6-2020

Stock market information and security prices

Haoyuan LI

Singapore Management University, haoyuan.li.2015@pbs.smu.edu.sg

STOCK MARKET INFORMATION AND SECURITY PRICES

HAOYUAN LI

SINGAPORE MANAGEMENT UNIVERSITY

2020

Stock Market Information and Security Prices

Haoyuan Li

Submitted to Lee Kong Chian School of Business
in partial fulfilment of the requirements for the
Degree of Doctor of Philosophy in Business (Finance)

Dissertation Committee Members:

Roger K. Loh (Chair)
Associate Professor of Finance
Singapore Management University

Weikai Li
Assistant Professor of Finance
Singapore Management University

Rong Wang
Associate Professor of Finance
Singapore Management University

Gennaro Bernile
Associate Professor of Finance
University of Miami

SINGAPORE MANAGEMENT UNIVERSITY
2020

I hereby declare that this PhD dissertation is my original work
and it has been written by me in its entirety.
I have duly acknowledged all the sources of information
which have been used in this dissertation.

This PhD dissertation has also not been submitted for any degree
in any university previously.

Li Haoyuan , 2020 Apr 28

Haoyuan Li
28 April 2020

Abstract

Chapter 1: Analyst report content and stock market anomalies

A series of recent papers document that security analyst recommendations tend to contradict stock-mispricing signals. This seems at odds with the large prior literature on the investment value of analyst recommendations. What justifications do analysts make when they write reports on mispriced stocks? I use the latest techniques in machine learning and textual analysis to categorize the qualitative information in a large sample of analyst reports. I find that report content can be intuitively classified into five categories or topics: 1) Growth, 2) Earnings, 3) New developments, 4) Management transactions, and 5) Conviction. I then relate the frequency of each topic and the tone surrounding the topic to stock-anomaly mispricing signals. I find that although analysts are incorrectly optimistic about overvalued stocks in general, reports on new developments and management transactions have investment value after controlling for the predictive power of the mispricing signals. For undervalued stocks, while analysts are on average incorrectly pessimistic, reports on growth, new developments, and management transactions have investment value. Overall, this paper helps to understand how analysts provide value in their reports even when the report ratings appear to contradict well-known signals of mispricing.

Chapter 2: The information cycle and return seasonality (with Roger Loh)

Heston and Sadka (2008) find that the monthly cross-sectional returns of stocks depend on their historical same calendar-month returns. We propose an information-cycle explanation for this seasonality anomaly—that firms' seasonal release of information coincide with higher returns during months with such dissolution of information uncertainty, and lower returns during months with no information releases. Using earnings announcements and changes in implied volatility as proxies for scheduled information releases, we find that seasonal winners in information-release months and seasonal losers in non-information release months indeed drive the seasonality anomaly. Our evidence shows that scheduled firm-level information releases can give rise to the appearance of an anomalous seasonal pattern when stock returns are in fact responding to information uncertainty.

Chapter 3: Managerial and analyst horizons during conference calls

It is alleged that public-firm managers face short-term pressures from investors. In this paper, I examine managers' tendency to talk about the short versus the long term by analyzing the language in quarterly analyst conference calls. Using the word embedding model, I determine whether conference calls focus on the short or long term. I find that when firms fail to meet analyst expectations, both managers and analysts focus on the short term rather than the long term. However, in macro bad times, analysts question managers about the short term rather than the long term, while managers maintain the same long term-short term balance whether in good or bad macro conditions. Finally, I show that firms whose conference call participants focus more on the long term have negative initial market reactions, but stock prices recover in the subsequent months. subsequent months. The results are consistent with Wall Street exerting excessive short-term pressures on public firm managers.

Contents

Acknowledgments

I would never have completed this work without help from many people. First and foremost, I sincerely thank my committee chair Roger Loh for his years of mentoring, advice, and encouragement. I have learned from him how to generate, develop, and defend research ideas. These skills are important not only for my PhD study but also for becoming a qualified scholar.

I especially thank Gennaro Bernile, who leads me to the research world. His energy and creativity encourage me to pursue exciting research. I appreciate his unconditional supports at the essential stages in my PhD life. I also thank my committee members, Rong Wang and Weikai Li, for many helpful discussions. Advice from Chishen Wei, Hao Liang, Jianfeng Hu, Dashan Huang, Jun Tu, and Ekkehart Boehmer is important for improving my PhD study.

Further, during the doctoral learning process, fellow students and friends help each other. I thank Li Guo. Though working on different research topics, we have learned from and inspired each other to explore ideas. I also thank Xin Zheng, Yijie Fei, Minrui Jiang, Han Chen, Ke Miao, Yiwei Yu, Yao Meng, Lin Sun, Sili Zhou, Yuan Zhuang, Shuyu Xue, Bin Li, and Jiangyuan Li. We have enjoyed the PhD life together at SMU.

I also thank my friend, Huan Zhao. We are lucky to know each other since the undergraduate. We keep learning and improving ourselves in different places. We have the common goal and will achieve it in the next ten years.

Finally, I thank Tingting for her love and support. I've been falling behind in the past years and hope to catch up in the future.

Chapter 1

Analyst report content and stock market anomalies

By Haoyuan Li

I thank Gennaro Bernile, Weikai Li, Roger Loh, Rong Wang, and Chishen Wei for comments and suggestions.

Abstract

A series of recent papers document that security analyst recommendations tend to contradict stock-mispricing signals. This seems at odds with the large prior literature on the investment value of analyst recommendations. What justifications do analysts make when they write reports on mispriced stocks? I use the latest techniques in machine learning and textual analysis to categorize the qualitative information in a large sample of analyst reports. I find that report content can be intuitively classified into five categories or topics: 1) Growth, 2) Earnings, 3) New developments, 4) Management transactions, and 5) Conviction. I then relate the frequency of each topic and the tone surrounding the topic to stock-anomaly mispricing signals. I find that although analysts are incorrectly optimistic about overvalued stocks in general, reports on new developments and management transactions have investment value after controlling for the predictive power of the mispricing signals. For undervalued stocks, while analysts are on average incorrectly pessimistic, reports on growth, new developments, and management transactions have investment value. Overall, this paper helps to understand how analysts provide value in their reports even when the report ratings appear to contradict well-known signals of mispricing.

Keywords: Qualitative information; textual analysis; market anomalies

1.1. Introduction

A large literature documents that analysts provide stock recommendations and earnings forecasts that have investment value to investors (see, e.g. Womack (1996) and Barber, Lehavy, McNichols, and Trueman (2001)). However, a recent series of studies show that analysts tend to have favorable recommendations on overvalued stocks and correspondingly unfavorable recommendations on undervalued stocks. In other words, analyst ratings appear to contradict the investment signals provided by a large set of well-known stock market market anomalies (Jegadeesh, Kim, Krische, and Lee (2004), Guo, Li, and Wei (2019), Grinblatt, Jostova, and Philipov (2018), Engelberg, McLean, and Pontiff (2019)). While there are some studies documenting the conflict of interests that analysts might face (e.g., Cowen, Groysberg, and Healy (2006), Malmendier and Shanthikumar (2014), and Chen and Matsumoto (2006)), and the behavioral biases that they might exhibit (e.g., Bondt and Thaler (1990), Hong and Kubik (2003)), these studies cannot explain why analysts would be systematically on the wrong side of stock-return anomaly signals when evaluating the investment prospects of stocks.

In this paper, I try to shed light this puzzle by examining the types of justifications provided by analysts when they write reports on mispriced stocks. When analysts provide recommendations and earnings forecasts in their reports, they accompany it with written content to justify the outlook in the report. Bagnoli, Watts, and Zhang (2008) for example, emphasize that institutions find the actual content in the report to be much more important than the explicit rating and earnings forecast in the report. The report content might come from the analyst's effort in collecting/intepreting data, interviewing management, or attending analyst conferences/events (e.g., see Kirk and Markov (2016)). Depending on the types of information which analysts draw upon to support their decisions, the investment value of recommendations may be different. It would be useful if one could systematically categorize the information in analyst reports. This will also enable one to examine whether analysts use similar types of justifications when they issue recommendations in apparently the wrong direction on mispriced stocks.

In this study I employ recent techniques in textual analysis and machine learning to systematically categorize the content of thousands of analyst reports. Past research usually employs research assistants to manually read reports and categorize report information. For example, Asquith, Mikhail, and Au (2005) read 1,126 all-star analyst reports individually and hand-code the words in the reports into 14 categories/topics. The 14 categories are—revenue growth, earnings growth, new product introductions, new projects, cost efficiencies, expectations met, mergers and acquisitions, repurchase programs, industry climate, management, international operations, leverage, competition, and risk—and they measure analysts' opinions towards each of these topics. This traditional method of classifying reports while useful is a subjective exercise and hard to employ for a large sample of reports. I use a recent technique in Natural Language Processing (NLP), the Word2Vec method, to classify and categorize a large sample of analysts reports. This replicable technique can categorize the textual content of analyst reports into five intuitive categories, which I then relate to the investment value of recommendations in general, as well as to the consistency of recommendations with mispricing signals.

I download from Thomson Reuters a sample of 34,531 Morgan Stanley analyst reports on U.S. firms from 2007 to 2014. I focus on one large broker for two reasons. The first is to ensure that the report quality is similar across firms. Second, and more importantly, the broker uses the same report format (e.g. location of headers, summaries, estimates, etc.) for all its reports, which allows for more efficient parsing of the report content. I proxy for the main content of the report by parsing the report's first page, which contains the synopsis and summary recommendation. The final sample includes analyst reports written for 1,561 unique firms, an average of 783 firms per year, each firm having five reports per year, and each report having 15 sentences and 297 words on the first page.

I classify the content of this large sample of reports into different categories/topics. Unlike the typical approach where the topics are first pre-specified by the researcher, the step-by-step technique I use allows the data to speak for itself to surface the topics. First, the textual content is grouped based on the semantic and syntactic similarity between words in pairs using neural networks (utilizing the Word2Vec model). Second, based on this computed similarity score, I build 15 information categories based on the most similar words related to the top 15 occurring words. From these 15 groups, I select five groups whose constituent words together can account for most of the words in the reports, *and* whose constituent words are the least similar between categories. I find that that analyst report content can be classified into these five categories/topics: 1) Growth, 2) Earnings, 3) New developments, 4) Management transactions, and 5) Conviction. Armed with this categorization, I then construct two new report attributes for each category—the frequency that the words of that category appears in the report cover, and the tone of the phrases adjacent to the category's words. To measure tone, I use the Loughran and McDonald (2011) word dictionary.

The first step to understanding analysts' decision-making process is to study the link between analyst recommendations and the frequency and tone of each information category. Jegadeesh et al. (2004) show that analysts are more likely to recommend glamour stocks with positive momentum, high sales growth, and large capital expenditure. Using the new report content categories I have identified, I can now classify the justifications used to recommend glamour stocks. I find that analysts are more likely to make favorable recommendations on glamour stocks when they are optimistic about growth, new developments, and conviction. In contrast, optimism about earnings is not the justification used in reports on glamour stocks. For example, A one standard deviation increase in analyst tone about growth increases the probability of a buy recommendation by 8%, while a one standard deviation increase in earnings tone increases the probability by 3%. Huang, Zang, and Zheng (2014) show that the overall tone of analyst reports is positively related to analyst recommendations. In this paper, I find that the tone of information categories have incremental value beyond the overall tone of the report. The results are robust with a battery of controls, including size, momentum, book-to-market ratio, and year and industry fixed effects.

Next, I consider how analysts' tone in each report category is related to stock-anomaly mispricing signals. Engelberg et al. (2019) show that analyst recommendations are contradictary to contrarian anomalies (or MGMT cluster based on Stambaugh and Yuan (2016)), but they are consistent with the signals of momentum anomalies (or PERF cluster based on Stambaugh and Yuan (2016)).[1] I compare these two aggregate mispricing scores to the analyst recommendations in my sample and confirm that recommendations are indeed contradicting contrarian anomaly signals and consistent with momentum anomaly signals. Although analysts are incorrectly optimistic about MGMT-cluster-based overvalued stocks in general, reports on new developments and management transactions have investment value even after controlling for the predictive power of the mispricing signals. Further, analysts' overall tone in the reports partly explains their favorable recommendations of overvalued stocks. For MGMT-cluster-based undervalued stocks, while analysts are on average incorrectly pessimistic, reports on new developments, management transactions, and overall tone in reports have investment value. Regarding the PERF cluster, as analysts recommendations line up correctly with anomaly signals, there is no additional investment value from conditioning on most of the report categories, except growth and overall tone, after controlling for the investment value of the anomaly signals.

Third, I explore how the information content contained in the five report topics are incorporated into prices. Regression results show that investors respond to analyst recommendations and all of the five information categories immediately. After controlling the overall tone of the reports, report topics on earnings, new developments, and management transactions still have incremental investment value. What happens in the weeks after the release of the reports? There is some reversal to the buy and sell recommendations, which shows that investors tend to overreact to the recommendations from this brokerage house. In contrast, there is no return reversal to different category tones, which shows that investors can digest information content properly. The category tone with the most predictive power for future cumulative risk-adjusted returns is the earnings report category. Calendar-time portfolio strategies based on a report's earnings tone can earn hedged returns of 0.7% per month.

[1]Anomalies in MGMT cluster are mainly related to managerial actions, including net stock issuance, composite net equity issuance, accounting accruals, net operating assets, asset growth, and investment to assets. Anomalies in PERF cluster are related to firm performance, including gross profitability, return on assets, failure probability, and bankruptcy probability. I exclude momentum anomaly from PERF cluster due to the momentum crash from 2009 to 2013 documented by Daniel and Moskowitz (2016)

This study contributes in several ways to the literature. First, by uncovering what are the major topics in analyst reports, it helps investors to understand the types of information contained in different types of analysts reports issued on different types of firms. This helps in our understanding of why analysts appear to line up on the wrong sign of anomaly signals. I show that although analysts are wrong on average, but when these reports on mispriced stocks touch on certain topics, they still have investment value. Knowing how to segregate analyst reports according to the topics these reports focus on might be crucial for institutions who rely on these reports for their investment decisions. Second, this paper contributes to the literature aiming to understand how the market reacts to qualitative information. Most studies on textual analysis talk about tone collectively, e.g., they look at the *overall* opinion in a piece of news (Tetlock (2007), Tetlock, Saar-Tsechansky, and Macskassy (2008)), 10-K files (Loughran and McDonald (2011)), textual records of earnings conference calls (Druz, Petzev, Wagner, and Zeckhauser (2020)), or analyst reports (Huang et al. (2014)). In this paper, I document the incremental value of tones in *different* information categories conditional on analysts' overall tone. Third, like Cong, Liang, and Zhang (2019), the method in this paper is based on the latest techniques in machine learning and textual analysis and hence can be applied to other settings where a large set of qualitative information needs to be categorized.

The rest of the paper is organized as follows. Section 2 discusses the methodology, data, and descriptive statistics. Section 3 examines how analyst recommendations are related to the frequency and tone of information categories. Section 4 shows whether the tone of a particular category can explain why analysts recommend mispriced stocks. Section 5 shows how the market reacts to the information categories. Finally, Section 6 concludes.

1.2. Data and methodology

1.2.1. Sample

The sample includes all U.S. company reports issued by Morgan Stanley during the 2007–2014 period, which covers a large number of firms, and its reports have been used in past research, e.g., Joos et al. (2016) and Loh and Stulz (2018). Also, the reports have a consistent format to organize the information in the sample period, which makes the textual content extractable by machine. To capture the relevant and critical information, I extract report date, company ticker, stock recommendation (buy, hold, and sell), industry view (attractive, cautious, and in-line), and the relevant textual content from the first page of reports. In the process of extracting information, I exclude irrelevant textual content, such as brokerage disclosure, tables, figures, and captions, as they are more likely to contain less meaningful template language Loughran and McDonald (2011). I then match the reports with Compustat and CRSP to obtain the accounting information and returns based on firm Ticker parsed from reports. The sample includes matched 34,531 reports and 1,561 firms. On average, each firm has five reports per year, and each report contains 15 sentences and 297 words on the first page.

1.2.2. Methods

After extracting the textual content, I group the information into five categories using neutral networks (Word2Vec). Then I relate the frequency of each category and the tone surrounding the topic to stock-anomaly mispricing signals. The variable construction includes three steps: 1) using Word2Vec algorithm to categorize the textual content; 2) selecting representative and dissimilar information categories; 3) computing the frequency and tone towards each information category. I provide the detailed descriptions below.

(1) Categorize the textual content in analyst reports

The current literature on textual analysis usually categorizes information based on subjective judgments (Asquith et al. (2005), Li, Lundholm, and Minnis (2013), Birru, Gokkaya, and Liu (2019) [2] or predetermined word dictionaries (Gao, Ren, and Zhang (2020)) [3]. However, the existent methods have two limitations: 1) predetermined word dictionaries are general-purpose—they fail to capture the features, structures, and terminologies of specific-purposed documents; and 2) word-list based on subjective judgment is hard to generalize or replicate.

To address these two limitations, I adopt a word classification approach, Word2Vec, a two-layer neural network to accurately guess a word's meaning and establish a word's association with other words (Mikolov, Sutskever, Chen, Corrado, and Dean (2013)). This approach can make highly accurate guesses about words' meanings based on their past appearances given enough data and contexts [4]. Specifically, Word2Vec produces a hundred-dimension vector space based on a large corpus of texts. Word vectors are positioned in the vector space such that words sharing common contexts in the corpus are located close to each other in the space. Each unique word in the corpus is assigned to a corresponding vector in the space. Based on the vectors, the similarity score in pairs could be computed to represent how close the two words are related semantically and syntactically [5].

To allow the Word2Vec algorithm to identify the semantic and syntactic similarity between words fully, I used all Compustat-matched analyst reports (34,531 reports) issued in 2007–2014.

[2] For example, Asquith et al. (2005) subjectively categorize words in reports into 14 categories, e.g., revenues, new product introduction, cost efficiencies. Li et al. (2013) use the occurrences of "competition, competitor, competitive, compete, competing" in firm's 10-K filing as the measure of the intensity of the competition faced by the management. Birru et al. (2019) build a keyword list for quantitative modelling, including variants of "mispricing", "cross-sectional", "overvalued", and "undervalued".

[3] For example, Gao, Ren, and Zhang (2020) use University of South Florida Free Association Norms (USFFAN) to obtain the associated words of six stimulus words, "weather", "disaster", "pollution", "terror/terrorist", "holiday", and "sport".

[4] A detailed description of the method can be found at https://code.google.com/archive/p/word2vec/ and https://www.tensorflow.org/tutorials/representation/word2vec

[5] Also, LDA (Latent Dirichlet Allocation) is another widely-used statistical instrument in textual analysis. It can fulfill the word-categorizing process. But the target of this algorithm is to extract unobserved topics from a large sample of words (Blei et al. (2003)). It can attenuate the subjective concerns in the nature of the methods of word dictionary and subjective judgment, as it can extract the subjects of reports (Huang et al. (2018)). But this method does not fit my reserach question. The current paper does not study the process of generating topics. I mainly use Word2Vec to measure the similarity of words in pairs.

Following the standard-setting, for each word, I set a 150-dimension vector, which means using 150 dimensions to represent each word (token). The output of the Word2Vec is a vocabulary in which each item has a vector attached to it, which can be queried to detect relationships between words. Based on the word vector, I can query the relations between words (cosine similarity). Stop words and words with lower frequency (less than 100 times) are not included in the analysis.

With the trained model, I choose the top 15 occurring words as seed words and cluster the related words around them, which are listed in table 1. The word frequency is calculated by the times of the word being mentioned on the first page. As shown in the table, "growth" is the word that is the most frequently mentioned by analysts, which is mentioned two times in each report on average. The second frequent word is "eps".

Then I use two steps to cluster words based on seed words. First, I choose 20 most similar words as the associated words for each of the 15 top-occurring words. Second, for each of the associated words, I choose 20 most similar words. Finally, I get a word list that contains 400 words with repeated cases. For example, "growth" has an associated word "sales", while "sales" is also related to "growth". After dropping the repeated words, each group contains 217 words on average.

[Table 1 inserts here]

(2) Select the most representative and dissimilar categories

The second step is to find categories that are representative but dissimilar from each other. The 15 categories from the first step are not mutually exclusive to each other. For example, 6 of the top 20 occurring words in "eps" and "revenue" are the same, showing that 30% of information is co-covered by two categories. To choose the representative and dissimilar categories, I start with the category which is chosen based on the highest occurring word, "growth", among the 15 categories. Then for the rest categories, I define a dissimilar score to capture the extent to which the current category covers different information from all the previously selected categories. Specifically, I count how many words in the current category have been included in the previously selected categories using the following equation:

$$Dissimilar\ score = 1 - \frac{\#\ of\ words\ covered\ by\ selected\ group}{20} \tag{1.1}$$

If the dissimilar score for a category is larger than 0.85, it is selected as a representative and dissimilar category. Specifically, I follow the following process to categorize the information:

1) Select the category with the highest occurring seed word "growth".

