and costs. *Amer. J. Political Sci.* 54(1):209–228.

Ramage D, Rosen E, Chuang J, Manning C, McFarland DA (2009) Topic modeling for the social sciences. In *NIPS Workshop on Applications for Topic Models: Text and Beyond*. (Whistler, Canada).

Ramnath S, Rock S, Shane P (2008) The financial analyst forecast literature: A taxonomy with suggestions for future research. *Internat. J. of Forecasting* 24(1):34-75.

Rogers JL, Van Buskirk A (2009) Shareholder litigation and changes in disclosure behavior. *J. Accounting Econom.* 47(1):136-156.

Rosen-Zvi M, Griffiths T, Steyvers M, and Smyth P (2004) The author-topic model for authors and documents. *In Proc. of the 20th conf. on Uncertainty in artificial intelligence*. (UAI, Banff, Canada), 487-494

Soltes E (2014) Private interaction between firm management and sell-side analysts. *J. Accounting Res*. 52 (1):245-272.

Steyvers M, Griffiths T (2006) Probabilistic topic models. *Handbook of Latent Semantic Analysis.*

Stickel SE (1992) Reputation and performance among security analysts*. J. Finance* 47(5):1811-1836

Verrecchia RE (1983) Discretionary disclosure. *J. Accounting Econom.* 5(1):179-194.

Verrecchia RE (2001) Essays on disclosure. *J. Accounting Econom.* 32(1-3):97-180.

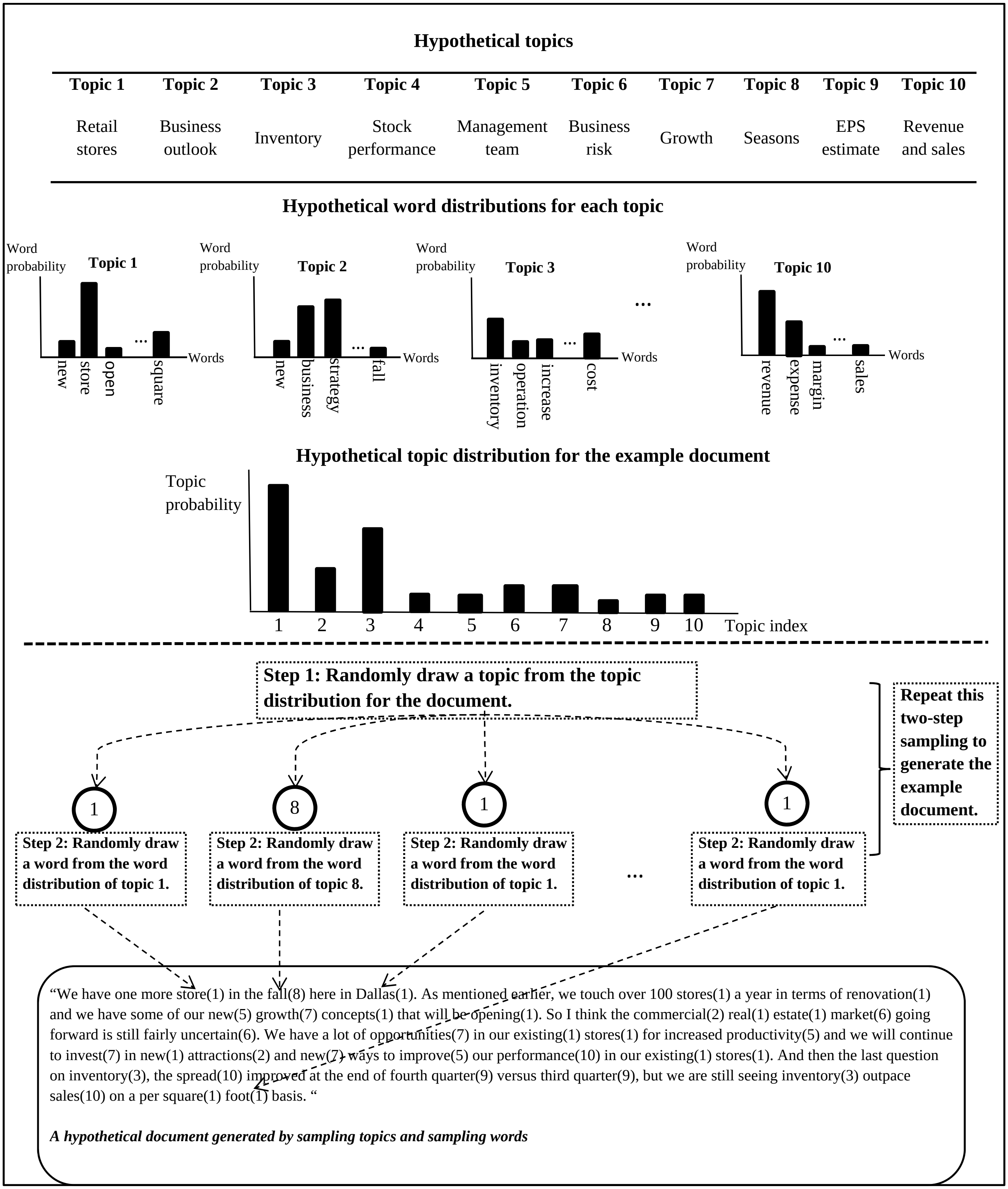
Appendix I

Intuition of Latent Dirichlet Allocation (LDA)

LDA assumes that a document is generated in two steps. First, a topic is randomly drawn based on the topic distribution of this document; next, a word is randomly drawn based on the word distribution of the topic selected in the previous step. Repeating these two steps word by word generates a complete document. This basic idea of LDA simulates how a human writes a document: He/she has a plan to discuss certain topics first and then selects words to explain each topic.