2) Then for the second top-occurring word, "eps", calculate how much the "eps" category is dissimilar from the "growth" category. As three words in the "eps" category appear in the "growth" category, including "beat", "outlook", and "result", the dissimilar score of "eps" category is 0.85.

3) For the "guidance" category, 13 of 20 top-occurring words have been covered by the "growth" and "eps" category, the dissimilar score for the "guidance" category is 0.35. Since a big part of the information in the "guidance" category has been covered in the previously selected categories, it is dropped.

4) For the "new" category, only 1 of 20 words are covered by the "growth" and "eps" category. Its dissimilar score with the prior selected groups is 0.95. So, the "new" category is selected.

5) None of the top 20 words in the "company" category has been included in previously selected categories and therefore is selected.

5) Similarly, the dissimilar score for the "estimate", "line" and "revenue" is 0.20 ,0.65, and 0.65, and thus are dropped.

6) As the dissimilar score of the "believe" category is 0.90, it is selected.

The dissimilar scores of all the categories after the "believe" category are smaller than 0.85. Finally, I select five categories, which are seeded with the words "growth", "eps", "new", "company", and "believe", respectively. I label these five categories, respectively, as growth, earnings, new developments, management transactions, and conviction. In the rest of the paper, I use either the labels the seed word to describe the categories, e.g. I will refer to the fifth category as the "believe" or "conviction" category. Figure 1 plots the word cloud maps for the five categories.

[Figure 1 inserts here]

(3) Measure analysts' tone of information categories

Analysts' tone of a specific information category is based on sentiment-word lists compiled by Loughran and McDonald (2011). [6] Specifically, I determine the tone of each word by checking whether the seven words on its left side and the seven words on its right side (except when a comma or a period appears within seven words) are within LM word list. If there is more than one surrounding word in the LM word list, the closest one is selected to determine the tone of the word. Also, I take the approach in Hu and Liu (2004) to account for sentiment negation. If the word distance between a sentiment word and a negation word ("no", "not", "none", "neither", "never", "nobody", "*n't") is smaller than five, the positive or negative polarity of the word is changed to the opposite of its original polarity. Following this process, I allocate each one a sentiment level (positive, negative or neutral).

I construct two sets of variables, which are the frequency and the tone of each information categories. Frequency of a specific category is the times of the words in a given category that appeared in the first page. Tone of a specific category is the difference between the number of

[6] I use the 2017 version from https://sraf.nd.edu/textual-analysis/resources/

non-negative words and the number of negative words.

$$
\begin{aligned}
Tone\_of\_category_{j,t} &= \#\ of\ positive\ words_{j,t} + \#\ of\ neutral\ words_{j,t} - \#\ of\ negative\ words_{j,t} \\
&= \#\ of\ non\_negative\ words_{j,t} - \#\ of\ negative\ words_{j,t}
\end{aligned}
\tag{1.2}
$$

1.2.3. Variables and summary statistics

I calculate analysts' tone separately for each of the five information categories based on Equation (2). The variable means the net number of non-negative words in each category on the first page. The larger number means analysts talk more about a specific information in a more positive tone. As shown in Table 2, analysts talk more about a firm's growth and earnings using 6.5 and 7.5 words on average, showing that they are two critical factors in their decision-making process. Firms' new developments and management transactions are less mentioned in the reports. *Tone_Believe* means analysts, on average, use three non-negative words related to conviction and prediction.

[Table 2 inserts here]

I also compute analysts' overall opinion on the first page of analyst reports based on the difference between the number of positive and negative words based on the sentiment word lists compiled by Loughran and McDonald (2011). As shown in Table 2, the average tone is 0.73, which is consistent with analysts' well-known optimism. On average, the first page uses 297 words and 15 sentences. Appendix A provides a detailed variable definition.

Panel B reports the correlation between variables. Analysts' tones towards growth and earnings are more likely to be mentioned in the same context. Analysts' overall opinion on the first page is more cognate to growth and new project information.

1.3. Information content and analyst recommendations

In a report, analysts include the justifications which they draw upon to support their recommendations. For each recommendation, analysts put varying weights on each information category by mentioning it more or less frequently. Also, analysts hold different tones toward each information category.

To test the relation between the information content of analyst reports and recommendations, I check both the frequency and tones of information categories in the following sections.

1.3.1. Frequency of information categories and recommendations

Based on the five information categories, I compute the frequency of words included in a given category appearing in the first page of reports. To address the concern that analysts' overall tone in the report may capture their tone towards different categories, I include the $Report\_Tone_{j,t}$ based on the net tone measure as used in prior studies(e.g., Henry and Leone (2016), Druz et al. (2020)). To address the concern that analysts from sell-side firms generally recommend "glamour" stocks (Jegadeesh et al. (2004), I control the firm size, market beta, 1-year momentum, and book-to-market ratio. Analysts may exploit different volumes of information in their decision-making process. Therefore, I control the total number of words as a proxy for the volume of information. To control for the industry- and market-wide conditions, I include industry and year fixed-effects. T-statistics are adjusted for heteroskedasticity and within-firm correlation by using heteroskedasticity-consistent standard errors clustered at the firm level. The results are also robust after controlling year-month fixed effects.

[Table 3 inserts here]

Table 3 shows how the frequency of information categories are related to analysts' recommendations. Among the five information categories, there is a weak pattern showing that growth, new developments, and conviction information are positively related to favorable recommendations. While the frequency of earnings is negatively related to analysts' recommendation level, which means that analysts tend to provide more hard information when they are making sell recommendations.

1.3.2. Tone of information categories and recommendations

To check how the recommendations are made based on information categories, I collect stock recommendations and their supportive content from analyst reports. Then I categorize the information into five categories using the method as described in the previous section and measure analysts' tones regarding the firm's growth, earnings, new developments, management transactions, and conviction.

Table 4 presents the regression results. The first five columns document the unconditional relation between analyst tone of information categories and recommendations. In general, higher tones that analysts hold towards the information category, the more favorable recommendation they will make. Analysts react differently to each category of information when deciding on their recommendations. As shown in Column (1) and Column (2), analysts are more likely to make more favorable recommendations when they are optimistic about a firm's growth. On average, a one-standard-deviation increase in "growth" tone increases the probability of buy recommendation by 8%. While a one-standard-deviation increase in "eps" tone only increases the probability of buy recommendation by 3%.

In column (6), I check the role of information in a multivariate test. Controlling all information categories, tone of growth, new developments, management transactions remain positive and statistically significant. Also, the four categories have almost equal contributions. A one-standard-deviation increase of analysts' tone is associated with a stock recommendation that is 7.5% higher relative to the mean recommendation level. Also, after controlling other information categories, the tone of earnings becomes insignificant in predicting recommendations.

In addition, Huang et al. (2014) document that analysts' overall opinions in reports are positively related to their recommendation. To check whether the overall tone can capture the tone of information categories, I control the overall tone of the report. As shown in column (7), including the report tone in the regression increases the adjusted R^2 by 12%, which shows that the tone towards each information category provides additional information beyond the overall tone of the report. Also, the tone of growth, earnings, new developments, management transactions, and conviction remain to be significant in predicting recommendation. The results are robust after controlling a series of variables in directions consistent with prior studies.

[Table 4 inserts here]

1.4. Information content and stock market anomalies

1.4.1. Stock recommendations conditional on anomalies and information tones

This section aims to understand analysts' decision-making process conditional on the market anomalies. Prior studies document that analysts tend to make favourable recommendations to anomaly-based overvalued stocks (Jegadeesh, Kim, Krische, and Lee (2004), Guo, Li, and Wei (2019), Grinblatt, Jostova, and Philipov (2018), Engelberg, McLean, and Pontiff (2019)). But none of these studies provide specific explanations of why analysts recommend the anomaly-sell stocks. Research focusing on the incentives of analysts provides an indirect explanation by implying that analysts are motivated to recommend overvalued stocks due to business opportunity (Cowen, Groysberg, and Healy (2006), Malmendier and Shanthikumar (2014)) or favoring the management (Chen and Matsumoto (2006)). Using a broader set of information categories from analyst reports, I could check whether analysts' bias towards specific information will cause the biased recommendations.

Figure 2 and Figure 3 plot the percentage scores of the frequency and tone of information categories, respectively, among different types of stocks and recommendations.

[Figure 2 and Figure 3 insert here]

To further check the relationship, I define stocks as over/undervaluation based on two mispricing factors proposed by Stambaugh and Yuan (2016). These anomalies are also used in the analyst context by Birru et al. (2019). Stambaugh and Yuan (2016) separate 11 anomalies into two clusters based on the correlation of time-series anomaly returns or cross-sectional anomaly rankings. Here I exclude medium-term momentum (Jegadeesh and Titman (1993)) from PERF cluster due to the momentum crash from 2009 to 2013 documented by Daniel and Moskowitz (2016). Below I list the 10 anomalies used in this paper.

MGMT
(1) Net stock issuance (NSI) (Ritter (1991), Loughran and Ritter (1995))
(2) Composite net equity issuance (CEI) (Daniel and Titman (2006))
(3) Accounting accruals (Accrual) (Sloan (1996))
(4) Net operating assets (NOA) (Hirshleifer, Hou, Teoh, and Zhang (2004))
(5) Asset growth (AG) (Cooper, Gulen, and Schill (2008))
(6) Investment-to-assets (IA) (Titman, Wei, and Xie (2004), Xing (2008))

PERF
(7) Gross Profitability (GP) (Robert (2013)))
(8) Return on Asset (ROA) (Fama and French (2006))
(9) Failure probability (Campbell, Hilscher, and Szilagyi (2008))
(10) O-Score bankruptcy probability (Ohlson (1980))

Following Stambaugh and Yuan (2016), I first calculate the two mispricing scores by ranking stocks into deciles for each anomaly and taking the average of the rankings across all anomalies in each cluster. When constructing the mispricing scores, I require that a stock have non-missing values for at least three (two) anomalies for MGMT (PERF) cluster. The higher mispricing scores represent the higher degree of overvaluation, which will earn lower future returns.

Next, to check if analyst's tone of a specific information category is optimistic, I compare the tone of a given information category issued in the reports in month t with the median value of the same information category among all the reports issued in month t-1. For the first month, I compare the tone of a category with the median value of the same category among all reports in the whole sample.

[Table 5 and Table 6 insert here]

Table 5 and Table 6 provide regression results for MGMT cluster and PERF cluster. Column (1) of Table 5 documents that analysts tend to recommend overvalued stocks with respect to short side anomaly predictors. Considering the mean recommendation level for this sample is 0.25 with a standard deviation of 0.68, the coefficient of 0.05 in Column 1 shows that stock recommendations are about 20% higher relative to the mean recommendation level for MGMT-cluster-based overvalued stocks, which 28% lower for MGMT-cluster-based undervalued stocks. These results are consistent with the results in Engelberg et al. (2019) that stock recommendations are contradictory to anomaly signals.

From Column (2) to (7), I check the tone of information categories separately. Generally, the tone of all categories are related to favorable recommendations conditional on anomaly signals. Specifically, analysts' overall tone in the reports partly explains their favorable recommendations of overvalued stocks. If analysts are optimistic about the "all aspect" of the firm, they are more likely to recommend favorable recommendations to overvalued stocks. In addition, although analysts are incorrectly optimistic about MGMT-cluster-based overvalued stocks in general, reports on new developments and management transactions have investment value even after controlling for the predictive power of the mispricing signals. As shown in Column (4) and (5), analysts have no favorable recommendations to overvalued stocks, if the recommendations are based on optimistic tone of new developments and management transactions.

Further, for MGMT-cluster-based undervalued stocks, while analysts are on average incorrectly pessimistic, they tend to provide buy recommendations if they have optimistic opinion of firms' new developments, management transactions, and overall aspects in reports. The recommendation based on these categories have investment value.

Current evidence shows that analysts on average issue anomaly-consistent recommendations in PERF cluster (Li et al. (2020)) or among momentum anomalies (Engelberg et al. (2019)). As shown in Table 6, stock recommendations are about 32% lower relative to the mean recommendation level for PERF-cluster-based overvalued stocks, which 56% high for PERF-cluster-based undervalued stocks, showing that analysts generally recommend stocks with better performance recently. For the relationship with information categories, as analysts recommendations line up correctly with anomaly signals, there is no additional investment value from conditioning on most of the report categories, except growth and overall tone, after controlling for the investment value of the anomaly signals.

1.4.2. Investment value of analysts' opinions conditional on anomalies

Previous section shows that analysts' opinions towards specific information could explain why they recommend overvalued stocks, which implies that the value of the stock recommendation should be conditional on the arguments in reports. Then a natural question is whether the information provided in reports can reverse the negative returns predicted by higher mispricing score.

To test this hypothesis, I use 22 trading days cumulative Fama and French (1993) adjusted returns, CAR[0,21], as a proxy for stock performance in the future one month. As analysts behave differently in terms of their recommendations in MGMT and PERF mispricing clusters, I analyze the two mispricing clusters separately. Consistent with previous section, MGMT mispricing score is constructed from six anomalies: net stock issuance, composite net equity issuance, accounting accruals, net operating assets, asset growth, and investment-to-assets. As the momentum anomaly in my sample period, 2007–2014, behaves differently from the literature , I use four anomalies to build the PERF mispricing score: gross profitability, return on asset, failure probability, and O-score bankruptcy probability. In this section, I employ a similar tests as the previous one, and the regression results are shown in Table 7 and Table 8.

[Table 7 and Table 8 insert here]

In Table 7 and Table 8, $CAR[0, 21]$ represents 1-month cumulative abnormal return for a firm from the issuing date of its report. $MGMT\_Sell$ ($Mgmt\_Buy$) is a dummy variable which equals to one if firm's mispricing scores in MGMT clusters are larger (smaller) than top quartile (bottom quartile) of the MGMT-cluster-based mispricing score in the sample. I define the stock as OPT_Growth if analysts have higher tone towards the information category of growth in month t comparing to the median value of tone towards growth categories among reports issued in month t-1. Optimistic tone of other categories are calculated following the same way. $OPT\_Report_{j,t}$ captures the overall tone of the reports. Detailed definitions are shown in Appendix 1.

Table 7 and Table 8 provide regression results for MGMT and PERF clusters, respectively. Column (1) in Table 7 shows that MGMT-overvalued stocks will generate negative returns, while MGMT-undervalued stocks have no positive return predictability. This results are consistent with the anomaly literature showing that the profitability of anomalies is mainly from the short-leg of the anomaly portfolios. Column (2) to (7) in Table 7 show the performance of the overvalued stocks conditional on analysts' tone of specific information categories. Specifically, Column (2), (4) and (5) show that analysts' optimism toward growth, new developments, and management transactions can dent the negative returns predicted by overvalued stocks. According to the previous section, analysts tend to recommend overvalued stocks if they are optimistic about "all aspects" of the firm. In Column (7), the results show that this "wrong" recommendation will generate significant lower returns. I also control stock recommendations in the regression. As the reports include all types, new recommendations, reiterate, downgrade, upgrade, the overall relation is not significant. As analysts' recommendations are consistent with the anomalies in PERF cluster, tone of the categories have no interaction effects for most of categories in Table 8.

1.5. Market reaction to analysts' tones

In this section, I examine how the market responds to analysts' opinions towards various information categories. Huang et al. (2014) show that investors respond to the overall opinion of reports beyond the quantitative summary measures. However, it is less evident that analysts' opinions towards information categories would be incorporated into the prices. In addition, given that analysts put different weights on the information categories, it is not certain that investors would put similar weights on the information categories. To test these questions, I estimate the following regression.

$$\begin{aligned} CAR_{j,t} =& \alpha_0 + \beta_1 Tone\_Growth_{j,t} + \beta_2 Tone\_Eps_{j,t} + \beta_3 Tone\_New_{j,t} \\ &+ \beta_4 Tone\_Comp_{j,t} + \beta_5 Tone\_Believe_{j,t} + \beta_6 Report\_Tone \\ &+ \beta_7 Buy + \beta_8 Sell + \sum_k \gamma_k Controls_{j,t,k} + \epsilon_{j,t} \end{aligned} \tag{1.3}$$

In Equation (5), I regress CAR[0,1] and CAR[2,21] on analysts' opinion towards five information categories. CAR[0,1] and CAR[0,21] are cumulative Fama and French (1993) adjusted returns in the first two days and next 19 days from report issuing date. $Tone\_Growth_{j,t}$, $Tone\_Eps_{j,t}$,

$Tone\_New_{j,t}$, and $Tone\_Comp_{j,t}$ represent analysts' opinion towards firm's growth, earnings, new developments, and management transactions. $Tone\_Believe_{j,t}$ represents analysts' confidence about their arguments. I also control some firm characteristics, including market beta, 1-year momentum, size, and book-to-market ratio. In addition, I control the number of words in the first page. Regression results are shown in Table 9.

[Table 9 inserts here]

1.5.1. Immediate stock market reactions

Columns (1) to (4) in Table 9 document how the market reacts to information tones on the issuing day of reports and the immediately following day. Column (1) shows that the market would react to the recommendation in reports immediately. Buy recommendation will generate 48 basis points higher returns, and Sell recommendation will generate 58 basis points lower returns comparing with the Hold recommendation.

In Column (2), I put all the information tones in the regression. In the previous section, analysts put more weights on growth and fewer weights on earnings when they are making recommendations. The market reaction has different patterns. For the tone of growth, a one-standard-deviation increase will lead to 37 basis points higher than the average value while a one-standard-deviation in tone of earnings will lead 33 basis points higher than the average value. In Column (3), controlling the overall tone of the report, tone of earnings, new developments, and management transactions can still predict higher returns, which means that the market could capture the difference of various information categories. Considering that all information categories are significantly related to analyst recommendations, the market has different opinions about the information, putting more weights on the firm's earnings than analysts in their recommendations.

In addition, once the tones are controlled, both buy and sell recommendations become less significant, and the adjusted R^2 increases. It shows that the information tones provide incremental information to the market conditional on their recommendations. This finding confirms and extends Asquith et al. (2005).

1.5.2. Event-time returns beyond the initial market reactions

The current results show that the market will react to the tones toward information in different categories. However, it is not clear whether and how the market reacts to the information in days or weeks later. The market could rational-price, over-price, or under-price specific information (rational pricing, continuation, or reversal). If the market prices specific information contained in the reports, there will be no relation between this information category and market returns in the post-issued period. If the investors under-react to some information tones, there will be a drift of returns in the direction of the initial market reaction. In the case of over-reaction, there will be a reversal. In addition, the market reacts to information tones differently. So, the post-issued influence among different information tones may be different.

Column (5) of Table 9 shows that the market tends to overreact to analyst recommendations in reports. The stocks with favorable recommendations will generate lower returns in the following period, and less favorable recommendations will generate higher returns. As the sample includes all report types, upgrades, downgrades, and reiterations, this result represents the average market reaction to all types of reports.

Column (6) and Column (7) show how the market reacts to information tone in the post-issuing period with and without the control of the overall report tone. Information tones behave differently. Although analysts weight less on earnings, investors take the tone of earnings as a valuable information source. After the immediate reaction, the earnings tone will continue to predict a positive market reaction. For the tone of other information categories, the coefficients are not significant, which shows that the market can immediately price the information adequately.

1.5.3. Calendar-time tests

Results in prior section implies that only earnings has investment value after the immediate-releasing period. In this part, I further test whether a portfolio strategy constructed to capture the potential lags in the revelation of information tones.

The portfolios are built to utilize the most recent information tones in analyst reports. At the beginning of each month, I rank all stocks with reports issued in t-1 into two groups based on a specific information tone. On average, each portfolio includes 500 firms. For each of the two portfolios, I then compute equal-weighted and value-weighted returns and Fama and French (2015) alphas. A portfolio that long (short) stocks with high (low) tone could capture market under-reaction to specific information. The results are presented in Table 9.

[Table 10 inserts here]

As shown in the table, a long-short portfolio based on Eps tone could generate returns 0.6% of 0.7% for equal- and value-weighted results. The alphas from Fama and French (2015) are still significant. The other information tones are not significant.

1.6. Conclusion

Analysts are important financial intermediaries. Most studies focus on analysts' forecasts on earnings, target prices, and recommendations, but few explore how analysts make these forecasts. As the information vendor, analysts' main job is more than to provide predictions periodically but to discover, interpret, and disseminate valuable information. Therefore, analyzing word-level information in analyst reports helps understand how analysts provide value in their reports even when these recommendations in the reports appear to be contradictory to well-known signals of mispricing.

Table 1. highest-occurring words in the first page of reports and in selected information categories

This table reports the word frequency and the highest-occurring words in each of five selected categories. Panel A reports the highest-occurring words in the first page of each analyst report. Panel B reports 20 word categories based on the Word2Vec (see details in the methodology section). I select five categories with the words “growth”, “eps”, “new”, “company”, and “believe”. I label them as growth, earnings, new developments, management transactions, and conviction, respectively.

Panel A: Highest-occurring words in the first page

Rank	Words	Freq	Rank	Words	Freq	Rank	Words	Freq
1	Growth	71,100	6	Estimate	30,213	11	Higher	27,046
2	Eps	59,490	7	Line	28,934	12	Result	25,992
3	Guidance	39,354	8	Revenue	28,457	13	Lower	25,294
4	New	33,051	9	Believe	28,042	14	Consensus	25,247
5	Company	31,797	10	Sales	27,993	15	Price	24,279

Panel B: Highest-occurring words in the 5 topics

Top seed word	Categorized words based on Word2Vec	Selected
Growth	growth, sales, result, strong, upside, beat, outlook, data, increase, trend, core, expansion, decline, performance, organic, improvement, average, comps, environment, strength	Yes
Eps	eps, guidance, estimate, revenue, **result**, consensus, expectation, target, end, report, model, forecast, valuation, **beat**, **outlook**, gross, million, rating, tax, range	Yes
Guidance	**eps**, **guidance**, **estimate**, **consensus**, expect, earnings, **expectation**, **target**, **end**, ebitda, **model**, **valuation**, **forecast**, **outlook**, **trend**, MSe (Morgan Stanley estimate), slightly, **est (estimate)**, street, come	No
New	new, next, **report**, last, change, full, raise, look, pipeline, lower, several, corp, previously, consumer, past, coming, momentum, development, recently, market	Yes
Company	company, management, mgmt, deal, capex, asset, plan, co, transaction, ceo, fda, wireless, brand, integration, cfo, fund, sense, midstream, expand, stake	Yes
Estimate	**eps**, **guidance**, **estimate**, **consensus**, expect, earnings, **expectation**, **target**, **end**, ebitda, **report**, **model**, **valuation**, **forecast**, **beat**, **outlook**, **revenue**, slightly, **range**, **est**	No
Line	**eps**, line, **strong**, **upside**, positive, **report**, **beat**, ahead, solid, **decline**, **range**, relative, weak, versus, flat, overall, good, unchanged, miss	No
Revenue	**eps**, **guidance**, **estimate**, **revenue**, **sales**, **consensus**, margin, earnings, operating, ebitda, cost, pricing, gross, **revenues**, cost, MSe, slightly, total, est, organic, segment	No
Believe	believe, see, expect, likely, view, **model**, rate, think, thesis, note, **look**, view, concern, like, exposure, provide, get, assume, make, suggest	Yes
Sales	**eps**, **revenue**, **sales**, price, margin, ebitda, **gross**, revenue, cost, **trend**, MSe, slightly, demand, organic, comp, flat, cons, volume, **comps**, international	No

Table 2. summary statistics

Panel A provides summary statistics for the main variables of interest. Panel B provides the correlation of the key variables. Information categories are defined from the information content in the first page of analysts reports. Refer to Appendix A for detailed description of variables.

Panel A: Summary statistics

Variable	Obs	Mean	Std	1%	10%	25%	50%	75%	90%	99%
Tone_Growth	34,531	6.54	4.10	-2	1	4	6	9	12	17
Tone_Eps	34,531	7.48	4.42	-1	2	4	7	11	13	18
Tone_New	34,531	4.79	3.09	-2	1	3	5	7	9	13
Tone_Comp	34,531	1.90	1.93	-2	0	1	2	3	4	8
Tone_Believe	34,531	3.78	2.51	-2	1	2	4	5	7	10
Report_Tone	34,531	0.73	5.00	-13	-6	-2	1	4	7	12
# of words	34,531	297.25	45.78	129	239	276	304	327	347	386
# of sentence	34,531	14.94	5.13	5	9	11	14	18	21	33
CAR[0,1] (%)	33,903	0.02	5.33	-19.09	-5.39	-2.09	0.03	2.17	5.45	18.11
CAR[0,21] (%)	33,903	0.15	10.55	-33.28	-11.77	-5.13	0.27	5.61	11.84	34.24
CAR[-2,-1] (%)	33,903	-0.00	0.03	-0.11	-0.03	-0.01	0	0.01	0.03	0.11
Size (billion)	29,762	23.34	41.15	0.25	1.06	2.60	7.60	21.69	62.68	217.78
Book-to-market	29,762	0.43	0.35	-0.34	0.09	0.21	0.34	0.58	0.88	1.81
Beta	29,576	1.11	0.50	0.28	0.52	0.74	1.05	1.41	1.79	2.58

Panel B: Correlation table

	Tone_Growth	Tone_Eps	Tone_New	Tone_Comp	Tone_Believe	Report_Tone
Tone_Growth	1					
Tone_Eps	0.42	1				
Tone_New	0.24	0.12	1			
Tone_Comp	0.07	-0.03	0.31	1		
Tone_Believe	0.18	0.12	0.27	0.16	1	
Report_Tone	0.52	0.33	0.41	0.24	0.31	1

Table 3. Stock recommendation and frequency of information categories

This table presents regression results for analysts' recommendation level between 2007 and 2014. The dependent variable is stock recommendation from analysts reports, which takes a value of 1 for "Overweight (Buy)", 0 for "equal-weight (Hold)" and -1 for "underweight (Sell)". Information categories are based on semantic and syntactic similarity of words in the first page of the reports. For each category, I compute the frequency as the times of words appeared in the first page of reports. Variables with "growth", "eps", "new", "comp", "believe" represent the frequency of information categories about growth, earnings, new developments, management transactions, and conviction. Refer to Appendix A for detailed description of variables. T-statistics are in parentheses with heteroskedastic-consistent standard errors clustered at company level. Industry and year fixed effects are included. *, **, and *** indicate statistical significance at the 10%, 5%, and 1%, respectively.

	Recommendation level (Buy = 1, Hold = 0, Sell = -1)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Freq_Growth	0.002*					0.004***	-0.003**
	(1.704)					(2.849)	(-1.994)
Freq_Eps		-0.005***				-0.006***	-0.004***
		(-5.247)				(-5.633)	(-3.745)
Freq_New			0.006***			0.005**	0.001
			(2.827)			(2.208)	(0.647)
Freq_Comp				0.001		-0.002	-0.004
				(0.466)		(-0.516)	(-1.400)
Freq_Believe					0.008***	0.007***	0.006**
					(3.419)	(2.912)	(2.444)
Report tone							0.026***
							(19.306)
Market beta							0.128***
							(3.631)
Momentum (t-11, t-1)							0.135***
							(5.823)
Size							0.134***
							(11.626)
Book-to-market							-0.035
							(-1.527)
Ln(Words)							0.010
							(0.191)
Constant	0.519	0.568	0.496	0.531	0.519	0.504	-0.141
	(1.486)	(1.573)	(1.429)	(1.484)	(1.441)	(1.473)	(-0.491)
Observations	33,080	33,080	33,080	33,080	33,080	33,080	27,737
R-squared	0.105	0.107	0.106	0.105	0.106	0.109	0.221
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 4. Stock recommendation and tone of information categories

This table presents regression results for analysts' recommendation level between 2007 and 2014. The dependent variable is stock recommendation from analysts reports, which takes a value of 1 for "Overweight (Buy)", 0 for "equal-weight (Hold)" and -1 for "underweight (Sell)". Information categories are based on semantic and syntactic similarity of words in the first page of the reports. For each of the word in a category, I determine the tone by checking whether there is a surrounded sentiment word based on Loughran and McDonald (2011) sentiment-word dictionary. The tone of a category is the sum of the tone of all the words in the category. Variables with "growth", "eps", "new", "comp", "believe" represent the frequency of information categories about growth, earnings, new developments, management transactions, and conviction. Refer to Appendix A for detailed description of variables. T-statistics are in parentheses with heteroskedastic-consistent standard errors clustered at company level. Industry and year fixed effects are included. *, **, and *** indicate statistical significance at the 10%, 5%, and 1%, respectively.

	Recommendation level (Buy = 1, Hold = 0, Sell = -1)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Tone_Growth	0.022***					0.018***	0.007***	0.005***
	(12.525)					(10.409)	(3.965)	(2.643)
Tone_Eps		0.007***				-0.002	-0.005***	-0.004***
		(4.475)				(-1.444)	(-3.805)	(-2.716)
Tone_New			0.024***			0.012***	0.004**	0.004*
			(11.832)			(6.063)	(2.173)	(1.927)
Tone_Comp				0.026***		0.014***	0.006*	0.005
				(7.859)		(4.187)	(1.849)	(1.518)
Tone_Believe					0.030***	0.020***	0.013***	0.010***
					(13.866)	(9.540)	(6.361)	(4.909)
Report_Tone							0.023***	0.022***
							(14.747)	(13.738)
Market beta								0.132***
								(3.764)
Momentum (t-12, t-1)								0.131***
								(5.666)
Size								0.135***
								(11.697)
Book-to-market								-0.035
								(-1.497)
Ln(Words)								-0.089*
								(-1.875)
Constant	0.618**	0.703***	0.600**	0.684***	0.635**	0.461*	0.556**	0.225
	(2.537)	(2.711)	(2.341)	(2.653)	(2.500)	(1.879)	(2.478)	(0.823)
Observations	34,455	34,455	34,455	34,455	34,455	34,455	34,455	27,742
R-squared	0.118	0.105	0.114	0.108	0.115	0.131	0.147	0.221
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 5. Stock recommendation conditional on market anomalies and tone of information categories: MGMT cluster

This table presents regression results for analysts' recommendation level between 2007 and 2014. The dependent variable is stock recommendation from analysts reports, which takes a value of 1 for "Overweight (Buy)", 0 for "equal-weight (Hold)" and -1 for "underweight (Sell)". Information categories are based on semantic and syntactic similarity of words in the first page of the reports. For each of the word in a category, I determine the tone by checking whether there is a surrounded sentiment word based on Loughran and McDonald (2011) sentiment-word dictionary. The tone of a category is the sum of the tone of all the words in the category. Opt_Growth equals one if a stock's tone of growth opportunities is larger than the median value of the same information category among all the reports issued in prior month. The other variables with *Opt_* prefix are defined similarly. Variables with "growth", "eps", "new", "comp", "believe" represent the frequency of information categories about growth, earnings, new developments, management transactions, and conviction. *Mgmt Sell* equals one if a stock's Mgmt-mispricing score is larger than 6.8 (highest quartile in the sample), and zero otherwise. *Mgmt Buy* equals one if a stock's Mgmt-mispricing score is lower than 4.2 (lowest quartile in the sample). Refer to Appendix A for detailed description of variables. T-statistics are in parentheses with heteroskedastic-consistent standard errors clustered at company level. Industry and year fixed effects are included. *, **, and *** indicate statistical significance at the 10%, 5%, and 1%, respectively.

Table 5 (continued)

	Recommendation level (Buy = 1, Hold = 0, Sell = -1)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Mgmt_Sell	0.048*	0.061*	0.062**	0.046	0.053	0.060**	0.016
	(1.851)	(1.958)	(1.987)	(1.578)	(1.611)	(2.163)	(0.525)
Mgmt_Buy	-0.073***	-0.090**	-0.071**	-0.102***	-0.103***	-0.086***	-0.118***
	(-2.654)	(-2.503)	(-2.037)	(-3.428)	(-3.378)	(-3.018)	(-3.593)
Opt_Growth		0.131***					
		(7.838)					
Mgmt_Sell × Opt_Growth		-0.018					
		(-0.677)					
Mgmt_Buy × Opt_Growth		0.034					
		(1.134)					
Opt_Eps			0.041**				
			(2.578)				
Mgmt_Sell × Opt_Eps			-0.024				
			(-1.012)				
Mgmt_Buy × Opt_Eps			-0.005				
			(-0.174)				
Opt_New				0.093***			
				(6.586)			
Mgmt_Sell × Opt_New				0.001			
				(0.041)			
Mgmt_Buy × Opt_New				0.059***			
				(2.675)			
Opt_Comp					0.059***		
					(3.616)		
Mgmt_Sell × Opt_Comp					-0.007		
					(-0.264)		
Mgmt_Buy × Opt_Comp					0.050**		
					(2.235)		
Opt_Believe						0.127***	
						(9.300)	
Mgmt_Sell × Opt_Believe						-0.020	
						(-0.982)	
Mgmt_Buy × Opt_Believe						0.026	
						(1.138)	
Opt_Report							0.202***
							(12.322)
Mgmt_Sell × Opt_Report							0.050**
							(1.974)
Mgmt_Buy × Opt_Report							0.094***
							(3.642)
Constant	0.507	0.422	0.486	0.413	0.441	0.433	0.330
	(1.356)	(1.124)	(1.277)	(1.105)	(1.179)	(1.114)	(0.939)
Observations	33,080	33,080	33,080	33,080	33,080	33,080	33,080
R-squared	0.109	0.118	0.109	0.115	0.111	0.117	0.138
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 6. Stock recommendation conditional on market anomalies and tone of information categories: PERF cluster

This table presents regression results for analysts' recommendation level between 2007 and 2014. The dependent variable is stock recommendation from analysts reports, which takes a value of 1 for "Overweight (Buy)", 0 for "equal-weight (Hold)" and -1 for "underweight (Sell)". Information categories are based on semantic and syntactic similarity of words in the first page of the reports. For each of the word in a category, I determine the tone by checking whether there is a surrounded sentiment word based on Loughran and McDonald (2011) sentiment-word dictionary. The tone of a category is the sum of the tone of all the words in the category. Opt_{Growth} equals one if a stock's tone of growth opportunities is larger than the median value of the same information category among all the reports issued in prior month. The other variables with Opt_{-} prefix are defined similarly.Variables with "growth", "eps", "new", "comp", "believe" represent the frequency of information categories about growth, earnings, new developments, management transactions, and conviction. *Perf Sell* equals one if a stock's Perf-mispricing score is larger than 6.8 (highest quartile in the sample), and zero otherwise. *Perf Buy* equals one if a stock's Perf-mispricing score is lower than 2.8 (lowest quartile in the sample). Refer to Appendix A for detailed description of variables. T-statistics are in parentheses with heteroskedastic-consistent standard errors clustered at company level. Industry and year fixed effects are included. *, **, and *** indicate statistical significance at the 10%, 5%, and 1%, respectively.

Table 6 (continued)

	Recommendation level (Buy = 1, Hold = 0, Sell = -1)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Perf_Sell	-0.076**	-0.083**	-0.074*	-0.072**	-0.075**	-0.079**	-0.109***
	(-2.409)	(-2.509)	(-1.915)	(-2.089)	(-2.065)	(-2.391)	(-3.038)
Perf_Buy	0.137***	0.121***	0.157***	0.126***	0.148***	0.150***	0.129***
	(4.269)	(3.549)	(4.354)	(3.716)	(4.001)	(4.458)	(3.497)
Opt_Growth		0.110***					
		(9.042)					
Perf_Sell × Opt_Growth		0.085***					
		(3.026)					
Perf_Buy × Opt_Growth		0.031					
		(1.232)					
Opt_Eps			0.033**				
			(2.236)				
Perf_Sell × Opt_Eps			0.002				
			(0.083)				
Perf_Buy × Opt_Eps			-0.035				
			(-1.484)				
Opt_New				0.108***			
				(7.486)			
Perf_Sell × Opt_New				-0.006			
				(-0.246)			
Perf_Buy × Opt_New				0.018			
				(0.706)			
Opt_Comp					0.076***		
					(4.666)		
Perf_Sell × Opt_Comp					-0.002		
					(-0.081)		
Perf_Buy × Opt_Comp					-0.018		
					(-0.671)		
Opt_Believe						0.130***	
						(9.746)	
Perf_Sell × Opt_Believe						0.009	
						(0.389)	
Perf_Buy × Opt_Believe						-0.028	
						(-1.287)	
Opt_Report							0.211***
							(13.576)
Perf_Sell × Opt_Report							0.086***
							(3.038)
Perf_Buy × Opt_Report							0.003
							(0.117)
Constant	0.623*	0.520	0.598	0.516	0.541	0.545	0.402
	(1.738)	(1.490)	(1.629)	(1.439)	(1.507)	(1.453)	(1.242)
Observations	33,080	33,080	33,080	33,080	33,080	33,080	33,080
R-squared	0.114	0.122	0.115	0.120	0.117	0.122	0.143
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 7: Investment value conditional on market anomalies and tone of information categories: MGMT cluster

This table presents regression results of investment value of market anomalies and information categories over 2007 and 2014. The dependent variable is is cumulative Fama and French (1993) adjusted 21-day abnormal returns starting from the report date. Information categories are based on semantic and syntactic similarity of words in the first page of the reports. For each of the word in a category, I determine the tone by checking whether there is a surrounded sentiment word based on Loughran and McDonald (2011) sentiment-word dictionary. The tone of a category is the sum of the tone of all the words in the category. Opt_{Growth} equals one if a stock's tone of growth opportunities is larger than the median value of the same information category among all the reports issued in prior month. The other variables with $Opt\_$ prefix are defined similarly. Variables with "growth", "eps", "new", "comp", "believe" represent the frequency of information categories about growth, earnings, new developments, management transactions, and conviction. *Mgmt Sell* equals one if a stock's Mgmt-mispricing score is larger than 6.8 (highest quartile in the sample), and zero otherwise. *Mgmt Buy* equals one if a stock's Mgmt-mispricing score is lower than 4.2 (lowest quartile in the sample). Refer to Appendix A for detailed description of variables. T-statistics are in parentheses with heteroskedastic-consistent standard errors clustered at company level. Industry and year fixed effects are included. *, **, and *** indicate statistical significance at the 10%, 5%, and 1%, respectively.

Table 7 (continued)

	CAR[0,21]						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Mgmt_Sell	-0.505**	-0.230	-0.715**	-0.261	-0.497	-0.572*	-0.833**
	(-2.274)	(-0.693)	(-2.144)	(-0.861)	(-1.475)	(-1.810)	(-2.499)
Mgmt_Buy	0.022	0.395	0.301	0.306	0.074	0.232	0.291
	(0.122)	(1.355)	(1.093)	(1.249)	(0.282)	(0.894)	(1.071)
Opt_Growth		1.032***					
		(5.258)					
Mgmt_Sell × Opt_Growth		-0.437					
		(-1.235)					
Mgmt_Buy × Opt_Growth		-0.650**					
		(-2.053)					
Opt_Eps			1.200***				
			(6.032)				
Mgmt_Sell × Opt_Eps			0.406				
			(1.122)				
Mgmt_Buy × Opt_Eps			-0.511				
			(-1.629)				
Opt_New				0.943***			
				(5.034)			
Mgmt_Sell × Opt_New				-0.433			
				(-1.250)			
Mgmt_Buy × Opt_New				-0.504*			
				(-1.765)			
Opt_Comp					0.505**		
					(2.445)		
Mgmt_Sell × Opt_Comp					-0.005		
					(-0.012)		
Mgmt_Buy × Opt_Comp					-0.095		
					(-0.323)		
Opt_Believe						0.469**	
						(2.576)	
Mgmt_Sell × Opt_Believe						0.124	
						(0.349)	
Mgmt_Buy × Opt_Believe						-0.377	
						(-1.342)	
Opt_Report							1.609***
							(7.868)
Mgmt_Sell × Opt_Report							0.547
							(1.384)
Mgmt_Buy × Opt_Report							-0.488
							(-1.594)
Buy		0.012	0.058	0.021	0.048	0.023	-0.145
		(0.064)	(0.314)	(0.115)	(0.260)	(0.125)	(-0.785)
Sell		0.310	0.297	0.288	0.272	0.258	0.486*
		(1.208)	(1.151)	(1.121)	(1.058)	(1.007)	(1.888)
Constant	-5.735**	-6.410***	-6.641**	-6.596***	-6.322***	-6.022**	-7.060***
	(-2.410)	(-2.699)	(-2.449)	(-2.852)	(-2.663)	(-2.473)	(-3.122)
Observations	33,078	33,078	33,078	33,078	33,078	33,078	33,078
R-squared	0.018	0.019	0.020	0.019	0.018	0.018	0.023
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 8: Investment value conditional on market anomalies and tone of information categories: PERF cluster

This table presents regression results of investment value of market anomalies and information categories over 2007 and 2014. The dependent variable is is cumulative Fama and French (1993) adjusted 21-day abnormal returns starting from the report date. Information categories are based on semantic and syntactic similarity of words in the first page of the reports. For each of the word in a category, I determine the tone by checking whether there is a surrounded sentiment word based on Loughran and McDonald (2011) sentiment-word dictionary. The tone of a category is the sum of the tone of all the words in the category. Opt_{Growth} equals one if a stock's tone of growth opportunities is larger than the median value of the same information category among all the reports issued in prior month. The other variables with $Opt\_$ prefix are defined similarly. Variables with "growth", "eps", "new", "comp", "believe" represent the frequency of information categories about growth, earnings, new developments, management transactions, and conviction. *Perf Sell* equals one if a stock's Perf-mispricing score is larger than 6.8 (highest quartile in the sample), and zero otherwise. *Perf Buy* equals one if a stock's Perf-mispricing score is lower than 2.8 (lowest quartile in the sample). Refer to Appendix A for detailed description of variables. T-statistics are in parentheses with heteroskedastic-consistent standard errors clustered at company level. Industry and year fixed effects are included. *, **, and *** indicate statistical significance at the 10%, 5%, and 1%, respectively.