Appendix Figure 1 illustrates a hypothetical example of this two-step document generation process. First, we assume that a collection of documents discusses ten different topics in total, including topics of retail stores, business outlook, inventory, etc. Each document in this collection discusses some of the ten topics. Put in statistical language, each document in this collection relates to the ten topics probabilistically: High-probability topics are discussed more heavily in the document than low-probability topics, and a zero-probability topic is one that is not discussed in the document at all. Next, each topic is related to words probabilistically. High-probability words in a topic mean that they are more likely to appear in this topic. For example, the four words with the highest probabilities in Topic 1 (Stores) are: “new,” “store,” “open,” and “square.” Note that a word can have high probabilities in more than one topic. The word “new,” for example, has high probabilities in Topic 1 (Stores), Topic 5 (Management), and Topic 7 (Growth and Expansion), because this word is used frequently in all three topics. Some words can relate to a topic with zero probability if they never appear in that topic.

To compare the topics of analyst reports and conference calls, we want to find the topic distribution of each document and the word distribution of each topic. The LDA algorithm achieves this by fitting a generative model to the actual words in documents and finding the best set of latent variables that describe the two sets of distributions. This is similar to how a maximum likelihood estimation method maximizes the “agreement” of the model with the observed data. Technical details of LDA are discussed in Internet Appendix I.



Appendix Figure 1. An illustration of how LDA assumes a document is generated

Appendix II

Applying LDA to Conference Call Transcripts and Analyst Reports

A. The corpus

Earnings conference call transcripts are obtained from Thomson Reuter's StreetEvents database, and analyst reports are obtained from Thomson Reuter's Investext database. Our corpus is composed of 18,607 earnings conference call transcripts and 476,633 analyst reports for S&P 500 firms from 2003 to 2012, all of which are used in the LDA to obtain the best representation of topics. We conduct the LDA analyses separately for each industry, because many topics are industry-specific. We use four-digit Standard & Poor's Global Industry Classification Standard (GICS) to identify industries. This classification is widely adopted by brokerages and analysts as their industry classification system and is superior to other industry classification schemes, such as SIC codes and NAICS codes, in identifying firms with their industry peers (Kadan et al. 2012; Boni and Womack 2006; Bhojraj et al. 2003).

B. Preprocessing of textual documents

The raw files of conference call transcripts that we obtained from Thomson Reuter's StreetEvents database are in XML (i.e., extensible markup language) format. We develop a Java program to parse the XML files and extract useful information, including company names, tickers, dates and times of the calls, participants and their titles, and textual dialogues. The downloaded analyst reports are in PDF (i.e., portable document format) format. We first use Adobe Acrobat to convert them into TXT (i.e., text file) format and then develop a Java program to extract the report issuance dates, analyst names, broker names, and the reports' textual content.

The next several steps prepare the textual data for the LDA analysis. First, we exclude venue-specific language or time-invariant information that is not associated with any economically relevant topics. For conference calls, we exclude narratives from operators, greeting words used by various speakers, and the safe harbor statements typically read by Investor Relation Officers. For analyst reports, we follow Huang et al. (2014) and remove the tables, graphs, and "brokerage disclosures." Brokerage disclosures contain explanations of the stock-rating system, disclosures regarding conflicts of interest, analyst certifications, disclaimers, glossaries, and descriptions of the brokerage firm. Second, we convert all words into lower case, remove all non-English characters (e.g., punctuation and numbers), and convert all plural nouns into their singular forms.[29]

Third, we remove high frequency functional words, also referred to as stop words. There are two benefits from removing stop words. First, these words, such as "a," "of," "the," "this," and "is," are extremely frequent, but convey little economic meaning. Second, stop words contain many deictic words (that is, words that cannot be fully understood without additional contextual information) that constitute the major difference between oral language and written language.[30] Removing them helps mitigate the concern that the difference between conference calls and analyst reports are due to the difference in language style.

Fourth, we follow Heylighen and Dewaele (2002) and delete more contextual words that distinguish oral language from written language. Heylighen and Dewaele (2002) develop a much simpler but coarser way

[29] We do not perform "stemming" (i.e., replacing words with their root form), because it is too aggressive for financial text, where words with the same stem are often not synonyms. For example, standard stemming would convert "marketing" into "market," "accounting" into "account," "investment" into "invest," and both "operating" and "operation" into "oper" (Porter 1980).

[30] Oral language is more context dependent than written language (Heylighen and Dewaele 2002; Levelt 1989; Lee 2016). Levelt (1989), for example, distinguishes four types of deixis: person (e.g., "we," "him," "my"), place (e.g., "here," "those," "upstairs"), time (e.g., "now," "later," "yesterday"), and discourse (e.g., "therefore," "yes," "however"), including exclamations or interjections (e.g., "oh," "well," "ok"). These deictic words are categorized as stop words and removed.

to identify contextual words by using grammatical categories, which include pronouns, adverbs, and interjections. Because pronouns and interjections are already excluded as stop words in the third step, in this step, we essentially remove high-frequency adverbs, such as “very,” “thus,” “really,” “actually,” and “basically.”[31]

Moreover, because financial and technical terminology is common in conference calls and analyst reports, we convert high-frequency phrases that constitute specific financial/technical terms into one word (or its common abbreviation if there is one). For example, “target price” is converted into “target-price,” “balance sheet” into “balance-sheet,” “earnings per share” into “EPS,” and “cost of goods sold” into “COGS.” This step helps retain financial/technical terms’ accurate meanings and disambiguate polysemous words and abbreviations.