Table 8 (continued)

	CAR[0,21]						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Perf_Sell	-0.407	-0.562**	-0.757**	-0.554*	-0.441	-0.514	-0.499
	(-1.571)	(-1.996)	(-2.043)	(-1.680)	(-1.067)	(-1.443)	(-1.368)
Perf_Buy	0.276	0.120	0.268	0.347	0.359	0.263	0.254
	(1.573)	(0.605)	(1.080)	(1.430)	(1.391)	(1.065)	(1.027)
Opt_Growth		0.565***					
		(3.896)					
Perf_Sell × Opt_Growth		0.905**					
		(2.385)					
Perf_Buy × Opt_Growth		0.387*					
		(1.699)					
Opt_Eps			0.927***				
			(4.836)				
Perf_Sell × Opt_Eps			0.858**				
			(2.166)				
Perf_Buy × Opt_Eps			-0.004				
			(-0.016)				
Opt_New				0.656***			
				(3.575)			
Perf_Sell × Opt_New				0.237			
				(0.629)			
Perf_Buy × Opt_New				-0.104			
				(-0.379)			
Opt_Comp					0.514***		
					(2.620)		
Perf_Sell × Opt_Comp					0.020		
					(0.045)		
Perf_Buy × Opt_Comp					-0.122		
					(-0.413)		
Opt_Believe						0.347**	
						(2.001)	
Perf_Sell × Opt_Believe						0.172	
						(0.444)	
Perf_Buy × Opt_Believe						0.024	
						(0.093)	
Opt_Report							1.529***
							(8.218)
Perf_Sell × Opt_Report							0.265
							(0.641)
Perf_Buy × Opt_Report							0.020
							(0.074)
Buy		-0.033	0.018	-0.024	0.004	-0.020	-0.187
		(-0.176)	(0.095)	(-0.129)	(0.022)	(-0.109)	(-1.008)
Sell		0.366	0.340	0.338	0.321	0.308	0.540**
		(1.414)	(1.306)	(1.304)	(1.241)	(1.196)	(2.083)
Constant	-5.564**	-6.161**	-6.394**	-6.299**	-6.119**	-5.738**	-6.916***
	(-2.167)	(-2.487)	(-2.143)	(-2.436)	(-2.377)	(-2.172)	(-2.848)
Observations	33,078	33,078	33,078	33,078	33,078	33,078	33,078
R-squared	0.018	0.019	0.020	0.019	0.018	0.018	0.022
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 9: Market reaction to tone of information categories

This table presents regression results of market reaction to information categories over 2007 and 2014. The dependent variable in columns in column (1) to (4) is CAR[0,1], the two-day, [0,1] cumulative Fama and French (1993) adjusted stock return in percent on and after the report issuing date. In column (5) to (8), the dependent variable is CAR[2,21], the 20 trading days [2,21] cumulative Fama and French (1993) adjusted stock return in percent from two days after the issuing date through the 21th day after that date. Refer to Appendix A for detailed description of variables. Variables with "growth", "eps", "new", "comp", "believe" represent the frequency of information categories about growth, earnings, new developments, management transactions, and conviction. T-statistics are in parentheses with heteroskedastic-consistent standard errors clustered at date level. Industry and year fixed effects are included. *, **, and *** indicate statistical significance at the 10%, 5%, and 1%, respectively.

Table 9 (continued)

	CAR[0, 1]				CAR[2, 21]			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Tone_Growth		0.091***	0.004	0.012		-0.027*	-0.017	-0.004
		(10.664)	(0.460)	(1.233)		(-1.944)	(-1.120)	(-0.246)
Tone_Eps		0.075***	0.048***	0.048***		0.043***	0.046***	0.050***
		(10.188)	(6.652)	(6.081)		(3.601)	(3.819)	(3.963)
Tone_New		0.089***	0.026**	0.038***		-0.028	-0.021	0.010
		(8.369)	(2.390)	(3.152)		(-1.477)	(-1.067)	(0.497)
Tone_Comp		0.098***	0.036**	0.035*		-0.029	-0.022	-0.063**
		(5.971)	(2.127)	(1.913)		(-1.044)	(-0.771)	(-2.108)
Tone_Believe		0.070***	0.017	0.026*		-0.030	-0.023	-0.006
		(5.299)	(1.291)	(1.802)		(-1.447)	(-1.136)	(-0.283)
Report_Tone			0.188***	0.183***			-0.022	-0.013
			(22.569)	(19.910)			(-1.640)	(-0.978)
Buy	0.482***	0.342***	0.204***	0.216***	-0.347***	-0.301***	-0.285***	0.021
	(6.959)	(4.979)	(2.980)	(2.885)	(-3.247)	(-2.797)	(-2.636)	(0.186)
Sell	-0.567***	-0.296***	-0.121	-0.187*	0.709***	0.676***	0.656***	0.293
	(-5.518)	(-2.893)	(-1.193)	(-1.694)	(4.284)	(4.057)	(3.927)	(1.607)
CAR[-2,-1]				-3.599**				7.228***
				(-2.495)				(3.169)
Market beta				0.207*				-0.174
				(1.849)				(-0.948)
Momentum (t-11, t-1)				-0.433***				-3.371***
				(-3.433)				(-14.928)
Size				-0.030				-0.113**
				(-1.034)				(-2.414)
Book-to-market				0.096				0.554***
				(1.613)				(5.302)
# of words				-0.642***				-0.525
				(-3.087)				(-1.484)
Constant	-1.377**	-3.561***	-2.806***	0.697	-5.684**	-5.493**	-5.580**	11.321***
	(-2.150)	(-5.526)	(-3.838)	(0.616)	(-2.204)	(-2.110)	(-2.136)	(5.945)
Observations	33,902	33,902	33,902	27,738	33,902	33,902	33,902	27,738
R-squared	0.011	0.032	0.049	0.053	0.024	0.024	0.024	0.050
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 10: Calendar investment value of tone of information categories

The table presents monthly calendar-time portfolio returns and Fama and French (2015) alpha to a information category strategy over 2007 and 2014. In the beginning of each month stocks are ranked into two portfolios based on their information opinion in the last month. Portfolio 1 contains stocks with the lowest information opinion, whereas portfolio 2 (high) contains the stocks with the highest information opinion. 2-1 is the long-short strategy which is long (short) portfolio 2 (1). The strategy is rebalanced monthly. Variables with "growth", "eps", "new", "comp", "believe" represent the frequency of information categories about growth, earnings, new developments, management transactions, and conviction. T-statistics are in parentheses.

		Growth tone			EPS tone			New tone		
		1 (< *med*)	2 (>= *med*)	2-1	1 (< *med*)	2 (>= *med*)	2-1	1 (< *med*)	2 (>= *med*)	2-1
EW	Raw return (%)	1.02	0.95	-0.07	0.65	1.25	**0.6**	0.871	1.02	0.15
		[1.38]	[1.59]	[-0.23]	[0.97]	[1.95]	**[3.53]**	[1.20]	[1.76]	[0.59]
	FF5 alpha	0.29	0.11	-0.17	-0.11	0.43	**0.53**	0.01	0.31	0.3
		[0.86]	[0.88]	[-0.59]	[-0.48]	[2.30]	**[2.98]**	[0.04]	[2.10]	[1.34]
VW	Raw return (%)	0.68	0.72	0.05	0.38	1.08	**0.7**	0.63	0.8	0.17
		[1.19]	[1.61]	[0.19]	[0.75]	[2.21]	**[3.13]**	[1.22]	[1.66]	[0.80]
	FF5 alpha	0.11	-0.01	-0.11	-0.17	0.33	**0.49**	0	0.09	0.09
		[0.64]	[-0.06]	[-0.52]	[-0.96]	[2.75]	**[2.10]**	[0.00]	[0.73]	[0.46]
		Company tone			Believe tone			Report tone		
		1 (< *med*)	2 (>= *med*)	2-1	1 (< *med*)	2 (>= *med*)	2-1	1 (< *med*)	2 (>= *med*)	2-1
EW	Raw return (%)	0.89	1	0.12	1.04	0.96	-0.08	0.81	1.1	0.28
		[1.28]	[1.59]	[0.62]	[1.44]	[1.59]	[-0.36]	[1.11]	[1.86]	[1.03]
	FF5 alpha	0.07	0.23	0.16	0.28	0.15	-0.13	0.05	0.29	0.23
		[0.29]	[1.24]	[0.80]	[0.97]	[1.08]	[-0.60]	[0.19]	[1.97]	[0.94]
VW	Raw return (%)	0.67	0.78	0.11	0.75	0.81	0.05	0.63	0.81	0.19
		[1.26]	[1.60]	[0.50]	[1.42]	[1.70]	[0.27]	[1.14]	[1.71]	[0.72]
	FF5 alpha	0.03	0.1	0.06	0.2	0.07	-0.12	0.04	0.07	0.03
		[0.22]	[0.85]	[0.27]	[1.26]	[0.64]	[-0.58]	[0.21]	[0.55]	[0.12]

Appendix A. Variable definitions

This table presents definitions of the variables in the study.

Variable	Definition
# of sentences	The total number of sentences in the first page. This number doesn't include the words in tables, captions, statements, and acknowledgements in the report.
# of words	The total number of words in the first page. This number doesn't include the words in tables, captions, statements, and acknowledgements in the report.
Beta	The stock's market beta estimated from the last four years of monthly returns.
Book-to-market ratio	Logarithm of the ratio of book value of equity divided by market capitalization at the end of its most recent fiscal year.
CAR[0,1]	The two-day, [0,1], cumulative Fama and French (1993) adjusted stock return in percent on and after the report issuing date.
CAR[2,21]	The 20 trading days, [2,21], cumulative Fama and French (1993) adjusted stock return in percent from 2 days after the report issuing date through the 21th day after that date.
CAR[0,21]	The 22 trading days, [0,21], cumulative Fama and French (1993) adjusted stock return in percent from the report issuing date through the 21th day after that date.
MGMT Sell	MGMT Sell is a dummy variable that equals 1 if the stock's composite mispricing score of the MGMT cluster is larger than 5. Following Stambaugh and Yuan (2016), composite MGMT mispricing score is the average ranking across 6 anomalies: net stock issues, composite equity issues, accruals, net operating assets, asset growth, and investment to assets.
OPT_Growth	A dummy variable equals one if a stock's tone of growth is larger than the median value of the same information category among all the reports issued in prior month.
OPT_Eps	A dummy variable equals one if a stock's tone of earnings is larger than the median value of the same information category among all the reports issued in prior month.
OPT_New	A dummy variable equals one if a stock's tone of new developments is larger than the median value of the same information category among all the reports issued in prior month.
OPT_Comp	A dummy variable equals one if a stock's tone of management transactions is larger than the median value of the same information category among all the reports issued in prior month.
OPT_Believe	A dummy variable equals one if a stock's tone of convition is larger than the median value of the same information category among all the reports issued in prior month.
Recommendation level	The sample reports now use three stock rating categories: Overweight (Buy), Equal-weight (Hold), and Underweight (Sell). Recommendation level takes a value of 1 for "Buys", 0 for "Holds", and -1 for "Sells".
Report_Tone	The difference of the number between positive and negative words based on the word lists compiled by Loughran and McDonald (2011) used in the relative textual contents in the first page of analysts reports.
PERF Sell	PERF Sell is a dummy variable that equals 1 if the stock's composite mispricing score of the PERF cluster is larger than 5. As momentum anomaly doesn't preform well in the period 2007—2014, I compute composite PERF mispricing score is the average ranking across 4 anomalies: distress, O-score, gross-profitability, and return on assets.

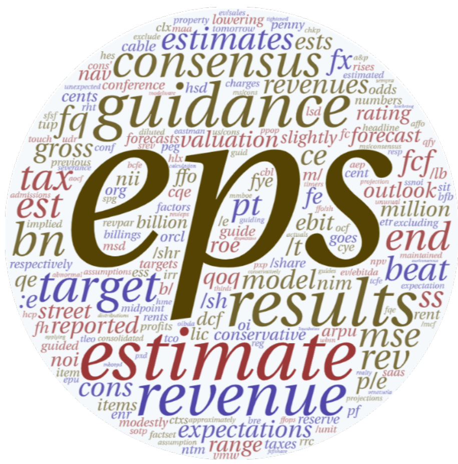
(continued)

Variable	Definition
Size	Logarithm of a firm's market capitalization at the end of its most recent fiscal year.
Tone_Growth	Analysts' tone of growth category, which is based on the "growth" related word cluster as shown in Table 1. Calculation process is shown in section 2.2. The variable is the difference of the number between non-negative related words and negative related words in "growth" word category. Negative related and non-negative related words are based on the word lists compiled by Loughran and McDonald (2011). Sentiment negation is considered in determining the tone of words.
Tone_Eps	Analysts' tone of earnings category, which is based on the "eps" related word cluster as shown in Table 1. See *OPT_Growth* for brief calculation process.
Tone_New	Analysts' tone of new developments category, which is based on the "new" related word cluster as shown in Table 1. See *OPT_Growth* for brief calculation process.
Tone_Comp	Analysts' tone of management transactions category, which is based on the "company" related word cluster as shown in Table 1. See *OPT_Growth* for brief calculation process.
Tone_Believe	Analysts' tone of conviction category, which is based on the "believe" related word cluster as shown in Table 1. See *OPT_Growth* for brief calculation process.

Figure 1: Five word clusters based on the methodology in Section 2.2

This figure shows the five most representative and dissimilar categories based on the methodology in Section 2.2. From Cluster 1 to 5, the categories are about the firms' growth, earnings, new developments, management transactions, and conviction.

Cluster 1: Growth



Cluster 2: Earnings

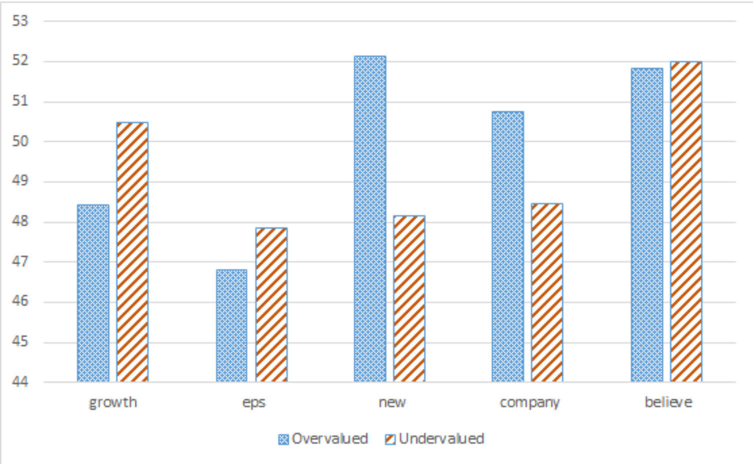
Cluster 3: New developments

Cluster 4: Management transactions

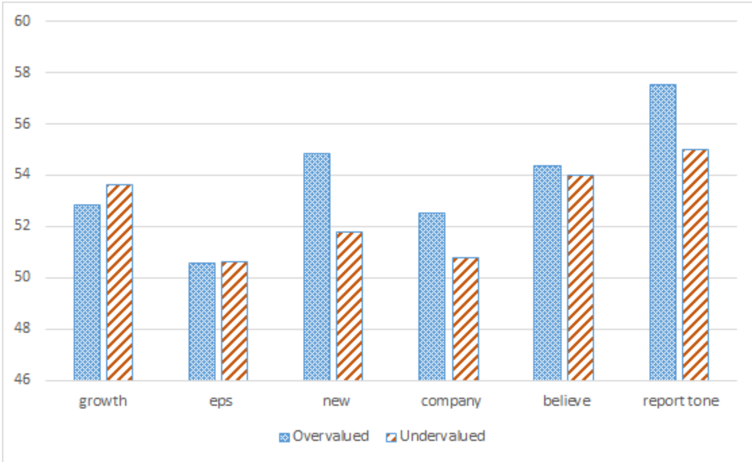
Cluster 5: Conviction

Figure 2: Percentage scores of information-categories' frequency and tone conditional on types of recommendations and MGMT-mispriced stocks

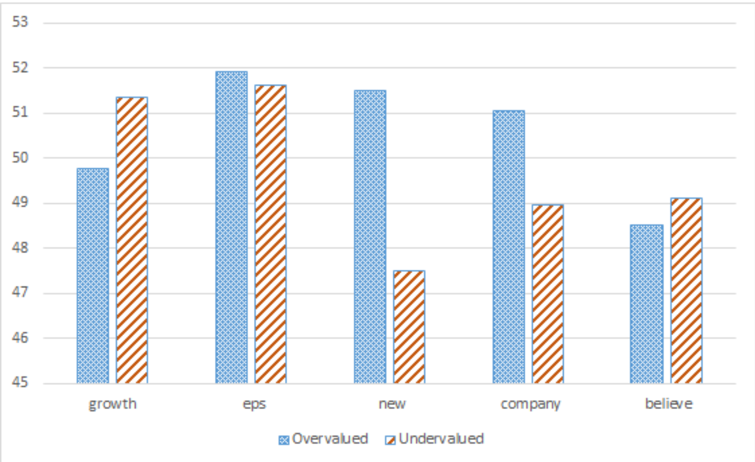
This figure shows the percentage scores of information-categories' frequency and tone conditional on stock recommendation levels and mispriced stocks based on MGMT mispricing score. I divide stocks into over/undervalued groups based using aggregate anomaly scores of MGMT cluster from Stambaugh and Yuan (2016). "Growth", "eps", "new", company", "believe" are the seed words of five information categories based on the methodology in Section 2.2, representing firms' growth, earnings, new developments, management transactions, and conviction. "Report tone" in Panel B and Panel D is analysts' overall opinion in reports. The sample period is from 2007 to 2014.



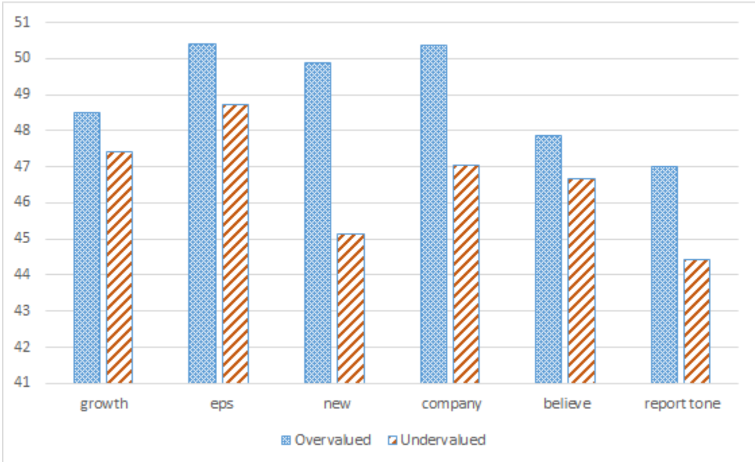
Panel A: Frequency score among Buy recommendations



Panel B:Tone score among Buy recommendation



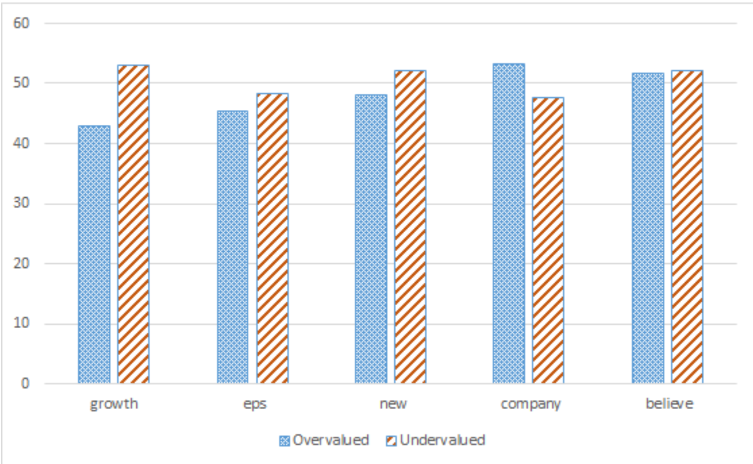
Panel C: Frequency score among Hold and Sell recommendations



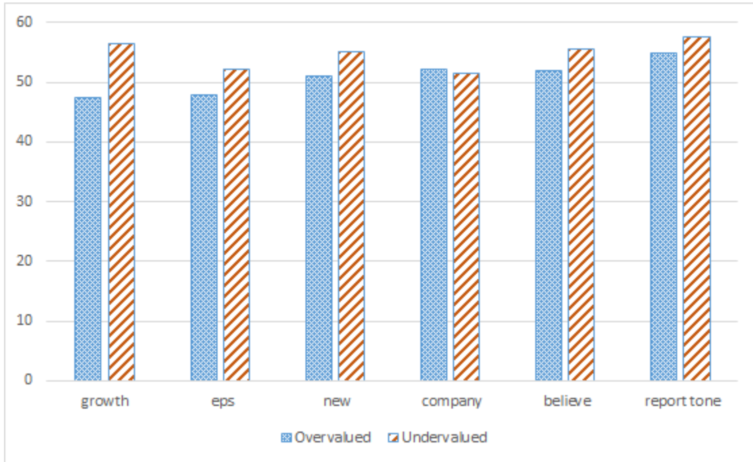
Panel D: Tone score among Hold and Sell recommendations

Figure 3: Percentage scores of information-categories' frequency and tone conditional on types of recommendations and PERF-mispriced stocks

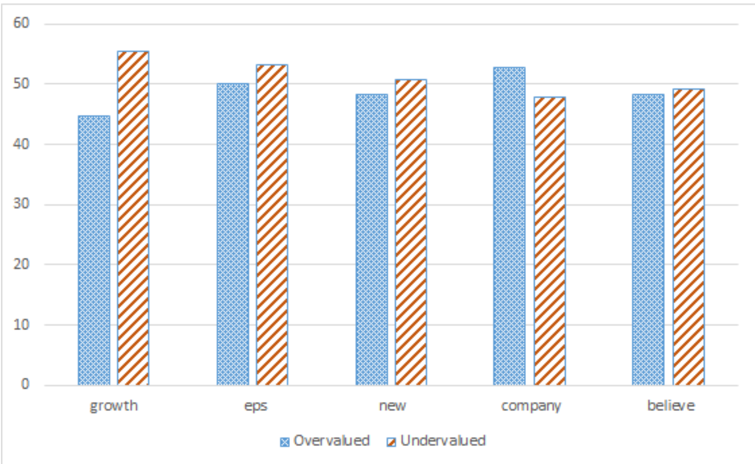
This figure shows the percentage scores of information-categories' frequency and tone conditional on stock recommendation levels and mispriced stocks based on PERF mispricing score. I divide stocks into over/undervalued groups based using aggregate anomaly scores of PERF cluster from Stambaugh and Yuan (2016). "Growth", "eps", "new", company", "believe" are the seed words of five information categories based on the methodology in Section 2.2, representing firms' growth, earnings, new developments, management transactions, and conviction. "Report tone" in Panel B and Panel D is analysts' overall opinion in reports. The sample period is from 2007 to 2014.



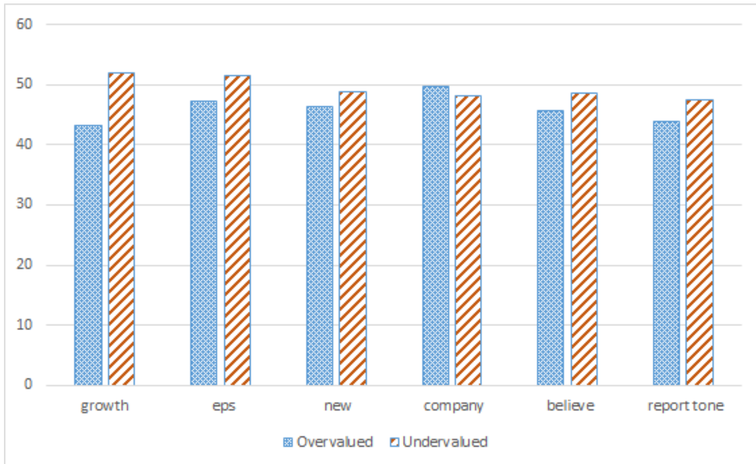
Panel A: Frequency score among Buy recommendations



Panel B:Tone score among Buy recommendation



Panel C: Frequency score among Hold and Sell recommendations



Panel D: Tone score among Hold and Sell recommendations

Chapter 2

The information cycle and return seasonality

By Haoyuan Li and Roger Loh

Roger Loh is at the Lee Kong Chian School of Business, Singapore Management University. 50 Stamford Road, Singapore 178899, Singapore. E-mail: rogerloh@smu.edu.sg. We thank Brad Barber, Gennaro Bernile, Hyun-Soo Doh, Kai Li, Weikai Li, Roni Michaely, Paul Schultz, Rong Wang, Chishen Wei, Xiaoyan Zhang, and conference/seminar participants at the 2019 SMU finance summer camp, 2019 SFS Cavalcade Asia Pacific, and Central University of Finance and Economics for helpful comments and suggestions.

Abstract

Heston and Sadka (2008) find that the monthly cross-sectional returns of stocks depend on their historical same calendar-month returns. We propose an information-cycle explanation for this seasonality anomaly—that firms' seasonal release of information coincide with higher returns during months with such dissolution of information uncertainty, and lower returns during months with no information releases. Using earnings announcements and changes in implied volatility as proxies for scheduled information releases, we find that seasonal winners in information-release months and seasonal losers in non-information release months indeed drive the seasonality anomaly. Our evidence shows that scheduled firm-level information releases can give rise to the appearance of an anomalous seasonal pattern when stock returns are in fact responding to information uncertainty.

Keywords: Return seasonality; Information cycle; Information Uncertainty; Earnings announcement premium

2.1. Introduction

Stocks have unevenly distributed monthly returns and tend to earn relatively high (low) returns in the same calendar month every year. This strong positive serial correlation of monthly U.S. stock returns at annual lags was first noted by Jegadeesh and Titman (1990), and later comprehensively documented by Heston and Sadka (2008). Keloharju, Linnainmaa, and Nyberg (2016) show that this anomaly also exists in international stock markets and other asset markets. To date, however, there has been no convincing explanation for this anomaly and whether it stems from mispricing or risk. Keloharju et al. (2016) argue that the seasonality anomaly might arise from systematic factors because it exists even in well-diversified portfolios and the strategy generates an extremely large return variance. However, it is not clear what this systematic factor might be. At the firm level, while it is well-known that some firm-specific events are also seasonal (e.g., see the review in Hartzmark and Solomon (2018)), Heston and Sadka (2008) show that the seasonal release of information about earnings or dividends cannot explain the return seasonality anomaly.

In this paper, we propose an information-cycle explanation for the seasonality anomaly. If a firm releases information seasonally, how would its information cycle look like? The firm will likely face high uncertainty prior to the scheduled information release, and this uncertainty gets resolved by the information release. If investors are unable to fully diversify such risk, they might require compensation for bearing this information uncertainty. Consequently, returns in the period prior to the announcement would be low, and returns in the announcement period would be high. We argue that if return seasonality is driven by an information cycle, seasonal losers will be more likely to coincide with the pre information-release period and seasonal winners more likely to occur in the information-release period.

To test this information-cycle explanation for the seasonality anomaly, we first sort stocks monthly based on their average return in the same calendar month in the last five years. To proxy for which stage of the information cycle these past seasonal returns are in, we check whether the majority of these past seasonal months are associated with an announcement of quarterly earnings. We find that the seasonality anomaly is much stronger when it is consistent with the information cycle (i.e., long seasonal winners in information-release months and short seasonal losers in non information-release months) compared to seasonality that is inconsistent with the information cycle (i.e., long seasonal winners in non information-release months and short seasonal losers in information-release months). Besides earnings announcements, we also find supportive evidence using the drop in option-implied volatility as an alternative proxy for uncertainty-reducing information releases.

Our study is related to a recent literature which documents that returns are higher in periods with important scheduled news releases—an announcement premium. At the firm-level, Frazzini and Lamont (2007) show that firms on average have higher returns in months that they are expected to announce quarterly earnings. Barber, De George, Lehavy, and Trueman (2013) argue that uncertainty about earnings information that is about to be disclosed is the primary driver of this earnings announcement premium. Linnainmaa and Zhang (2019) show that there is a distinct price pattern around the quarterly earnings announcement cycle. At the macro level, Cieslak, Morse, and Vissing-Jorgensen (2019) document that the entire U.S. equity risk premium is earned in weeks that coincide with the Fed's release of information about monetary policy in predictable intervals around Federal Open Market Committee (FOMC) meetings (see also Lucca and Moench (2015)). Hu, Pan, Wang, and Zhu (2019) show that the positive drift before FOMC meetings and other schedule macro announcements is associated with the dissolution of uncertainty in the periods just prior to the announcements. Ai, Han, Pan, and Xu (2020) show that firms that are more sensitive to monetary policy announcements experience a larger amount of implied variance reduction after macro announcements and earn a larger macro announcement premium. We apply this general idea to the seasonality anomaly and show that seasonal information releases at even the firm-level matter for stock returns and can give rise to the appearance of an anomalous seasonal return pattern when stock returns are responding predictably to firm-level information uncertainty.

We are certainly not the first to investigate an information explanation for the seasonality anomaly; e.g. Heston and Sadka (2008) also propose this idea. However, once an information event proxy has been identified, the approach typically used in the literature is to compare months with the information event versus months without. These tests always find (and we also confirm this) that return seasonality is strong in both event months (such as earnings announcement months) and non-event months and conclude that information events do not explain the anomaly. Our approach importantly differs from this by comparing seasonality that is *consistent* with the information cycle (i.e. long information-based winners and short non information-based losers) versus seasonality that is *inconsistent* with the information cycle (i.e. long non information-based winners and short information-based losers). We find that the seasonality anomaly mainly resides in the former case. This method of forming hedge portfolios where the extreme anomaly portfolios are obtained from the diagonals based on their consistency with a second signal has also been used by papers such as Jegadeesh, Kim, Krische, and Lee (2004), Piotroski and So (2012), and Guo, Li, and Wei (2020).

Our results imply that firms release value-relevant information about their businesses in fixed periodic intervals. For example, Apple's stock returns are the highest in October and this coincides with a high likelihood uncertainty-reducing news about the company. For a typical company, we show that the higher-return calendar months are associated with a higher probability of information release, while the lower-return months are associated with a lower probability of information release. These results obtain whether we use earnings announcements or a decrease in implied volatility as proxies for information release events in a given month.

Our results here shed light on the source of the seasonality anomaly. It is interesting that none of the recent factor models are able to dent the seasonality anomaly (Fama and French (2015), Stambaugh and Yuan (2016), Hou, Mo, Xue, and Zhang (2020), and Hou, Xue, and Zhang (2018)). The existing literature has conjectured that seasonality could stem from systematic risk but is unable to identify what this risk factor might be (Keloharju et al. (2016)). Bogousslavsky (2016) offers a theoretical model that uses the infrequent trading behavior of institutions as a possible source of the seasonality anomaly. However, it is not clear what drives institutions to trade infrequently. Our information-cycle story provides a simple reason for such infrequent and periodic trading—that information is released in a seasonal manner by firms. By linking the seasonal release of information to return seasonality, we provide a simple and intuitive angle to understand this anomaly.

Our key finding is that the seasonality anomaly fits the pattern of an information cycle where investors demand a premium as compensation for the bearing the information uncertainty prior to a firm's news release. One might object this interpretation because for such a pattern to persist, it should be that investors are not able to sufficiently diversify away such risks, which appears unlikely. However, there is evidence that even institutions do not hold sufficiently diversified portfolios (e.g., see Bessembinder (2018)). A second concern might be that the pattern which fits an information cycle could also be consistent with investor misreaction, such as overreaction to earnings announcements. For example, Keloharju et al. (2019) argue that the twin patterns of positive returns in seasonal months and negative returns in non-seasonal months seems more consistent with mispricing than risk and Johnson, Kim, and So (2019) propose that analysts walk down estimates to create more beatable earnings forecasts that causes stock prices on average to rise in earnings announcement months. With regards to the manipulation explanation, we show that our results also hold in subsamples with few or no analyst coverage, where analyst walk-down is presumably less of an issue. Overall, we argue that the seasonality pattern lines up well with the direction and timing of returns expected by the information cycle, whereas for the mispricing explanation, whether one should expect investors to over or underreact to information releases a priori is not obvious.

The rest of the paper is organized as follows. Section 2 describes the information cycle and return seasonality. Section 3 reports summary statistics. Section 4 presents our main results, and Section 5 concludes.

2.2. Cross-sectional distribution of returns and information

In this section, we discuss the cross-sectional relation between seasonal returns and firms' information releases. Heston and Sadka (2008) and Keloharju et al. (2016) document seasonality by showing that when monthly cross-sectional returns are regressed on their sequential lagged-month returns, the estimated coefficients show a puzzling seasonal pattern. If a firm earned relatively lower (higher) returns in the same month over the past five, ten, or even twenty years, this firm continues to earn lower (higher) returns in the same calendar months in the future.

Another way to illustrate this anomalous pattern is to sort firms cross-sectionally based on their past same-month returns and check their future performance. Hence, each month, we sort stocks into quintiles based on their average return in the same calendar month in the last five years following Heston and Sadka (2008). Figure 1 plots the average monthly returns of seasonal loser-to-winner quintiles where 1 in the x-axis denotes the quintile of firms which earn the lowest returns in prior seasonal months, 2 denotes the quintile of firms which earn the second lowest returns, and so on. The y-axis denotes the average monthly returns of the firms in each quintile. All firms (CRSP ordinary shares) with non-missing quarterly earnings report dates in the last five years are included. If stock returns do not exhibit seasonality, the distribution of average monthly returns should be fairly uniform across seasonal losers to winners quintiles. However, as the figure shows, firms with relatively higher returns in the past seasonal months continue to earn higher returns in the future—the difference in average raw returns between firms with the highest seasonal returns and firms with the lowest seasonal returns is more than 0.5% per month.

[Figure 1 inserts here]

To observe whether firms' seasonal returns are related to the information cycle, we use Compustat's earnings announcements dates (item RDQ) as a simple proxy for information releases. A firm-month is defined as associated with seasonal information releases if the majority of its same calendar months in the past five years contain earnings announcements. Based on Frazzini and Lamont (2007) and Barber et al. (2013), firms earn higher returns during earnings announcement months than during non-announcement months. If earnings announcements are related with return seasonality, we should expect a larger fraction of firms with earnings announcements in the seasonal winner quintile compared to the seasonal loser quintile. Panel A of Figure 2 plots this relation.

Comparing the extreme quintiles, the plotted bars show that a higher fraction of seasonal winners are associated with prior information releases comparing to seasonal losers. In other words, firms in the seasonal winner quintile are more likely to release earnings information in their past seasonal months compared to firms in the seasonal loser quintile. Across the five quintiles, the pattern is slightly U-shaped with quintile 1 having a slightly higher fraction of information releases than the middle three quintiles.

[Figure 2 inserts here]

We also examine an alternative proxy for information releases. Earnings announcements represent scheduled information releases at the quarterly frequency. Chang, Hartzmark, Solomon, and Soltes (2017) show that not every earnings announcement is equally important for the firm. In addition, important periodic information releases might not always occur in earnings announcement months. We propose another proxy for information releases—the monthly change in option-implied volatility. In contrast, an increase in the implied volatility indicates the build-up of uncertainty. The change in implied volatility serves as a proxy for uncertainty reducing information releases by the firm. The disadvantage of this measure is that we do not know for sure if the uncertainty-reducing announcements are scheduled and more limited data availability (due to the need for traded options data).

Panel B shows the relation between average change in implied volatility for the return seasonality quintiles. We can see that seasonal winner firms are indeed associated with a reduction in implied volatility in their relevant seasonal months. This shows that uncertainty-reducing news for a firm tend to coincide with seasonal winner months and uncertainty-increases tend to coincide with seasonal loser months.

Together, Figure 2 provides visual evidence of the relation between the seasonality of stock returns and the likelihood of information releases.

We then examine how information uncertainty changes around earnings announcements. Figure 3 plots the average change in implied volatility around the earnings announcements at the monthly (Panel A) and weekly (Panel B) horizons. The sample includes all firms with available earnings announcement dates and option-implied volatilities around the dates. The x-axis denotes the event time where 0 represents the earnings announcement event. As shown in Panel A, information uncertainty, as proxied by implied volatility, builds up one month prior to the announcement and is resolved with the release of earnings information. After the announcement, uncertainty slowly builds up over time and again peaks at the month before the next scheduled announcement. In Panel B, the weekly pattern is very similar—there is a reduction of uncertainty at the earnings announcement event and an increase uncertainty in the non-announcement periods. The patterns we document here are consistent with the evidence in Gao, Ren, and Zhang (2020).

[Figure 3 inserts here]

This figure illustrates the information cycle of a typical firm—uncertainty builds up before the scheduled announcement, drops sharply around the announcement, and builds up again afterwards. Earnings announcements reveal the profitability of firms. In non-announcement periods, investors cannot fully back out how well the firm is doing and the posterior variance about earnings will be high. If firm-specific uncertainty matters to investors, for example if investors are not well-diversified, they might require compensation for this uncertainty due to uncertainty aversion. The asset price will then tend to be low prior to the announcement and be higher upon the announcement which resolves this uncertainty. In the macroeconomic announcement literature, the same sort of effects occur around macro announcements (e.g. Ai and Bansal (2018) and Hu et al. (2019)). While it seems logical that investors would care about the resolution of macroeconomic uncertainty, it is less obvious why investors would care about diversifiable firm-specific uncertainty. However, Frazzini and Lamont (2007) and Barber et al. (2013) document that prices of stocks indeed tend to go up in periods where earnings are scheduled to be announced.

We aim to relate the earnings announcement premium to the seasonality anomaly. Figure 4 plots the change in implied volatility for firms with (Panel B) and without (Panel A) earnings announcements. Firms with earnings announcements are defined as firms where the majority of the same calendar months in past five years were associated with earnings announcements. The implied volatility changes are measured for the period associated with the past seasonal months. Comparing seasonal winners and losers, we see that the most uncertainty resolution occurs for seasonal winners associated with announcements, and the most uncertainty build-up occurs for seasonal losers without announcements. Indeed, this is the pair that is consistent with the information cycle—i.e., seasonal winners associated with announcements and seasonal losers associated with no announcements—and we predict that this pair will exhibit the strongest seasonality effect.

[Figure 4 inserts here]

2.3. Data and summary statistics

Our sample covers all NYSE, NASDAQ, and Amex common stocks on the CRSP from 1971 to 2017. To measure seasonal information releases using earnings announcements, we require a firm to have non-missing earnings report dates (Compustat's item RDQ) for similar quarters in the last five years. This five-year requirement is used by Chang et al. (2017) and is also the minimum window used to observe seasonality in Keloharju et al. (2016).

To measure return seasonality, we form a measure *Return seasonality* which is the average of returns from 12, 24, 36, 48, and 60 months ago following Heston and Sadka (2008). This is the variable that will be used to sort firms into quintiles to determine whether firms that did well in same calendar month on average during the last five years also do better in the same month this year. We also report several firm characteristic, namely, $MOM_{t-12,t-2}$ is the month $t-12$ to $t-2$ buy-and-hold return, $Size$ is the firm's market capitalization in \$millions, B/M is the book-to-market ratio, and $Beta$ is the stock's market beta estimated from the last four years of monthly returns. Summary statistics of these variables are reported in Table 1.

[Table 1 inserts here]

Option-implied volatility data are from the WRDS Option Suites using the average of the interpolated implied volatility from puts and calls with 30 days to expiration and a delta of 50. The sample period of option data is from 1996 to 2017.

Data on stock prices come from the CRSP monthly stock file (ordinary shares). Data on the excess market return, risk-free rate, small-minus-big (SMB), high-minus-low (HML), robust-minus-weak (RMW), and conservative-minus-aggressive (CMA) portfolios are from Ken French's website. Table 1 shows the descriptive statistics of the main variables in our paper. In general, the summary statistics are consistent with those of prior research. Implied volatility data are available for about a quarter of the firm-month observations.

We use three measures in this paper to proxy for a firm's information cycle. Although public firms are required by regulations to publish their financial situations at fixed periodic intervals, they often file the earnings announcements earlier or later than the roughly-scheduled dates, and sometimes earnings reports can be cancelled. We check whether earnings are likely to be announced in the current month by examining whether the probability of observing an earnings announcement in the same calendar month over the past five years is greater than 50%. $Frac.RDQ$ is the fraction of earnings-announcement months during all the same-calendar months over the past five years. We include only observations where we have non-missing earnings announcements dates for the same fiscal quarter for the past five years. A current firm-month is defined as likely to be an information-release month if the majority of its past same-calendar months have had an earnings announcement. E.g., supposing that a firm's announcements of fourth quarter earnings were made in Feb one year ago, Feb two years ago, Feb three years ago, Mar four years ago, and Apr five years ago, then $Frac.RDQ$ for the current Feb will be 0.6, for Mar and Apr it will be 0.2. Hence, Feb is the only month marked as highly likely to be associated with a seasonal information release. Across, our sample, the mean of $Frac.RDQ$ is 0.34, consistent with the fact that about one-third of firms would have a quarterly earnings announcement in any given month.