Lastly, we remove S&P 500 company names and tickers to prevent LDA from identifying companies as topics.

All of these preprocessing steps enhance the interpretability of the topics identified and reduce the computational burden of the LDA model. After these steps, we have approximately 303 million words in our corpus.

C. Determining the parameters of the LDA algorithm

The LDA algorithm we use is the “Stanford Topic Modeling Toolbox” developed by the Stanford Natural Language Processing Group (Ramage et al. 2009). It requires the researcher to set three parameters for the assumed statistical model including the total number of topics in the entire collection of documents, and α and β, which determines how smooth the topic and word distributions are, respectively (please see Internet Appendix I for a detailed explanation).

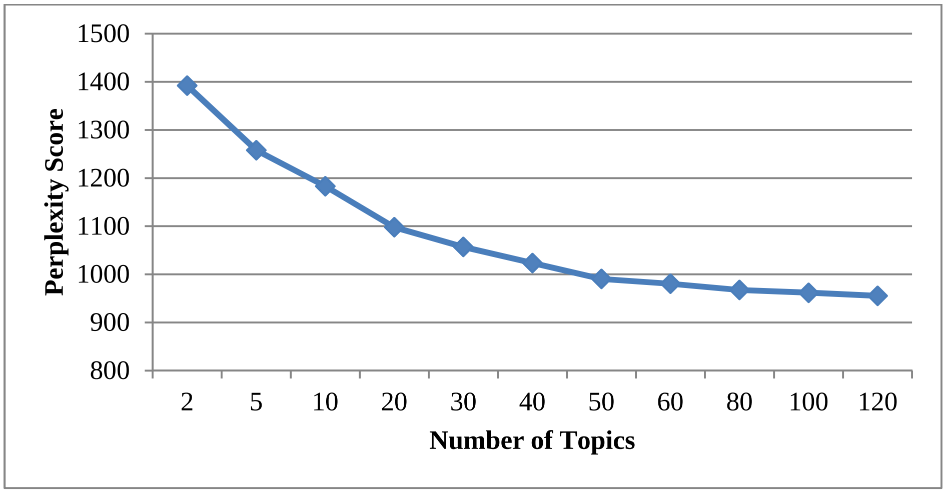
The number of topics of the model affects the interpretability of the results. Setting the number too low can result in topics that are too broad and ambiguous. Conversely, setting the number too high may introduce economically meaningless topics. To select the optimal number of topics, we follow the computational linguistic literature and calculate the *Perplexity Score* of the LDA model based on different numbers of topics (Blei et al., 2003; Rosen-Zvi et al. 2004). The perplexity score measures the ability of an LDA model estimated on a subset of documents (training data) to predict the word choices in the remaining documents (testing data). It is defined as the exponential of the negative normalized predictive likelihood under the model. Accordingly, the perplexity score is monotonically decreasing in the likelihood of observing the testing data, given the model estimated from the training data. A lower perplexity score indicates that the model has better generalization performance. Formally, for a testing data (D_{test}) with M documents, the perplexity score is equal to:

$$perplexity\ score(D_{test}) = \exp\left\{-\frac{\sum_{d=1}^{M} log p(w_d)}{\sum_{d=1}^{M} N_d}\right\},$$

where N_d is the number of words in document d; w_d is a vector of all the words in document d; and $p(w_d)$ is the probability of observing the word vector w_d in document d given the LDA model estimated from the training data.

[31] Note that we only remove stop words and contextual words for LDA because they are either meaningless (stop words) or just reflect linguistic styles (contextual words) and do not help identify economically meaningful topics. For tone classification using the naïve Bayes classification, we follow the procedures described in Huang et al. (2014) and do not remove contextual words. Other variables based on text, i.e., $Uncertain$, $Qualitative$, and $Competition$, are calculated using original sentences.

Following the literature (Blei et al., 2003; Rosen-Zvi et al., 2004), we compute and plot the perplexity scores of the LDA model for different numbers of topics, ranging from 2 to 120. As Appendix Figure 2 shows, the perplexity score improves with the number of topics, but the improvement is marginally decreasing. The improvement diminishes significantly once the number of topics exceeds 60. Therefore, we choose 60 as the number of topics in our corpus.[32]



Appendix Figure 2. Perplexity of LDA model for different numbers of topics

The choice of the values of α and β depends on the specific textual genre, the number of topics and the vocabulary size. We choose values of 0.1 and 0.01 for α and β, respectively, based on the recommended values in the literature (Steyvers and Griffiths 2006; Kaplan and Vakili 2015).

D. Constructing a topic vector of a document

The output from the LDA algorithm is a topic-word probability matrix Φ in which an element, p_{ik}, is word w_i's probability in topic k. With the LDA output, we construct the topic vector (T_d) of a document d using the following procedures: First, for each sentence in d, we sum the probabilities of its words in each topic to obtain this sentence's probabilities in all topics. Next, we assign each sentence to the topic in which it has the highest probability.[33] Lastly, we calculate the fraction of document d that is dedicated to topic k (S_{dk}) as document d's proportion of sentences assigned to topic k. Formally, the topic vector of document d is defined as

$$Topic\ vector\ of\ document\ d = T_d = (S_{d1}, S_{d2}, \ldots S_{d60}).$$

[32] The suitable number of topics depends on the specific samples employed by different studies. For example, Ball et al. (2013) use 100 topics for MD&A text; Quinn et al. (2010) use 42 topics for political text; Atkins et al. (2012) use 100 topics for couples-therapy transcripts. In addition to using the perplexity score, we also compare LDA outputs manually based on 30, 60, and 100 topics. Based on our comparison, we conclude that the LDA results with 60 topics outperform other specifications in terms of its ability to identify economically important topics without generating too many uninterpretable topics.