As discussed earlier, another proxy for an information release is the change in option-implied volatility, $\Delta Implvol$. We define $\Delta Implvol$ as the month-end to month-end percentage change of daily option-implied volatility (e.g., if the implied volatility changes from 0.5 to 0.45, the percentage change is -10%). A drop in implied volatility occurs when implied volatility is larger at the beginning of the month than it is at the end of the month. Daily implied volatility data are from the WRDS Option Metrics, using the average of the interpolated implied volatility from puts and calls with 30 days to expiration and a delta of 50. Similar to the sample requirement for $Frac.RDQ$, we measure the seasonal pattern of uncertainty-reducing information releases by taking the average of the change in implied volatility in the same-calendar months over past five years. We require non-missing values for the past two years instead of the five years so as not to further reduce the already small Options Metrics sample. On average, the average monthly percentage change in implied volatility equals to 2.33%.

We also have a third proxy for information release which is used in a robustness test. This is the measure used in Chang et al. (2017) where they show that not every earnings announcement is equally important for a firm. Firms may announce remarkably higher or lower earnings in some quarters compared to a typical quarter due to its underlying business cycle. Following Chang et al. (2017), to capture the importance of a given quarter, we compute *Quarterly Earnings Rank* for each quarter q where we rank the prior five-year earnings from quarter $q-23$ to $q-4$ from largest to smallest based on its Compustat earnings per share excluding extraordinary items (item EPSPXQ). Then we calculate the average rank of quarters $q-4$, $q-8$, $q-12$, $q-16$, and $q-20$ as the seasonality rank of the upcoming earnings quarter. A rank of 1 or a smaller rank indicates a quarter when remarkably low earnings are expected. A higher rank indicates a quarter when remarkably high earnings are expected.

2.4. Results

2.4.1. Seasonal returns and earnings announcements

To test the role of the information cycle in explaining the seasonality anomaly, we sort stocks into quintiles at the end of each month based on their *Return seasonality* measure and hold stocks for one month. Stocks are equally weighted within each portfolio. We see in Panel A of Table 2 that a seasonal portfolio which longs seasonal winners and shorts seasonal losers (labelled "All Stocks") earns a hedged return of 0.68% per month and has high statistical significance ($t = 5.71$), confirming the results in prior studies that stocks exhibit a puzzling seasonal pattern to their returns.

Our main goal is to examine if the information cycle can explain the seasonality anomaly. To proxy for the likelihood of information releases in the portfolio holding month, we now split the stocks in each seasonal quintile into two. Stocks with prior seasonal information releases are those where the majority of the past seasonal months contain earnings announcements. This further sort results in 5×2 portfolios. High prior seasonal information release serves as a proxy for the information release likelihood in the portfolio holding month.

Panel C reports the average number of firms per month in each portfolio and indeed we see that the numbers of stocks in each group is pretty balanced where about one-third of the firms fall in the no information release groups.

We then form a seasonality portfolio that is *consistent* with the information cycle—i.e. long the seasonal winners with high information release likelihood and short losers with low information release likelihood. In Panel A of Table 2, colored in blue represent the average returns of these seasonality portfolios that are consistent with the information cycle and we see a return spread of 0.93% ($t = 7.28$) per month.

How about a portfolio that is *inconsistent* with the information cycle—i.e., one that longs seasonal winners with low information release likelihood and shorts losers with higher information release likelihood? We see that such a seasonality portfolio (average returns in green) earns hedged returns of only 0.43% which is significantly lower (t-statistic of the difference is 3.23) than the earlier return spread of 0.93%. Panel B reports similar results using Fama-French five factor alphas. Table 3 repeats the Table 2 analysis using value-weighted returns instead of equal-weighted returns and we see that the results are even stronger. Indeed the seasonality anomaly defined over winners and losers that are inconsistent with the information cycle is not even statistically significant whereas the information cycle-consistent seasonality portfolio earns a robust 1.14% per month alpha (Panel B of Table 3). In unreported results, we show that our results are also robust (and stronger in some cases) to sorts that use NYSE breakpoints instead of universal breakpoints.

[Table 2 and Table 3 insert here]

To look at these results from another angle. In month 0, we define a seasonal winner quintile which is the top quintile of stocks based on their average same calendar month returns in the past five years. We then hold this portfolio for 13 months from month 0 to month 12. Figure

5 plots the probability that stocks in this quintile remains winners quintile in future months. Panel A shows this for unconditional return seasonality. One can see that seasonal winner firms have a higher probability of being winners in month 0 (about 0.23 likelihood versus a baseline of 0.2). This is the seasonality anomaly. They also are more likely to be winners every three months (this is the earnings announcement premium). Finally, they are also much more likely to be winners in month 12, which is the second annual spike in returns after sorting. This overall U-shaped pattern is consistent with the information cycle explanation.

Next, we focus on the information-driven winners, which are firms in the seasonal winner quintile that have a majority of their past same calendar months associated with earnings announcements. Indeed, we see the sharp U-shaped pattern. In contrast, firms in the seasonal winners quintile that are not associated with an earnings announcement show a comparatively flatter pattern.

[Figure 5 inserts here]

To provide further visual evidence that the seasonality pattern is driven by firms whose returns are consistent with the information cycle, we show in Figure 6 the value-weighted buy-and-hold monthly cumulative returns to the unconditional return seasonality strategy (shown with a red dashed line), the info-cycle-driven seasonality strategy (blue line), and the counter info-cycle seasonality strategy (black line) from 1977–2017. One can clearly see that the information-cycle driven seasonality strategy outperforms the other two strategies.

[Figure 6 inserts here]

We believe that our evidence supports an information-cycle explanation for the seasonality anomaly, where seasonal winners have high returns because they are associated with a higher information release likelihood and seasonal losers have low returns because they are associated with a lower information release likelihood. The build-up of uncertainty in the low information release-likelihood months coincides with the lower returns and precedes an upcoming scheduled announcement.

This result is also related to the earnings announcement premium, where stocks expected to announce earnings earn higher returns (Frazzini and Lamont (2007) and Barber et al. (2013)). The months with high probability of information release are predicted to earn higher returns and this drives up the long side of the information-cycle consistent portfolio. The months with low probability of information release are not predicted to have high returns according to the earnings announcement premium and this drives down the short side of the information cycle-driven seasonality. Conversely, for the information cycle-inconsistent seasonality, seasonal winners with no predicted announcements have less kick and seasonal losers with predicted announcements get an earnings announcement premium boost hence denting spread of the seasonality strategy.

One natural question one might have after seeing these results is why the literature has not conducted such a test? The literature has indeed implemented tests which double sort seasonality portfolios into announcement and non-announcement months. Heston and Sadka (2008) for example do this very type of sorts. However, they typically then directly compare whether the seasonality anomaly in stronger in the information release group versus the non-information release group. Finding that the seasonality anomaly is strong in both groups, they conclude that the seasonal release of information does not explain the seasonality anomaly. We can see from our Tables 2 and 3 that when we implement these exact sorts, we also get the same results that seasonality is strong in stocks with prior seasonal information release *and* stocks with no prior information release (see Panel B where the alphas are 0.55% and 0.45% respectively). Our contribution is to identify seasonality that is *consistent* with the information cycle—information winners minus non-information losers—which is a cross-diagonal comparison indicated by portfolio return numbers in blue. Such cross-diagonal comparisons of extreme portfolios on two dimensions have also been used by papers such as Li et al (2020) and Jegadeesh et al. (2004).

2.4.2. Seasonal returns and change in option-implied volatility

The limitation of quarterly earnings announcements is that it occurs an a quarterly frequency while the seasonality we examine is in an annual frequency. In this section, we use the monthly change in option-implied volatility as an alternative proxy for information releases, where a decrease in option-implied volatility indicates information events which reduce firm uncertainty. The relation between option trading and news arrivals at both firm and market levels have been well documented (Chowdhry and Nanda (1991), An, Ang, Bali, and Cakici (2014), Bernile, Hu, and Tang (2016). The change in option-implied volatility can also be reasonably considered a measure of news arrivals (An et al. (2014)). Unlike earnings announcements which capture earnings news, this measure captures all types of firm information events that reduce uncertainty, e.g., news about products, management, or organizational restructuring.

We use the option-implied volatility reduction in the same-calendar months over past the five years to measure the stage of the information cycle that the seasonal returns are in. Because the availability of implied volatility data begins only in 1996, we require a minimum of two years of implied volatility data for inclusion (compared to requiring a minimum of five years when we use earnings announcements to measure information releases). Table 4 and Table 5 report respectively the equal-weighted and value-weighted returns of seasonality portfolios sorted on the average change in implied volatility in the seasonal months.

Similar to the earlier section, we build and compare the long-short portfolios from the diagonals to determine the profitability of seasonal strategies that are consistent versus inconsistent with the information cycle. In the beginning of each month, we *independently* sort stocks into seasonality quintiles and average uncertainty change ($\Delta impvol$) quintiles. Seasonality is the firm's average same-calendar return over five years (minimum of two years of past calendar month returns for inclusion), and $\Delta impvol$ is the average percentage change in month-end implied volatility over these five past seasonal calendar months. Because the availability of implied volatility data begins only in 1996, we require $\Delta impvol$ to be available for the most recent two years for inclusion (compared to requiring a minimum of all prior five years when we use earnings announcements to proxy for information releases). If the option-implied volatility on average declines in the same calendar month in the past years, this means that the firm tends to release information in the same calendar month and might be expected to do so in the current month. The low uncertainty change quintile contains firms with the most uncertainty-reducing information releases in their past seasonal months. In total there are 5×5 portfolios where each portfolio on average has about 50 stocks and the total number of firms per month is 1,619. The portfolios run from 1998 to 2017.

[Table 4 and Table 5 insert here]

From Table 4, we can see that return seasonality is not significant in this reduced sample—unconditionally, seasonal winners earn returns that are insignificantly different from that of seasonal losers. Because the options sample start from 1996, this coincides with the period where some anomalies have been documented to be weaker (McLean and Pontiff (2016)).

Although the unconditional return seasonality anomaly is not significant, interestingly, we see that seasonality consistent with the information cycle indeed earns positive and significant returns in both equal and value-weighted portfolios. In Table 4, *info-driven winners* outperform *noninfo-driven losers* by 55 basis points ($t = 2.06$) in equal-weighted portfolios. In contrast, the seasonality anomaly return and alpha defined counter to the information cycle are not significantly different from zero. The difference between the *info-cycle-driven seasonality* and *counter info-cycle seasonality* is significantly different from zero [1].

[1] The results are robust using simple difference rather than percentage change in option-implied volatility.

2.4.3. Seasonal returns and remarkable earnings announcements

In this section, we use the most important quarter of earnings as a proxy for information releases. Chang et al. (2017) show that not every quarter's earnings announcement is important—e.g., retail businesses might make most of their profits at the year end making Q4 the most important quarter for earnings. This tests whether return seasonality is stronger when the information cycle is defined based on the most important information release within a year.

Following Chang et al. (2017), we measure earnings seasonality by ranking five years of quarterly earnings and taking the average rank of prior five earnings announcements from the same fiscal quarter. A high (low) earnings rank means that the current quarter historically has higher (lower) earnings than the other quarters. Then we match this quarterly earnings rank to observations where the seasonal earnings-announcement month variable (($Frac.RDQ > 0.5$). Each month for this subsample, we divide stocks with seasonal announcements into quintile portfolios based on their earnings rank. Firms with "remarkable" earnings are defined as those in either the top and bottom quintiles which means that these are the most informative months (either very high earnings or very low earnings). Seasonality quintiles are defined independently. Results are presented in Tables 6 and 7.

[Table 6 and Table 7 insert here]

We show that return seasonality is stronger when it is consistent with the earnings announcement cycle where the cycle is demarcated by the quarters that have remarkable earnings. As shown in Table 6, seasonal winners associated with remarkable earnings (*remarkable info-driven winners*) earn the highest returns among the 5×2 portfolios, while seasonal losers with no earnings announcements (*noninfo-driven losers*) earn the lowest returns. The difference between *remarkable info-driven winners* and *noninfo-driven losers* is 116 basis points for equal-weighted portfolios and 123 basis points for value-weighted portfolios. In contrast, the counter remarkable-information-cycle seasonality is very weak—the raw return differential between seasonal winners with no earnings announcements and seasonal losers with remarkable earnings—is not significant in both equal-weighted and value-weighted portfolios.

Overall, we find evidence supporting the information cycle explanation for the seasonality anomaly. While there are multiple earnings announcements for a firm in a year, the historically most important earning announcements square up well with the seasonal pattern in returns.

2.5. Robustness tests

2.5.1. Cross-sectional regressions

We also report our results using cross-sectional Fama and MacBeth (1973) regressions to verify if the information-cycle explanation of return seasonality is robust to additional firm characteristics controls. We control for firm characteristics that have been shown to predict the cross section of stock returns, namely, firm size, short-term reversal (stock return in month $t-1$), book-to-market ratio, 1-year momentum (skipping a month), and market beta.

We use indicator explanatory variables obtained from independently sorting stocks based on their historical seasonal return and the stage of the information cycle. Then we investigate the return predictability of four intersecting portfolios, namely, information-driven winners, non information-driven losers, non information-driven winners, and information-driven losers. The regression is as follows:

$$\begin{aligned} Ret_{i,t} = & \alpha + \beta_1 Seasonal\ Winner_t \times Info_t + \beta_2 Seasonal\ Loser_t \times Non\_info_t \\ & + \beta_3 Seasonal\ Winner_t \times Non\_info_t + \beta_4 Seasonal\ Loser_t \times Info_t \\ & + \sum \beta_k X_{k,i,t-1} + \epsilon_{i,t-1} \end{aligned} \tag{2.1}$$

Table 8 presents the estimates from the Fama-MacBeth regressions of returns on the seasonal portfolios interacted with information cycle indicators. $Seasonal\ Winner_t(Loser_t)$ is a dummy variable that equals one for stocks in the top (bottom) quintile based on the average return of the same-calendar months over the past five years. $Seasonal\ Info_t(Non\_info_t)$ is a dummy variable that equals one if the majority of the same-calendar months over prior five years have (does not have) an earnings announcement.

We start the analysis by showing the existence of seasonality, as displayed in the first two columns. As $Seasonal\ Info_t$ and $Seasonal\ Non\_info_t$ are mutually exclusive, we only include $Seasonal\ Info_t$ in the regression. The analysis shows that seasonal winners earn higher returns and seasonal losers earn lower returns. Firms also earn higher returns in the months with seasonal earnings announcements ($Frac.RDQ > 0.5$) and this is the earnings announcement premium (Frazzini and Lamont (2007)). These results are similar after controlling firm characteristics. In column (3) and (4), the information-cycle-driven seasonality predicts return with and without controls. Seasonal winners generate 0.49% ($t = 5.36$) higher returns when they are associated with seasonal earnings announcements and seasonal losers generate 0.31% ($t = -4.19$) lower returns when they do not have announcements. The other two portfolios have weaker effects—seasonal winners with seasonal earnings announcements earn 0.19% ($t = 2.57$) higher returns, and seasonal losers with no seasonal earnings announcements cannot predict returns.

[Table 8 inserts here]

Table 9 reports regression results where the information cycle is measured by the seasonal change in option-implied volatility. $Seasonal\ Winner_t(Loser_t)$ is a dummy variable that equals one for stocks in the top (bottom) quintile based on the average return of the same-calendar months over past five years (minimum of two years of data for inclusion). $Seasonal\ Info_t(Non\_info_t)$ is a dummy variable that equals one for stocks in the quintile based on the lowest (highest) average change in implied volatility.

The first column shows that the return seasonality anomaly does not exist in this subsample, consistent with our earlier results that the seasonality anomlay is weak in this time period and small cross-section. The results are similar after controlling firm characteristics, except that seasonal losers now earn 0.2% lower returns, as shown in the second column. As before, we see that once we identify seasonality that is consistent with the information cyle, i.e. in Column (4), the seasonality anomaly become significant. *Info-driven winners* earn 0.43% higher returns while *noninfo-driven losers* generate 0.45% lower returns. The other two portfolios, *noninfo-driven winners* and *info-driven winners*, have no significant effect on returns.

[Table 9 inserts here]

Table 10 shows Fama-MacBeth regression results where the information cycle is measured by remarkable earnings. $Seasonal\ Winner_t(Loser_t)$ is a dummy variable that equals one for stocks in the top (bottom) quintile based on the average return of the firm in the same calendar month over the past five years. *Seasonal Remarkable* $Annc_t$ is a dummy variable that equals one for stocks having remarkable seasonal earnings announcements. $Seasonal\ Non\_annc$ is a dummy variable that equals one if the majority of the same-calendar months over prior five years does not have an earnings announcement.

Consistent with Heston and Sadka (2008), the first and second columns show that return seasonality is significant among firms with earnings announcements—seasonal winners (losers) predict higher (lower) returns. Consistent with the earnings announcement premium, remarkable earnings earn 0.27% higher returns, and months with no earnings announcements earn 0.26% lower returns. The last two columns show that return seasonality is stronger when it is consistent with the remarkable earnings information cycle. The results hold with and without controls. In column (4), seasonal winners earn 0.70% ($t = 5.56$) higher returns when they are associated with remarkable earnings announcements and seasonal losers earn 0.34% ($t = -5.16$) lower returns when they are associated with no earnings announcements.

[Table 10 inserts here]

2.5.2. Weekly returns

Keloharju et al. (2016) show that return seasonality does not only occur in the same calendar month every year, but is also evident in the same calendar week every year. Recent studies also show that the positive drift of returns usually concentrates in the short window surrounding earnings announcements (e.g., Johnson et al. (2019) and Linnainmaa and Zhang (2019)). If the information cycle can explain return seasonality at the monthly horizon, might our main results also hold when identifying seasonality based on returns in the same week in the past years?

We construct weekly return seasonality and information cycle following the manner in which we constructed monthly variables. To determine the number of weeks within a year, we set week 1 of the year as the week including both January 4th and the first Thursday of the year (Monday is considered the first day of the week). The maximum value of the number of week is 53. Weeks with the same week values are considered the same week across different years. A firm-week is defined as associated with seasonal information releases if the majority of its same weeks in the past five years contain quarterly earnings announcements.

[Table 11 and Table 12 insert here]

Table 11 and Table 12 show that stocks have strong return seasonality at the weekly horizon. An equal-weighted weekly seasonal portfolio which longs seasonal winners and shorts losers (labelled "All stocks") earns a hedge return of 21 basis points per month.[2] Similar to the earlier section, we build and compare two seasonality portfolios based on the information cycle. We find that the abnormal returns of the counter information-cycle seasonality strategy are

[2] Note that the "equal"-weighting here uses the method in Asparouhova, Bessembinder, and Kalcheva (2010) (i.e. one plus the lag return) to correct the upward bias of daily returns which were compounded into weekly returns for this analysis.

not statistically significant in either equal-weighted or value-weighted portfolios. In contrast, the seasonality portfolio consistent with the information cycle earns positive abnormal returns. Thus our weekly results are similar to our monthly horizon results.

2.5.3. Alternative expectations-management explanation

Johnson et al. (2019) show that firms have incentives to manage earnings expectations toward beatable levels before earnings announcements and that this will lead to earnings announcement premia and return seasonality. In this section, we examine whether the information cycle explanation is robust controlling firms' incentives of earnings expectations management.

We use the number of analysts covering a firm as reported by the I/B/E/S Summary History file as a measure for the likelihood of such earnings expectations manipulation. According to Johnson et al. (2019), analyst coverage is highly correlated with their Expectations Management Incentives (EMI) measure, which is computed from analyst coverage, institutional ownership, sales growth, and distress risk. We repeat our cross-diagonal comparisons across four partitions based on the monthly analyst coverage.

[Table 13 and Table 14 insert here]

Table 13 and Table 14 show the equal- and value-weighted strategies of info-cycle-driven seasonality, counter info-cycle seasonality, and the difference between the two strategies in the partitioned samples. The group with the highest analyst coverage are defined by firms with more than 11 analysts. Firms with between 7 to 11 analysts form the second group, then firms with between 3 to 6 analysts, and finally firms with less than 2 or zero analysts. Presumably, firms with lower analyst coverage and the firms uncovered by I/B/E/S database as a group might experience the lowest incentives to manage expectations. We show that our main results that information driven seasonality is stronger than counter-information cycle seasonality in all four analyst coverage groups and even in the lowest analyst coverage group. Hence, we believe that this alternative earnings expectation management story does not explain our findings.

2.5.4. Investor-mood explanation of return seasonality

Hirshleifer, Jiang, and Meng (2019) show that stocks' relative performance recur in high and low investor mood periods and can cause return seasonalities. Using January and March as a proxy for good mood months, and September and October as a proxy for bad mood months, they show that assets that tend to do well in good mood months will continue to do well in future good mood months (seasonality) but will do worse in bad mood months (reversals).

In unreported tests, we add the investor mood variable, which is the average returns from the past January and March, and average return from the past September and October as a control variable in our cross-sectional regressions. We show that our results are robust to such controls.