[33] In a sensitivity test, we assign each sentence to three topics with the highest probabilities. Our empirical results remain qualitatively similar.

Appendix III

Variable Definitions

Variable Name	Definition
$Discovery$	The number of sentences labeled by LDA as non-CC topics in AR scaled by the total number of sentences in AR. CC topics are the topics in which the discussion length exceeds 2% of the CC.
$NewLanguage$	The average of one minus within-topic cosine word similarity between CC and AR in the CC topics. The within-topic cosine word similarity between CC and AR for a given topic k is calculated as $\frac{\sum_{j=1}^{N}(w_{jk} \cdot v_{jk})}{\sqrt{\sum_{j=1}^{N}(w_{jk})^2} \cdot \sqrt{\sum_{j=1}^{N}(v_{jk})^2}}$, where, w_{jk} is word j's frequency in the discussion of topic k in AR; v_{jk} is word j's frequency in the discussion of topic k in CC; N is the total number of unique words in CC and AR. CC topics are the topics in which the discussion length exceeds 2% of the CC.
$Tone\_Discovery$, $Tone\_Interpret$, $Tone\_NewLanguage$, $Tone\_SimilarLanguage$	$Tone\_Discovery$ and $Tone\_Interpret$ are the textual opinions of the sentences labeled by LDA as non-CC and CC topics in AR, respectively. $Tone\_NewLanguage$ ($Tone\_SimilarLanguage$) is the textual opinion of the sentences labeled by LDA as CC topics in AR using new (similar) language. A topic is defined as using new language if the Pearson's chi-square test for the homogeneity between AR and CC with respect to their word distributions in this topic is significant at the 10% level. The textual opinion of the sentences is calculated as the percentage of positive sentences minus the percentage of negative sentences as classified by the naïve Bayes approach (Huang et al. 2014).
$CAR[0,1]$	The cumulative abnormal return over the [0, 1] window relative to the conference call date, winsorized at the top and bottom 1%, where the abnormal return is calculated as the raw return minus the buy-and-hold return on the NYSE/Amex/Nasdaq value-weighted market index.
$Tone\_CC$	The textual opinion of the sentences labeled by LDA as CC topics in CC minus that of the same firm's previous CC. The textual opinion of the sentences is calculated as the percentage of positive sentences minus the percentage of negative sentences as classified by the naïve Bayes approach (Huang et al. 2014).
EPS_Surp	Earnings surprise, calculated as the actual EPS minus the last consensus EPS forecast before the earnings announcement date, both from I/B/E/S, scaled by the stock price 10 days prior to the earnings announcement date, winsorized at the top and bottom 1%.
$Miss$	An indicator variable that equals one if the actual EPS is less than the last consensus EPS forecast before the earnings announcement date, both from I/B/E/S, and zero otherwise.
$Prior\_CAR$	The cumulative 10-day abnormal returns ending two days before the conference call, winsorized at the top and bottom 1%, where abnormal return

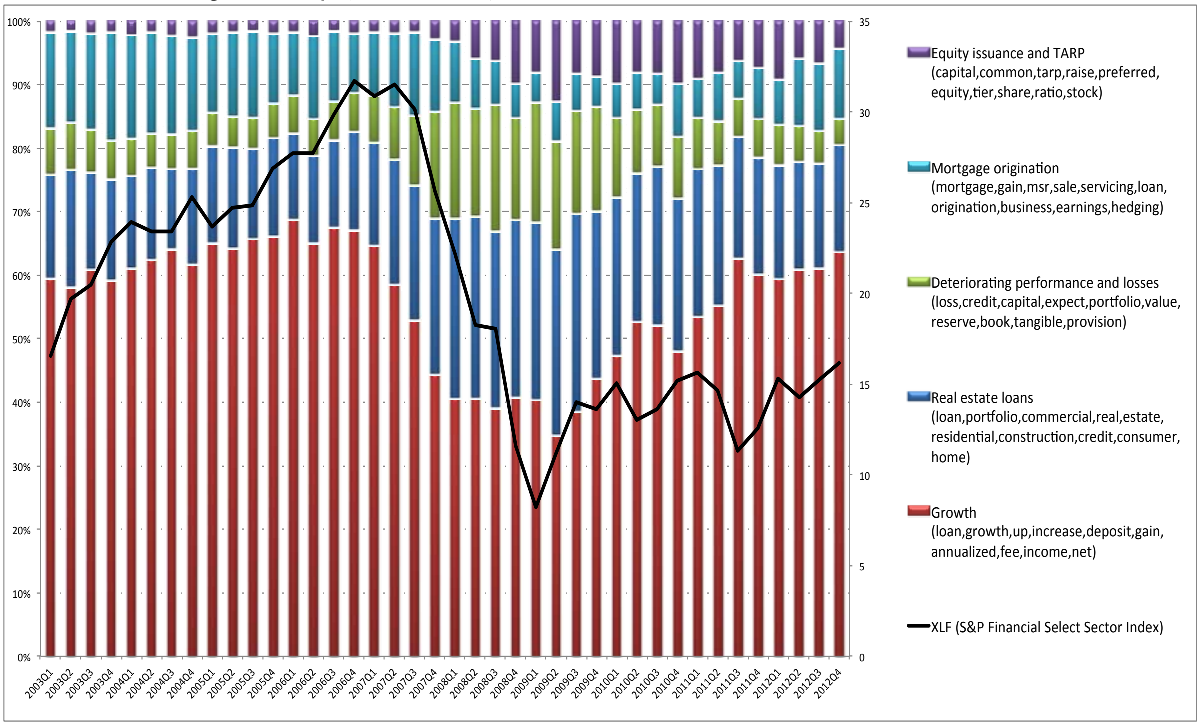
	is calculated as the raw return minus the buy-and-hold return on the NYSE/Amex/Nasdaq value-weighted market index.
$Size$	The natural log of the market value of equity of the firm (*CSHOQ*×*PRCCQ*) at the end of the quarter prior to the conference call, winsorized at the top and bottom 1%.
$BtoM$	The book value of equity (*CEQ*) scaled by the market value of equity (*CSHOQ*×*PRCCQ*) of the firm at the end of the quarter prior to the conference call, winsorized at the top and bottom 1%.
$\#Analysts$	The number of analyst reports issued on the day of or the day following the conference call.
$\#Questions$	The natural log of one plus the number of questions raised by analysts in the conference call's Q&A session.
$Competition$	Percentage of competition-related words in CC in the firm's previous conference call. Following Li et al. (2013), competition related words include "competition," "competitor," "competitive," "compete," and "competing." We include words with an "s" appended and do not count words in phrases that contain negation, such as "less competitive," and "few competitors."
$LitigRisk$	The standard deviation of the monthly return of the firm in the 12 months prior to the conference call, winsorized at the top and bottom 1%.
$Expr$	The average experience of analysts who issue reports on the day of or the day following the conference call. Experience is measured as the number of years since the analyst first issued a forecast in I/B/E/S.
$Star$	The number of analysts who are ranked as Institutional Investor All-Star analysts, scaled by the total number of analysts who issued reports on the day of or the day following the conference call.
$Uncertain$	The number of words in CC that are in the Uncertainty word list created by Loughran and McDonald (2013), scaled by the total number of words in CC.
$Qualitative$	One minus the percentage of sentences that contain "$" or "%."
$\#Segments$	The natural log of a firm's number of segments.
ABS_EPS_Surp	The absolute value of the earnings surprise, calculated as the absolute value of the difference between the actual EPS and the last consensus EPS forecast before the earnings announcement date, both from I/B/E/S, scaled by the stock price 10 days prior to the earnings announcement date, winsorized at the top 2%.
AR_Length	The number of sentences in analyst reports issued on the day of or the day following the conference call.

Figure 1

Temporal Variation in the Distribution of Key Topics

This figure presents the relative weights in the five topics with the highest variability in the banking and telecommunication industries, along with their respective sector indices (Financial Sector SPDR – XLF and iShares US Telecommunications – IYZ respectively) in our sample period of 2003-2012.

Panel A: Banking industry (GICS 4010)



Panel B: Telecommunication industry (GICS 5010)

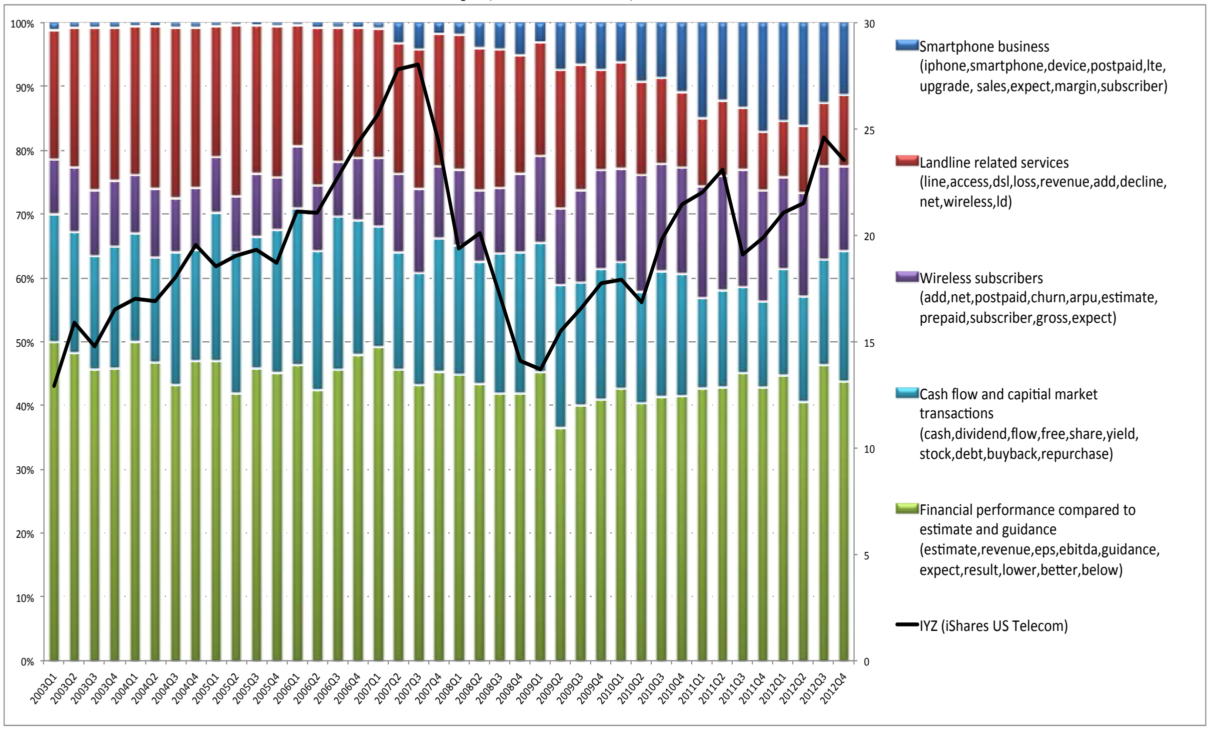


Table 1

Highest Probability Words in the Top Ten Topics of Two Large Industries

This table reports the top 20 words in each of the top ten topics and our inferred topic labels for two of the five largest industries in terms of the total number of conference calls in our sample.