2.6. Conclusion

Return seasonality is a robust anomaly that to date has no satisfactory explanation. In this paper, we propose an information-cycle explanation for the seasonality anomaly. We sort stocks monthly based on their average return in the same-calendar months in the past five years and the stage of their information cycle in those past seasonal months to proxy for whether the upcoming month is likely to contain an information release or not. Using past quarterly earnings announcements or past declines in option-implied volatility as proxies for information releases, we find that seasonality anomaly is much stronger when it is consistent with information cycle, i.e. seasonal winners in information-release months and seasonal losers in non-information-release months drive the seasonality anomaly. These return patterns are evidence that the seasonality anomaly fits the pattern of uncertainty dissolution during the scheduled announcement of information and investors demanding a premium for holding stocks that contain such information uncertainty.

Table 1: Summary Statistics

This table reports summary statistics of the main variables used in the study. Return seasonality is the average of monthly returns from 12, 24, 36, 48, and 60 months ago, where we require all of these five past months of data to be available before inclusion in the sample (Heston and Sadka (2008)). $Frac.RDQ$ is average fraction that these past seasonal months are associated with earnings announcement (checking Compustat item RDQ, where we require non-missing observations for all prior five similar fiscal quarters). Implied volatility ($Implvol$) is from the WRDS option Metrics using the average of the interpolated implied volatility from puts and calls with 30 days to expiration and a delta of 50. $\Delta Implvol$ is monthly change in option-implied volatility, which is defined as the month-end to month-end percentage change of daily option-implied volatility (e.g., if the implied volatility changes from 0.5 to 0.45, the percentage change is -10%). At least two past years of non-missing implied volatility data are required for this computation. Quarterly earnings rank is formed by ranking five years of quarterly earnings and taking the average rank of prior five earnings from the same fiscal quarter (Chang et al. (2017)). The Compustat item used is earnings per share excluding extraordinary items, EPSPXQ. $MOM_{t-12,t-2}$ is the month $t-12$ to $t-2$ buy-and-hold return, $Size$ is the firm's market capitalization, B/M is the book-to-market ratio, β is the stock's market beta based on four years of past monthly returns. Compustat earnings data begin in 1971 so that available CRSP sample period for computing seasonality is from January 1976 to December 2017. Option Metrics data start only in 1996.

Variable	Mean	Std dev	P10	P25	Median	P75	P90	N
Frac.RDQ	0.34	0.42	0	0	0	0.8	0.8	1,259,558
Implvol	0.45	0.24	0.21	0.28	0.39	0.54	0.75	385,743
ΔImplvol(%)	2.33	27.10	-20.33	-10.33	-0.46	10.91	25.45	385,743
Quarterly Earnings Rank	10.61	2.79	7.00	8.80	10.61	12.40	14.00	399,739
Return seasonality (%)	1.47	6.98	-5.86	-2.21	1.11	4.57	8.83	1,240,370
MOM (t-12, t-2)	0.15	0.65	-0.38	-0.15	0.08	0.32	0.67	1,241,790
Size (million)	3,143	15,900	18	59	271	1,275	4,864	1,247,212
B/M	0.72	2.29	0.18	0.35	0.60	0.96	1.45	1,241,924
Beta	1.11	0.76	0.29	0.62	1.02	1.47	2.01	1,242,640

Table 2: Equal-weighted seasonal portfolios double sorted on information release likelihood proxied by prior earnings announcements

This table reports equal-weighted returns and alphas of portfolios sorted by prior seasonal returns and information release likelihood. Each month, we independently sort stocks into quintiles based on their average return in the same calendar month in the last five years, and into two groups based on whether the majority of these past seasonal months are associated with information releases. Information releases are defined as quarterly earnings announcements (Compustat item RQD). We define *Information-cycle-driven seasonality* as a hedge portfolio with a long position in seasonal winners with prior information releases and a short position in seasonal losers with no prior seasonal information releases. In contrast, *Counter information-cycle seasonality* is a hedge portfolio with a long position in seasonal winners with no prior seasonal information releases and a short position in seasonal losers with prior seasonal information releases. We hold stocks for one month in each portfolio. Panel A reports raw portfolio returns, Panel B reports Fama and French (2015) five-factor alphas, and Panel C reports the average number of firms in a typical month. t-statistics are in brackets and the portfolios hold stocks from 1976–2017.

Panel A: Raw Return (%)

	Prior Seasonal Return					
Seasonal Portfolio	1 (loser)	2	3	4	5 (winner)	5-1 Diff
All Stocks	1.09	1.19	1.35	1.49	1.77	0.68
	[4.05]	[5.47]	[6.29]	[6.72]	[6.36]	[5.71]
Stocks with prior seasonal info release	1.22	1.46	1.59	1.69	1.96	0.73
	[4.17]	[6.22]	[6.79]	[7.27]	[7.14]	[4.57]
Stocks with no prior seasonal info release	1.03	1.11	1.26	1.42	1.65	0.62
	[3.86]	[4.99]	[5.85]	[6.25]	[5.73]	[4.99]
Info-cycle-driven seasonality						0.93
						[7.28]
Counter info-cycle seasonality						0.43
						[2.62]
Info - Counter info						0.51
						[3.23]

Panel B: Fama-French five-factor alphas (%)

	Prior Seasonal Return					
Seasonal Portfolio	1 (loser)	2	3	4	5 (winner)	5-1 Diff
All stocks	-0.20	-0.10	0.06	0.25	0.61	0.81
	[-2.19]	[-1.80]	[1.07]	[3.82]	[5.62]	[6.75]
Stocks with prior seasonal info release	-0.11	0.18	0.27	0.45	0.75	0.86
	[-0.08]	[1.74]	[3.17]	[4.69]	[6.53]	[5.21]
Stocks with no prior seasonal info release	-0.25	-0.19	-0.02	0.16	0.51	0.75
	[-2.62]	[-2.92]	[-0.24]	[2.26]	[4.22]	[5.95]
Info-cycle-driven seasonality						1.00
						[7.52]
Counter info-cycle seasonality						0.62
						[3.70]
Info - Counter info						0.37
						[2.28]

Panel C: Average number of firms in each portfolio per month

	Prior Seasonal Return				
Seasonal Portfolio	1 (loser)	2	3	4	5 (winner)
Stocks with prior seasonal info release	173	157	157	167	192
Stocks with no prior seasonal info release	333	348	348	338	312

Table 3: Value-weighted seasonal portfolios double sorted on information release likelihood proxied by prior earnings announcements

This table reports value-weighted returns and alphas of portfolios sorted by prior seasonal returns and information releases. Each month, we independently sort stocks into quintiles based on their average return in the same calendar month in the last five years, and into two groups based on whether the majority of these past seasonal months are associated with information releases. Information releases are defined as quarterly earnings announcements (Compustat item RQD). We define *Information-cycle-driven seasonality* as a hedge portfolio with a long position in seasonal winners with prior information releases and a short position in seasonal losers with no prior seasonal information releases. In contrast, *Counter information-cycle seasonality* is a hedge portfolio with a long position in seasonal winners with no prior seasonal information releases and a short position in seasonal losers with prior seasonal information releases. We hold stocks for one month in each portfolio. Panel A reports raw portfolio returns, Panel B reports Fama and French (2015) five-factor alphas, and Panel C reports the average number of firms in a typical month. t-statistics are in brackets and the portfolios hold stocks from 1976–2017.

Panel A: Raw Return (%)

	Prior Seasonal Return					
Seasonal Portfolio	1 (loser)	2	3	4	5 (winner)	5-1 Diff
All Stocks	0.82	0.91	1.03	1.02	1.26	0.44
	[3.47]	[4.60]	[5.48]	[5.16]	[5.15]	[2.72]
Stocks with prior seasonal info release	1.21	1.20	1.48	1.30	1.71	0.50
	[4.13]	[5.22]	[6.61]	[5.64]	[6.50]	[2.11]
Stocks with no prior seasonal info release	0.70	0.80	0.88	0.96	1.02	0.33
	[2.87]	[3.78]	[4.44]	[4.61]	[3.98]	[1.96]
Info-cycle-driven seasonality						1.02
						[5.63]
Counter info-cycle seasonality						-0.19
						[-0.80]
Info - Counter info						1.22
						[4.88]

Panel B: Fama-French five-factor alphas (%)

	Prior Seasonal Return					
Seasonal Portfolio	1 (loser)	2	3	4	5 (winner)	5-1 Diff
All stocks	-0.28	-0.28	-0.06	-0.03	0.27	0.54
	[-2.71]	[-4.39]	[-1.39]	[-0.56]	[2.85]	[3.37]
Stocks with prior seasonal info release	0.15	0.04	0.35	0.18	0.69	0.55
	[0.75]	[0.31]	[2.83]	[1.48]	[5.32]	[2.28]
Stocks with no prior seasonal info release	-0.45	-0.39	-0.22	-0.08	0.00	0.45
	[-4.20]	[-4.74]	[-3.31]	[-1.04]	[0.04]	[2.69]
Info-cycle-driven seasonality						1.14
						[6.26]
Counter info-cycle seasonality						-0.14
						[-0.58]
Info - Counter info						1.29
						[4.94]

Panel C: Average number of firms in each portfolio per month

	Prior Seasonal Return				
Seasonal Portfolio	1 (loser)	2	3	4	5 (winner)
Stocks with prior seasonal info release	173	157	157	167	192
Stocks with no prior seasonal info release	333	348	348	338	312

Table 4: Equal-weighted seasonal portfolios double sorted on information release likelihood proxied by prior implied volatility changes

This table reports equal-weighted returns and alphas of portfolios sorted by prior seasonal returns and uncertainty-reducing information releases proxied by implied volatility changes $\Delta impvol$. Each month, we independently sort stocks into quintiles based on their average return in the same calendar month in the last five years, and into quintiles based on their average option-implied volatility change in these past seasonal months. We require a minimum of two years of implied volatility data for inclusion. We define $\Delta impvol$ as the month-end to month-end percentage change of daily option-implied volatility. *Information-cycle-driven seasonality* is a hedge portfolio with a long position in seasonal winners with prior information releases and a short position in seasonal losers with no prior information releases. In contrast, *Counter information-cycle seasonality* is a hedge portfolio with a long position in seasonal winners with no prior information releases and a short position in seasonal losers with prior information releases. We hold stocks for one month in each portfolio. Panel A reports raw portfolio results, Panel B reports Fama and French (2015) five-factor alphas, and Panel C reports the average number of firms in a typical month. t-statistics are in brackets and the portfolios hold stocks from 1998–2017.

Table 4 (continued)

Panel A: Raw Return							
		Prior Seasonal Return					
Seasonal Portfolio		1 (loser)	2	3	4	5 (winner)	5-1 Diff
All stocks		1.24	1.10	1.09	1.14	1.35	0.11
		[2.74]	[3.04]	[3.12]	[3.09]	[2.88]	[0.52]
Δ Implied Volatility	1 (Low)	1.43	1.27	1.42	1.19	1.58	0.14
		[3.01]	[3.31]	[3.59]	[3.19]	[3.34]	[0.53]
	2	1.40	1.17	1.09	1.01	1.42	0.02
		[3.04]	[3.08]	[3.11]	[2.78]	[3.06]	[0.07]
	3	1.31	1.07	1.07	1.15	0.98	-0.33
		[2.82]	[2.94]	[3.02]	[2.94]	[2.14]	[-1.34]
	4	1.27	0.97	0.89	1.30	1.48	0.22
		[2.63]	[2.68]	[2.53]	[3.43]	[2.93]	[0.77]
	5 (High)	1.03	1.12	1.01	1.06	1.15	0.12
		[2.23]	[2.84]	[2.65]	[2.57]	[2.11]	[0.44]
Info-cycle-driven seasonality							0.55
							[2.06]
Counter Info-cycle seasonality							-0.28
							[-0.87]
Info - Counter info							0.83
							[2.31]

Panel B: Fama-French five-factor alphas							
		Prior Seasonal Return					
Seasonal Portfolio		1 (loser)	2	3	4	5 (winner)	5-1 Diff
All stocks		0.19	0.07	0.09	0.20	0.45	0.27
		[1.27]	[0.58]	[0.85]	[1.68]	[2.64]	[1.37]
Δ Implied Volatility	1 (Low)	0.39	0.27	0.43	0.27	0.68	0.29
		[1.84]	[1.44]	[2.09]	[1.72]	[3.32]	[1.06]
	2	0.33	0.12	0.14	0.11	0.51	0.19
		[1.66]	[0.75]	[1.07]	[0.73]	[2.48]	[0.74]
	3	0.27	0.03	0.04	0.13	0.05	-0.22
		[1.47]	[0.25]	[0.26]	[0.82]	[0.25]	[-0.89]
	4	0.16	-0.05	-0.14	0.33	0.63	0.48
		[0.84]	[-0.33]	[-1.14]	[2.26]	[2.61]	[1.70]
	5 (High)	0.01	0.07	0.00	0.15	0.23	0.21
		[0.07]	[0.41]	[0.02]	[0.89]	[0.88]	[0.78]
Info-cycle-driven seasonality							0.67
							[2.50]
Counter Info-cycle seasonality							-0.16
							[-0.49]
Info - Counter info							0.83
							[2.23]

Panel C: Average number of firms in each portfolio per month						
		Prior Seasonal Return				
Seasonal Portfolio		1 (loser)	2	3	4	5 (winner)
	1 (Low)	47	51	59	71	95
	2	53	61	67	72	72
Δ Implied Volatility	3	61	67	69	66	60
	4	73	73	68	61	50
	5 (High)	93	71	60	53	46

Table 5: Value-weighted seasonal portfolios double sorted on information release likelihood proxied by prior implied volatility changes

This table reports value-weighted returns and alphas of portfolios sorted by prior seasonal returns and uncertainty-reducing information release proxied by implied volatility changes $\Delta impvol$. Each month, we independently sort stocks into quintiles based on their average return in the same calendar month in the last five years, and into quintiles based on their average option-implied volatility change in these past seasonal months. We require a minimum of two years of implied volatility data for inclusion. We define $\Delta impvol$ as the month-end to month-end percentage change of daily option-implied volatility. *Information-cycle-driven seasonality* is a hedge portfolio with a long position in seasonal winners with prior information releases and a short position in seasonal losers with no prior information releases. In contrast, *Counter information-cycle seasonality* is a hedge portfolio with a long position in seasonal winners with no prior information releases and a short position in seasonal losers with prior information releases. We hold stocks for one month in each portfolio. Panel A reports raw portfolio results, Panel B reports Fama and French (2015) five-factor alphas, and Panel C reports the average number of firms in a typical month. t-statistics are in brackets and the portfolios hold stocks from 1998–2017.

Table 5 (continued)

Panel A: Raw Return							
		Prior Seasonal Return					
Seasonal Portfolio		1 (loser)	2	3	4	5 (winner)	5-1 Diff
All stocks		0.82	0.70	0.67	0.65	0.84	0.02
		[2.24]	[2.51]	[2.44]	[2.18]	[2.20]	[0.07]
Δ Implied Volatility	1 (Low)	0.67	0.65	0.92	0.81	1.07	0.40
		[1.50]	[1.78]	[2.96]	[2.44]	[2.59]	[1.08]
	2	0.77	1.03	0.64	0.38	0.72	-0.04
		[1.98]	[3.46]	[1.98]	[1.17]	[1.77]	[-0.12]
	3	0.75	0.67	0.79	0.81	0.54	-0.22
		[1.94]	[2.19]	[2.56]	[2.54]	[1.28]	[-0.62]
	4	1.18	0.71	0.67	0.82	0.87	-0.32
		[2.85]	[2.25]	[2.25]	[2.55]	[1.96]	[-0.86]
	5 (High)	0.86	0.54	0.63	0.80	0.87	0.01
		[2.01]	[1.62]	[1.88]	[2.05]	[1.85]	[0.03]
Info-cycle-driven seasonality							0.22
							[0.64]
Counter Info-cycle seasonality							0.20
							[0.49]
Info - Counter info							0.01
							[0.03]

Panel B: Fama-French five-factor alphas							
		Prior Seasonal Return					
Seasonal Portfolio		1 (loser)	2	3	4	5 (winner)	5-1 Diff
All stocks		0.09	-0.09	-0.09	-0.06	0.18	0.09
		[0.55]	[-1.13]	[-1.30]	[-0.71]	[1.28]	[0.36]
Δ Implied Volatility	1 (Low)	-0.24	-0.16	0.19	-0.02	0.43	0.67
		[-0.80]	[-0.68]	[1.05]	[-0.10]	[1.97]	[1.74]
	2	-0.03	0.23	0.13	-0.30	0.08	0.11
		[-0.14]	[1.27]	[-0.76]	[-1.69]	[0.35]	[0.32]
	3	0.11	-0.18	-0.08	0.01	-0.20	-0.31
		[0.45]	[-1.22]	[-0.49]	[0.06]	[-0.87]	[-0.85]
	4	0.40	-0.13	-0.14	0.01	0.18	-0.22
		[1.66]	[-0.88]	[-1.01]	[0.08]	[0.70]	[-0.59]
	5 (High)	0.10	-0.29	-0.09	-0.01	0.07	-0.03
		[0.44]	[-1.65]	[-0.52]	[-0.05]	[0.25]	[-0.09]
Info-cycle-driven seasonality							0.32
							[0.95]
Counter Info-cycle seasonality							0.31
							[0.73]
Info - Counter info							0.01
							[0.03]

Panel C: Average number of firms in each portfolio per month						
		Prior Seasonal Return				
Seasonal Portfolio		1 (loser)	2	3	4	5 (winner)
	1 (Low)	47	51	59	71	95
	2	53	61	67	72	72
Δ Implied Volatility	3	61	67	69	66	60
	4	73	73	68	61	50
	5 (High)	93	71	60	53	46

Table 6: Equal-weighted seasonal portfolios double sorted on whether prior earnings quarters are remarkable

This table reports equal-weighted returns and alphas of portfolios sorted by prior seasonal returns and information-release importance. Each month, we independently sort stocks into quintiles based on their average return in the same calendar months during the last five years, and into two groups based on whether the majority of the past seasonal months are associated with remarkable earnings announcements. Firms with remarkable earnings are defined as those with seasonal announcements in either the top and bottom quintiles of their quarterly earnings rank defined in Chang et al. (2017). We define *Remarkable info-cycle-driven seasonality* as a hedge portfolio with a long position in seasonal winners with prior remarkable earnings announcements and a short position in seasonal losers with no earnings announcements. In contrast, *Counter remarkable info-cycle seasonality* is a hedge portfolio with a long position in seasonal winners with no earnings announcements and a short position in seasonal losers with prior remarkable earnings announcements. We hold stocks for one month in each portfolio. Panel A reports raw portfolio results, Panel B reports Fama and French (2015) five-factor alphas, and Panel C reports the average number of firms in a typical month. *t*-statistics are in brackets and the portfolios hold stocks from 1976–2017.

Panel A: Raw Return (%)

	Prior Seasonal Return					
Seasonal Portfolio	1 (loser)	2	3	4	5 (winner)	5-1 Diff
All stocks	1.30	1.51	1.60	1.70	1.96	0.65
	[4.44]	[6.37]	[6.80]	[7.38]	[7.19]	[4.10]
Stocks with remarkable info release	1.47	1.49	1.39	1.71	2.19	0.73
	[4.87]	[5.90]	[5.86]	[7.49]	[7.97]	[3.67]
Stocks with no prior seasonal info release	1.03	1.07	1.24	1.36	1.65	0.62
	[3.83]	[4.75]	[5.68]	[5.99]	[5.70]	[5.05]
Remarkable info-cycle seasonality						1.16
						[7.63]
Counter remarkable-info-cycle seasonality						0.18
						[0.98]
Remarkable - Counter remarkable info						0.97
						[4.45]

Panel B: Fama-French five-factor alphas (%)

	Prior Seasonal Return					
Seasonal Portfolio	1 (loser)	2	3	4	5 (winner)	5-1 Diff
All stocks	-0.04	0.21	0.28	0.46	0.71	0.75
	[-0.29]	[1.96]	[3.20]	[4.72]	[6.30]	[4.59]
Stocks with remarkable info release	0.15	0.19	0.09	0.49	0.93	0.78
	[0.84]	[1.33]	[0.73]	[4.24]	[6.40]	[3.79]
Stocks with no prior seasonal info release	-0.23	-0.21	-0.02	0.13	0.53	0.76
	[-2.43]	[-3.16]	[-0.27]	[1.84]	[4.42]	[6.05]
Remarkable info-cycle seasonality						1.16
						[7.63]
Counter remarkable-info-cycle seasonality						0.18
						[0.98]
Remarkable - Counter remarkable info						0.97
						[4.45]

Panel C: Average number of firms in each portfolio per month

	Prior Seasonal Return				
Seasonal Portfolio	1 (loser)	2	3	4	5 (winner)
Stocks with remarkable info release	65	63	64	66	71
Stocks with no prior seasonal info release	325	330	331	321	295

Table 7: Value-weighted seasonal portfolios double sorted on whether prior earnings quarters are remarkable

This table reports value-weighted returns and alphas of portfolios sorted by prior seasonal returns and information-release importance. Each month, we independently sort stocks into quintiles based on their average return in the same calendar months during the last five years, and into two groups based on whether the majority of the past seasonal months are associated with remarkable earnings announcements. Firms with remarkable earnings are defined as those with seasonal announcements in either the top and bottom quintiles of their quarterly earnings rank defined in Chang et al. (2017). We define *Remarkable info-cycle-driven seasonality* as a hedge portfolio with a long position in seasonal winners with prior remarkable earnings announcements and a short position in seasonal losers with no earnings announcements. In contrast, *Counter remarkable info-cycle seasonality* is a hedge portfolio with a long position in seasonal winners with no earnings announcements and a short position in seasonal losers with prior remarkable earnings announcements. We hold stocks for one month in each portfolio. Panel A reports raw portfolio results, Panel B reports Fama and French (2015) five-factor alphas, and Panel C reports the average number of firms in a typical month. t-statistics are in brackets and the portfolios hold stocks from 1976–2017.