Topic Label	Top 20 Words
	Capital Goods (GICS 2010)
Comparing financial performance with expectation	margin, estimate, guidance, EPS, expect, consensus, operating, revenue, lower, consensus, sales, expectation, below, per-share, segment, management, forecast, in-line, beat, outlook
Sales and revenue	sales, improve, operating, margin, price, profit, revenue, estimate, decline, volume, share, segment, operating-income, improve, off-set, rise, lower, cost, currency, earnings
Growth	growth, revenue-growth, organic, strong, sales, digit, acquisition, business, rate, expect, EPS-growth, strength, grow, line, margin, solid, core, single, segment, guidance
Business outlook	business, expansion, good, term, margin, positive, looking, rate, big, difficult, customer, forward, guidance, market, pressure, area, line, opportunity, issue, new
Financial outlook	revenue, growth, operating, margin, segment, increase, business, expect, year-over-year, forecast, result, acquisition, higher, estimate, decline, compare, income, report, strong, EPS
Valuation model	multiple, price, stock, earnings, target-price, valuation, estimate, DCF (discount cash flow), cycle, risk, growth, EPS, current, price-to-earnings, group, relative, view, investor, peak, upside
Defense contracts	system, program, defense, contract, space, service, budget, electronic, aircraft, information, ship, missile, government, technology, international, sales, air, support, navy, DOD (Department of Defense)
Cash flows and financing	cash, flow, free, share, capital, net, dividend, debt, balance, repurchase, increase, strong, sheet, margin, stock, free-cash-flow, earnings, growth, program, cash-flow
Raw materials and input price	cost, price, increase, material, pricing, margin, higher, raw, volume, expect, incremental, inflation, impact, commodity, product, off-set, operating, steel, inventory, benefit
Geographic segments	market, growth, China, Europe, global, emerging, America, demand, region, Asia, India, investment, country, north, economy, expect, middle, economic, European, east
	Health Care Equipment & Services (3510)
Growth	growth, margin, revenue, expect, operating, business, rate, gross, digit, market, expansion, improvement, EPS, organic, mix, increase, drive, single, grow, new
Earnings guidance and expectations	estimate, EPS, guidance, share, expect, range, management, result, expectation, consensus, growth, earnings, impact, per-share, lower, new, in-line, revenue, report, stock
Geographic segments	sales, division, currency, constant, growth, report, expect, divisional, product, FX (foreign exchange), gross, rate, Europe, business, impact, margin, international, foreign, tax, Japan
Income statement items	income, net, revenue, expense, operating, after-tax, EPS, margin, gross, interest, share, cost, profit, rate, SG&A, dilute, pre-tax, amortization, item, adjust
Valuation	estimate, EPS, target-price, multiple, price, share, risk, growth, valuation, price-to-earnings, stock, earnings, rating, base, trade, industry, group, forward, premium, peer
Medical cost	enrollment, MLR (medical loss rate), commercial, cost, trend, medical, earnings, share, Medicare, expect, ratio, membership, higher, prior, SG&A, live, projection, increase, report, premium
Business outlook and opportunities	business, positive, term, good, market, future, guidance, impact, looking, forward, rate, new, product, performance, opportunity, better, call, cost, issue, start
Cash flow and financing	cash, debt, flow, share, net, asset, capital, cash-flow, liability, repurchase, balance, equity, note, investment, free, free-cash-flow, stock, dividend, sheet, expense
Medicare and Medicaid	Medicare, plan, commercial, member, Medicaid, advantage, health, premium, care, benefit, cost, membership, group, enrollment, business, contract, government, risk, Tricare, individual
Drug trial	announce, disease, drug, product, category, treatment, trial, patient, update, system, new, agreement, Humira (a drug name), study, clinical, program, hub, pharmaceutical, administration, phase

Table 2

Sample Selection

Panel A presents the sample selection procedures for the earnings conference calls. Panel B presents the sample selection procedures for the analyst reports. Revision reports are analyst reports that contain a revision in at least one of the analysts' quantitative measures (earnings forecast, stock recommendation, and target price).

Panel A: Sample selection – earnings conference calls

Earnings conference calls of S&P 500 firms in the 2003-2012 period	18,607
Less earnings conference calls not on days [0, +1] relative to the earnings announcement dates	371
Less earnings conference calls without accompanying analyst reports	486
Earnings conference calls on days [0, +1] relative to the earnings announcement dates, with accompanying analyst reports	17,750

Panel B: Sample selection – analyst reports

	All Reports	Revision Reports
Analyst reports issued for S&P 500 firms in the 2003-2012 period	476,633	220,723
Less analyst reports not within [0, +1] relative to the earnings conference call dates	313,316	114,034
Less analyst reports issued before the start time of the earnings conference calls	4,107	4,107
Number of analyst reports issued on days [0, +1] after the earnings conference calls (denoted, $\boldsymbol{AR}$)	159,210	102,582
$\boldsymbol{AR}$ as a percentage of all analyst reports issued for S&P 500 firms	33.4%	46.5%

Table 3

Descriptive Statistics

This table reports the summary statistics for the variables used in the empirical analyses. Variable definitions are provided in Appendix III.

Variables:	# of obs.	Mean	Median	Std	Q1	Q3
$Discovery$	17,749	0.314	0.303	0.104	0.238	0.379
$NewLanguage$	17,749	0.540	0.538	0.082	0.483	0.593
$Tone\_Discovery$	17,731	0.188	0.193	0.159	0.094	0.289
$Tone\_Interpret$	17,749	0.217	0.223	0.178	0.100	0.343
$Tone\_NewLanguage$	17,480	0.213	0.221	0.214	0.083	0.355
$Tone\_SimilarLanguage$	17,694	0.224	0.231	0.205	0.100	0.357
Determinants of Discovery and Interpretation:						
$Competition$ (%)	17,131	0.071	0.052	0.073	0.019	0.101
$LitigRisk$	17,724	0.086	0.074	0.048	0.053	0.104
$Uncertain$ (%)	17,749	0.836	0.811	0.265	0.651	0.986
$Qualitative$ (%)	17,749	80.699	79.412	7.961	74.038	84.615
$\#Segments$	17,749	0.751	0.693	0.747	0.000	1.386
$Miss$	17,632	0.222	0.000	0.416	0.000	0.000
EPS_Surp	17,622	0.001	0.001	0.005	0.000	0.002
$Expr$	17,327	8.016	7.859	2.423	6.424	9.400
$Star$	17,749	0.224	0.182	0.235	0.000	0.347
Other Variables:						
$CAR[0,1]$	17,733	0.000	0.000	0.057	-0.030	0.030
$Tone\_CC$	17,064	0.000	0.002	0.097	-0.061	0.063
$\#Questions$	17,749	3.045	3.135	0.577	2.833	3.367
ABS_EPS_Surp	17,622	0.002	0.001	0.004	0.000	0.003
AR_Length	17,749	359.909	320.000	227.985	187.000	487.000
$Prior\_CAR$	17,699	0.003	0.002	0.049	-0.024	0.028
$Size$	17,723	9.339	9.233	1.083	8.594	9.952
$BtoM$	17,745	0.468	0.393	0.326	0.248	0.609
$\#Analysts$	17,749	8.954	8.000	4.967	5.000	12.000

Table 4

Investors' Reaction to Analyst Information Discovery and Information Interpretation

This table reports the coefficient estimates and *t*-statistics from estimating Eq. (1). All variables are defined in Appendix III. *t*-stats based on standard errors clustered at the firm and year levels are displayed in parentheses below the coefficient estimates. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively, using two-tailed tests.

	Dependent Variable $CAR[0,1]$
$Tone\_Discovery$	0.041***
	(10.9)
$Tone\_Interpret$	0.061***
	(16.7)
$Tone\_CC$	0.039***
	(7.9)
EPS_Surp	2.828***
	(10.6)
$Miss$	-0.013***
	(-8.9)
$EPS\_Surp * Miss$	-2.503***
	(-7.8)
$Prior\_CAR$	-0.060***
	(-4.8)
$Size$	-0.000
	(-0.5)
$BtoM$	0.016***
	(8.3)
$\#Analysts$	-0.000***
	(-3.6)
$Intercept$	-0.021***
	(-3.9)
Year Fixed Effect	Yes
Observations	16,923
Adjusted R^2	0.138

Table 5

Determinants and Relative Value of Analyst Information Discovery and Information Interpretation

Panel A: Determinants of the value of analyst information discovery and interpretation roles

This panel reports the coefficient estimates and *t*-statistics from estimating Eq. (2). All variables are defined in Appendix III. *t*-stats based on standard errors clustered at the firm and year levels are displayed in parentheses on the right of the coefficient estimates. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively, using two-tailed tests.

	Dependent Variable $CAR[0,1]$	
$Tone\_Discovery$	0.005	(0.6)
$Tone\_Discovery * Competition$	0.002**	(2.2)
$Tone\_Discovery * LitigRisk$	0.008***	(7.4)
$Tone\_Discovery * Miss$	0.013*	(1.7)
$Tone\_Discovery * EPS\_Surp$	-0.111	(-0.1)
$Tone\_Discovery * Expr$	-0.002**	(-2.0)
$Tone\_Discovery * Star$	0.000	(0.2)
$Tone\_Interpret$	0.070***	(6.4)
$Tone\_Interpret * Uncertain$	0.038**	(2.0)
$Tone\_Interpret * Qualitative$	0.001**	(2.3)
$Tone\_Interpret * \#Segments$	-0.001	(-1.3)
$Tone\_Interpret * Miss$	-0.014**	(-2.1)
$Tone\_Interpret * EPS\_Surp$	-0.030	(-0.0)
$Tone\_Interpret * Expr$	0.001**	(2.2)
$Tone\_Interpret * Star$	-0.001	(-1.2)
$Tone\_CC$	0.038***	(7.7)
$Competition$	-0.000	(-1.6)
$LitigRisk$	-0.001***	(-4.2)
$Uncertain$	0.001**	(2.1)
$Qualitative$	-0.000	(-0.9)
$\#Segments$	-0.000	(-0.3)
$Miss$	-0.013***	(-6.3)
EPS_Surp	2.893***	(8.0)
$EPS\_Surp * Miss$	-2.619***	(-6.4)
$Expr$	0.000	(1.3)
$Star$	0.000	(1.2)
$Prior\_CAR$	-0.063***	(-5.0)
$Size$	-0.000	(-0.1)
$BtoM$	0.017***	(8.9)
$\#Analysts$	-0.000***	(-4.0)
$Intercept$	-0.021***	(-3.1)
Year Fixed Effect	Yes	
Observations	16,615	
Adjusted R^2	0.146	

Panel B: The relative value of information discovery and information interpretation

This panel reports the coefficient estimates and t-statistics of the main variables from estimating Eq. (1). In Columns (1) and (2), (3) and (4), (5) and (6), and (7) and (8), we separately estimate Eq. (1) for sub-samples of conference calls in the bottom and top deciles in terms of $Competition$, $LitigRisk$, $Uncertain$, and $Qualitative$, respectively. All variables are defined in Appendix III. t-stats based on standard errors clustered at the firm and year levels are displayed in parentheses below the coefficient estimates. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively, using two-tailed tests.

Dependent Variable	$CAR[0,1]$							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Partition Variables	$Competition$		$LitigRisk$		$Uncertain$		$Qualitative$	
	Bottom Decile	Top Decile	Bottom Decile	Top Decile	Bottom Decile	Top Decile	Bottom Decile	Top Decile
$Tone\_Discovery$	0.031***	0.054***	0.024***	0.075***	0.016*	0.026***	0.044***	0.039***
	(4.8)	(5.1)	(4.3)	(4.9)	(1.8)	(2.7)	(4.5)	(3.3)
$Tone\_Interpret$	0.054***	0.057***	0.041***	0.071***	0.074***	0.071***	0.045***	0.056***
	(8.3)	(5.9)	(7.7)	(4.6)	(7.4)	(7.0)	(4.5)	(4.9)
F-test of equality between $Tone\_Discovery$ and $Tone\_Interpret$	5.14**	0.03	3.19*	0.03	12.73***	7.16***	0.01	0.66*
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,162	1,681	1,718	1,612	1,710	1,671	1,698	1,652
Adjusted R^2	0.111	0.171	0.159	0.128	0.210	0.108	0.132	0.162

Table 6

Determinants of Analyst Information Discovery and Interpretation Roles

This table reports the coefficient estimates and t-statistics from OLS regressions of $Discovery$ and $NewLanguage$ on their determinants and control variables. Variable definitions are provided in Appendix III. t-stats based on standard errors clustered at the firm and year levels are displayed in parentheses below the coefficient estimates. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively, using two-tailed tests.

	Dependent Variables	
	$Discovery$	$NewLanguage$
	(1)	(2)
$Competition$	0.027**	
	(2.4)	
$LitigRisk$	0.096***	
	(4.0)	
$Uncertain$		0.021***
		(4.1)
$Qualitative$		0.001***
		(7.7)
$\#Segments$		0.000
		(0.2)
$Miss$	0.005***	0.000
	(3.2)	(0.0)
ABS_EPS_Surp	-0.103	0.165
	(-0.4)	(0.6)
$Expr$	-0.002*	-0.002
	(-1.9)	(-0.2)
$Star$	0.005	0.010
	(0.8)	(0.6)
$\#Questions$	0.001	-0.009***
	(0.6)	(-3.4)
$Size$	0.005***	-0.000
	(3.3)	(-0.3)
$BtoM$	-0.004	-0.001
	(-0.8)	(-0.4)
AR_Length	-0.000	-0.000***
	(-0.6)	(-17.6)
$\#Analysts$	0.000	-0.001**
	(0.4)	(-2.4)
$Intercept$	0.262***	0.536***
	(14.5)	(23.3)
Industry and Year Fixed Effects	Yes	Yes
Observations	16,704	17,291
Adjusted R^2	0.190	0.415

Table 7

Investors' Value of Analyst Information Discovery and Interpretation Efforts

This table reports the coefficient estimates and *t*-statistics from estimating Eq. (3). All variables are defined in Appendix III. *t*-stats based on standard errors clustered at the firm and year levels are displayed in parentheses to the right of the coefficient estimates. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively, using two-tailed tests.

	Dependent Variable $CAR[0,1]$	
$Tone\_Discovery$	-0.008	(-1.0)
$Tone\_Discovery * Discovery$	0.173***	(6.5)
$Discovery$	-0.014*	(-1.9)
$Tone\_Interpret$	-0.016	(-0.9)
$Tone\_Interpret * (1 - Discovery)$	0.209***	(6.3)
$Tone\_Interpret * (1 - Discovery) * NewLanguage$	0.184***	(4.1)
$Tone\_CC$	-0.016	(-0.5)
$Tone\_CC * NewLanguage$	0.101*	(1.8)
$NewLanguage$	0.029**	(2.4)
$Miss * NewLanguage$	-0.008	(-0.5)
$EPS\_Surp * NewLanguage$	-0.761	(-0.4)
$Miss$	-0.008	(-0.9)
EPS_Surp	1.683	(1.4)
$Prior\_CAR$	-0.060***	(-4.7)
$Size$	-0.001*	(-1.7)
$BtoM$	0.022***	(11.6)
$\#Analysts$	-0.000***	(-2.8)
$Intercept$	-0.028***	(-3.0)
Year Fixed Effect	Yes	
Observations	16,923	
Adjusted R^2	0.135	

Table 8

Investors' Reaction to Analyst Information Discovery, New Language, and Confirmation

This table reports the coefficient estimates and *t*-statistics from estimating Eq. (4). All variables are defined in Appendix III. *t*-stats based on standard errors clustered at the firm and year levels are displayed in parentheses below the coefficient estimates. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively, using two-tailed tests.

	Dependent Variable $CAR[0,1]$
$Tone\_Discovery$	0.049***
	(12.2)
$Tone\_NewLanguage$	0.036***
	(12.5)
$Tone\_SimilarLanguage$	0.013***
	(5.1)
$Tone\_CC$	0.043***
	(8.5)
$Miss$	-0.013***
	(-9.2)
EPS_Surp	1.297***
	(8.0)
$Prior\_CAR$	-0.057***
	(-4.4)
$Size$	-0.001
	(-1.3)
$BtoM$	0.021***
	(10.9)
$\#Analysts$	-0.000***
	(-3.4)
$Intercept$	-0.016***
	(-3.0)
Year Fixed Effect	Yes
Observations	16,627
Adjusted R^2	0.130