Panel A: Raw Return (%)

	Prior Seasonal Return					
Seasonal Portfolio	1 (loser)	2	3	4	5 (winner)	5-1 Diff
All stocks	1.22	1.22	1.46	1.32	1.65	0.44
	[4.18]	[5.19]	[6.50]	[5.77]	[6.35]	[1.79]
Stocks with remarkable info release	1.35	1.38	1.39	1.62	1.87	0.52
	[4.73]	[5.65]	[5.86]	[7.05]	[6.60]	[1.99]
Stocks with no prior seasonal info release	0.64	0.77	0.87	0.93	1.01	0.37
	[2.64]	[3.57]	[4.33]	[4.39]	[3.92]	[2.26]
Remarkable info-cycle seasonality						1.23
						[5.67]
Counter remarkable-info-cycle seasonality						-0.34
						[-1.43]
Remarkable - Counter remarkable info						1.57
						[5.28]

Panel B: Fama-French five-factor alphas (%)

	Prior Seasonal Return					
Seasonal Portfolio	1 (loser)	2	3	4	5 (winner)	5-1 Diff
All stocks	0.17	0.04	0.34	0.20	0.60	0.44
	[0.86]	[0.28]	[2.75]	[1.61]	[4.78]	[1.79]
Stocks with remarkable info release	0.11	0.19	0.37	0.52	0.75	0.64
	[0.53]	[1.06]	[2.25]	[3.53]	[3.98]	[2.38]
Stocks with no prior seasonal info release	-0.49	-0.40	-0.20	-0.08	0.00	0.49
	[-4.63]	[-4.68]	[-2.95]	[-1.10]	[0.02]	[2.95]
Remarkable info-cycle seasonality						1.24
						[5.54]
Counter remarkable-info-cycle seasonality						-0.10
						[-0.43]
Remarkable - Counter remarkable info						1.34
						[4.35]

Panel C: Average number of firms in each portfolio per month

	Prior Seasonal Return				
Seasonal Portfolio	1 (loser)	2	3	4	5 (winner)
Stocks with remarkable info release	65	63	64	66	71
Stocks with no prior seasonal info release	325	330	331	321	295

Table 6: Fama-MacBeth regressions with information releases proxied by earnings announcements

This table reports Fama and MacBeth (1973) regressions of monthly stock returns on seasonality interacted with prior information release indicators. $Seasonal\ Winner_t(Loser_t)$ is a dummy variable that equals one for a stock in the top (bottom) quintile based on the average return in the same calendar month in the prior five years. $Seasonal\ Annc_t(Non\_annc_t)$ is a dummy variable that equals one if the majority of these past seasonal months have (do not have) earnings announcements. Earnings announcements are defined using Compustat item RDQ. $MOM_{t-12,t-2}$ is the stock's month $t-12$ to $t-2$ buy-and-hold return. $Size$ is the firm's market capitalization. B/M is the book-to-market ratio. $Beta$ is the stock's market beta based on four years of past monthly returns. The reported t-statistics account for heteroskedasticity and autocorrelation using Newey and West (1987) correction with 12 lags, with *, **, and *** indicating statistical significance at the 10%, 5%, and 1% levels respectively. The cross-sectional regressions are estimated from January 1976 to December 2017.

	Monthly Return$_t$			
	(1)	(2)	(3)	(4)
Seasonal Winner$_t$	0.415***	0.298***		
	(4.072)	(3.904)		
Seasonal Earnings Annc$_t$	0.288***	0.363***		
	(5.316)	(6.224)		
Seasonal Loser$_t$	-0.236**	-0.259***		
	(-2.145)	(-3.687)		
Seasonal Winner$_t$ × Annc$_t$			0.596***	0.488***
			(5.352)	(5.357)
Seasonal Loser$_t$ × Non_annc$_t$			-0.273**	-0.313***
			(-2.458)	(-4.190)
Seasonal Winner$_t$ × Non_annc$_t$			0.323***	0.190**
			(2.961)	(2.574)
Seasonal Loser$_t$ × Annc$_t$			-0.167	-0.142
			(-1.067)	(-1.250)
Monthly Return$_{t-1}$		-0.055***		-0.055***
		(-11.106)		(-11.322)
MOM$_{t-12,t-1}$		0.444***		0.424**
		(2.602)		(2.559)
Ln(Size)$_{t-1}$		-0.161***		-0.150***
		(-2.963)		(-2.936)
Ln(B/M)$_{t-1}$		0.146**		0.152**
		(2.063)		(2.140)
Beta$_{t-1}$		0.075		0.073
		(0.563)		(0.542)
Constant	1.229***	3.034***	1.310***	3.017***
	(5.763)	(4.375)	(6.200)	(4.511)
Observations	1,201,341	1,201,341	1,201,341	1,201,341
R-squared	0.013	0.062	0.013	0.062

Table 7: Fama-MacBeth regressions with information releases proxied by implied volatility changes

This table reports Fama and MacBeth (1973) regressions of monthly stock returns on seasonality interacted with indicators of prior uncertainty-reducing information releases. $Seasonal\ Winner_t (Loser_t)$is a dummy variable that equals one for a stock in the top (bottom) quintile based on the average return in the same calendar month in prior five years. $Seasonal\ Info_t$ (Non_info_t) is a dummy variable that equals one for a stock in the bottom (top) quintile based on the average change in implied volatility ($\Delta impvol$) in the same calendar month in prior five years. We require a minimum of two years of implied volatility data for inclusion. We define $\Delta impvol$ as the last available observation minus the first available observation of a firm's daily implied volatility within a month. $MOM_{t-12,t-2}$ is the stock's month $t-12$ to $t-2$ buy-and-hold return. $Size$ is the firm's market capitalization. B/M is the book-to-market ratio. $Beta$ is the stock's market beta based on four years of past monthly returns. The t-statistics account for heteroskedasticity and autocorrelation using Newey and West (1987) correction with 12 lags, with *, **, and *** indicating statistical significance at the 10%, 5%, and 1% levels respectively. The cross-sectional regressions are estimated from January 1998 to December 2017.

	Monthly Return$_t$			
	(1)	(2)	(3)	(4)
Seasonal Winner$_t$	0.230	0.231		
	(1.166)	(1.498)		
Seasonal Info$_t$($Low\ \Delta implvol$)	0.360***	0.143*		
	(2.723)	(1.897)		
Seasonal Loser$_t$	0.058	-0.200**		
	(0.448)	(-2.214)		
Seasonal Non_info$_t$($High\ \Delta implvol$)	0.085	-0.101		
	(0.765)	(-1.214)		
Seasonal Winner$_t$ × Info$_t$			0.508*	0.431**
			(1.944)	(2.401)
Seasonal Loser$_t$ × Non_info$_t$			-0.188	-0.450***
			(-1.146)	(-3.143)
Seasonal Winner$_t$ × Non_info$_t$			0.380	0.183
			(1.539)	(1.149)
Seasonal Loser$_t$ × Info$_t$			0.356	-0.042
			(1.614)	(-0.267)
Monthly Return$_{t-1}$		-0.020***		-0.020***
		(-2.763)		(-2.744)
MOM$_{t-12,t-1}$		-0.349		-0.345
		(-0.786)		(-0.771)
Ln(Size)$_{t-1}$		-0.239***		-0.239***
		(-3.256)		(-3.285)
Ln(B/M)$_{t-1}$		0.007		0.002
		(0.073)		(0.021)
Beta$_{t-1}$		0.053		0.061
		(0.238)		(0.265)
Constant	1.046***	4.385***	1.145***	4.390***
	(3.464)	(3.779)	(3.514)	(3.816)
Observations	300,570	300,570	300,570	300,570
R-squared	0.016	0.087	0.012	0.086

Table 10: Fama-MacBeth regressions with information releases proxied by whether prior earnings quarters are remarkable

This table reports Fama and MacBeth (1973) regressions of monthly stock returns on seasonality interacted with importance of prior information releases. $Seasonal\ Winner_t (Loser_t)$is a dummy variable that equals one for a stock in the top (bottom) quintile based on the average return in the same calendar month in prior five years. Each month, we divide stocks with seasonal announcements into quintile portfolios based on their earnings rank (Chang et al. (2017)). Firms with remarkable earnings are defined as those in either the top and bottom quintiles, while firms with unremarkable earnings as those in the middle three quintiles. $Seasonal\ Remarkable\ Info_t$ is a dummy variable that equals one for the month with seasonal remarkable earnings. $Seasonal\ Unremarkable\ Info_t$ is a dummy that equals one the month with seasonal unremarkable earnings. $MOM_{t-12,t-2}$ is the stock's month $t-12$ to $t-2$ buy-and-hold return. $Size$ is the firm's market capitalization. B/M is the book-to-market ratio. $Beta$ is the stock's market beta based on four years of past monthly returns. The t-statistics account for heteroskedasticity and autocorrelation using Newey and West (1987) correction with 12 lags, with *, **, and *** indicating statistical significance at the 10%, 5%, and 1% levels respectively. The cross-sectional regressions are estimated from January 1976 to December 2017.

	Monthly Return$_t$			
	(1)	(2)	(3)	(4)
Seasonal Winner$_t$	0.384***	0.300***		
	(3.770)	(3.932)		
Seasonal Remarkable Annc$_t$	0.142*	0.275**		
	(1.962)	(2.284)		
Seasonal Loser$_t$	-0.243**	-0.260***		
	(-2.271)	(-3.903)		
Seasonal Non_Annc$_t$	-0.239***	-0.262***		
	(-3.568)	(-4.421)		
Seasonal Winner$_t$ × Remarkable Annc$_t$			0.803***	0.695***
			(6.272)	(5.556)
Seasonal Loser$_t$ × Non_annc$_t$			-0.307***	-0.344***
			(-3.017)	(-5.157)
Seasonal Winner$_t$ × Non_annc$_t$			0.294***	0.165***
			(2.928)	(2.656)
Seasonal Loser$_t$ × Remarkable Annc$_t$			0.104	0.069
			(0.546)	(0.435)
Monthly Return$_{t-1}$		-0.059***		-0.055***
		(-8.262)		(-11.340)
MOM$_{t-12,t-1}$		0.404**		0.448**
		(2.451)		(2.573)
Ln(Size)$_{t-1}$		-0.109***		-0.112***
		(-2.843)		(-2.925)
Ln(B/M)$_{t-1}$		0.173**		0.170**
		(2.142)		(2.165)
Beta$_{t-1}$		0.120		0.082
		(0.943)		(0.635)
Constant	1.489***	2.658***	1.347***	2.567***
	(6.341)	(4.758)	(6.290)	(4.761)
Observations	1,140,958	1,140,958	1,140,958	1,140,958
R-squared	0.010	0.056	0.007	0.055

Table 11: Equal-weighted seasonal portfolios double sorted on prior earnings announcements: weekly rebalance

This table reports equal-weighted returns and alphas of portfolios sorted by prior seasonal returns and information release likelihood. Each week, we independently sort stocks into quintiles based on their average return in the same week in the last five years, and into two groups based on whether the majority of these past seasonal weeks are associated with information releases. Information releases are defined as quarterly earnings announcements (Compustat item RQD). We define *Information-cycle-driven seasonality* as a hedge portfolio with a long position in seasonal winners with prior information releases and a short position in seasonal losers with no prior seasonal information releases. In contrast, *Counter information-cycle seasonality* is a hedge portfolio with a long position in seasonal winners with no prior seasonal information releases and a short position in seasonal losers with prior seasonal information releases. We hold stocks for one week in each portfolio. Daily returns are adjusted following Asparouhova et al. (2010). Panel A reports raw portfolio returns, Panel B reports Fama and French (2015) five-factor alphas. t-statistics are in brackets and the portfolios hold stocks from 1976–2017.

Panel A: Raw return (%)						
	Prior Seasonal Return					
Seasonal Portfolio	1 (loser)	2	3	4	5 (winner)	5-1 diff
All stocks	0.16	0.25	0.27	0.32	0.37	0.21
	[3.00]	[5.18]	[5.75]	[6.60]	[6.59]	[10.06]
Stocks with prior seasonal info release	0.29	0.45	0.32	0.40	0.48	0.20
	[3.47]	[5.56]	[4.45]	[5.90]	[6.61]	[2.64]
Stocks with no prior seasonal info release	0.14	0.23	0.25	0.29	0.32	0.18
	[2.34]	[4.38]	[5.01]	[5.66]	[5.25]	[8.25]
Info-cycle-driven seasonality						0.34
						[6.97]
Counter info-cycle seasonality						0.03
						[0.52]
Info - Counter info						0.31
						[4.09]

Panel B: Fama-French five-factor alphas (%)						
	Prior Seasonal Return					
Seasonal Portfolio	1 (loser)	2	3	4	5 (winner)	5-1 diff
All stocks	-0.09	-0.03	-0.01	0.04	0.11	0.20
	[-4.84]	[-2.05]	[-0.75]	[3.46]	[5.95]	[9.74]
Stocks with prior seasonal info release	0.03	0.18	0.05	0.12	0.21	0.18
	[0.56]	[3.08]	[1]	[2.66]	[4.37]	[2.39]
Stocks with no prior seasonal info release	-0.10	-0.04	-0.02	0.03	0.07	0.17
	[-5.05]	[-2.91]	[-1.34]	[2.17]	[3.68]	[7.93]
Info-cycle-driven seasonality						0.31
						[6.34]
Counter info-cycle seasonality						0.03
						[0.57]
Info - Counter info						0.28
						[3.62]

Table 12: Value-weighted seasonal portfolios double sorted on prior earnings announcements: weekly rebalance

This table reports value-weighted returns and alphas of portfolios sorted by prior seasonal returns and information release likelihood. Each week, we independently sort stocks into quintiles based on their average return in the same week in the last five years, and into two groups based on whether the majority of these past seasonal weeks are associated with information releases. Information releases are defined as quarterly earnings announcements (Compustat item RQD). We define *Information-cycle-driven seasonality* as a hedge portfolio with a long position in seasonal winners with prior information releases and a short position in seasonal losers with no prior seasonal information releases. In contrast, *Counter information-cycle seasonality* is a hedge portfolio with a long position in seasonal winners with no prior seasonal information releases and a short position in seasonal losers with prior seasonal information releases. We hold stocks for one week in each portfolio. Panel A reports raw portfolio returns, Panel B reports Fama and French (2015) five-factor alphas. t-statistics are in brackets and the portfolios hold stocks from 1976–2017.

Panel A: Raw return (%)						
	Prior Seasonal Return					
Seasonal Portfolio	1 (loser)	2	3	4	5 (winner)	5-1 diff
All stocks	0.19	0.22	0.23	0.26	0.32	0.13
	[3.43]	[4.62]	[4.96]	[5.4]	[5.63]	[3.92]
Stocks with prior seasonal info release	0.33	0.34	0.37	0.32	0.33	-0.01
	[3.84]	[4.13]	[4.12]	[4.06]	[3.93]	[-0.05]
Stocks with no prior seasonal info release	0.17	0.21	0.22	0.25	0.31	0.14
	[2.81]	[4.11]	[4.47]	[4.72]	[4.98]	[3.74]
Info-cycle-driven seasonality						0.16
						[2.41]
Counter info-cycle seasonality						-0.02
						[-0.32]
Info - Counter info						0.18
						[1.93]

Panel B: Fama-French five-factor alphas (%)						
	Prior Seasonal Return					
Seasonal Portfolio	1 (loser)	2	3	4	5 (winner)	5-1 diff
All stocks	-0.06	-0.03	-0.03	0.003	0.09	0.15
	[-2.91]	[-2.56]	[-3.24]	[0.27]	[4.28]	[4.44]
Stocks with prior seasonal info release	0.08	0.08	0.09	0.06	0.09	0.01
	[1.15]	[1.24]	[1.19]	[1.05]	[1.5]	[0.14]
Stocks with no prior seasonal info release	-0.08	-0.04	-0.04	-0.01	0.08	0.16
	[-3.59]	[-2.96]	[-3.34]	[-0.46]	[3.44]	[4.36]
Info-cycle-driven seasonality						0.17
						[2.63]
Counter info-cycle seasonality						0.003
						[0.97]
Info - Counter info						0.17
						[1.82]

Table 13: Information-cycle explanation conditional on analyst coverage (equal-weighted portfolio)

This table reports equal-weighted returns and alphas of portfolios sorted by prior seasonal returns and information release likelihood according to independently sorted analyst coverage groups. Firms are divided into four groups based on the number of analysts covering a firm, namely, [0,2] analysts, [3,6] analysts, [7,11] analysts, and > 11 analysts according to the analyst coverage reported by the I/B/E/S Summary File. Each month, we independently sort stocks into quintiles based on their average return in the same calendar month in the last five years, and into two groups based on whether the majority of these past seasonal months are associated with information releases. Information releases are defined as quarterly earnings announcements (Compustat item RQD). *info-cycle* is a hedge portfolio with a long position in seasonal winners with prior information releases (*winner & info*) and a short position in seasonal losers with no prior seasonal information releases (*loser & noninfo*). In contrast, *Counter infor-cycle* is a hedge portfolio with a long position in seasonal winners with no prior seasonal information releases (*winner & noninfo*) and a short position in seasonal losers with prior seasonal information releases (*loser & info*). Panel A reports raw portfolio returns, Panel B reports Fama and French (2015) five-factor alphas. t-statistics are in brackets and the portfolios hold stocks from 1976–2017.

Panel A: Raw return (%)

Number of analysts	winner & info	loser & noninfo	winner & noninfo	loser & info	info-cycle	counter info-cycle	Diff
Q4 (> 11)	1.97	0.89	1.19	1.04	1.11	0.13	0.96
	[5.85]	[2.74]	[2.98]	[2.7]	[4.66]	[0.45]	[2.9]
Q3 ($\geq$ 7 & $\leq$ 11)	2.19	1.31	1.19	1.48	1.10	-0.004	1.17
	[5.98]	[3.89]	[2.84]	[3.53]	[4.42]	[-1.55]	[3.25]
Q2 ($\geq$ 3 & $\leq$ 6)	1.87	0.92	1.17	1.44	0.95	0.22	0.65
	[5.51]	[2.98]	[3.19]	[3.96]	[4.97]	[1.03]	[2.49]
Q1 ($\geq$ 0 & $\leq$ 2)	1.96	1.17	1.73	1.36	0.80	0.37	0.43
	[6.72]	[4.18]	[5.62]	[4.31]	[5.01]	[2.04]	[2.00]

Panel B: Fama-French five-factor alphas (%)

Number of analysts	winner & info	loser & noninfo	winner & noninfo	loser & info	info-cycle	counter info-cycle	Diff
Q4 (> 11)	0.77	-0.45	0.01	-0.34	1.21	0.32	0.88
	[4.38]	[-2.72]	[0.03]	[-1.55]	[4.94]	[1.1]	[2.56]
Q3 ($\geq$ 7 & $\leq$ 11)	0.68	-0.34	-0.23	0.14	1.03	-0.39	1.02
	[3.48]	[-1.99]	[-1.04]	[0.62]	[3.9]	[-1.21]	[2.64]
Q2 ($\geq$ 3 & $\leq$ 6)	0.61	-0.47	0.17	-0.27	1.07	0.39	0.66
	[3.37]	[-3.83]	[1.18]	[-1.63]	[5.41]	[1.77]	[2.39]
Q1 ($\geq$ 0 & $\leq$ 2)	0.79	-0.03	0.66	0.10	0.82	0.56	0.26
	[4.85]	[-0.22]	[4.24]	[0.53]	[4.90]	[3.01]	[1.16]

Table 14: Information-cycle explanation conditional on analyst coverage (value-weighted portfolio)

This table reports value-weighted returns and alphas of portfolios sorted by prior seasonal returns and information release likelihood according to independently sorted analyst coverage groups. Firms are divided into four groups based on the number of analysts covering a firm, namely, [0,2] analysts, [3,6] analysts, [7,11] analysts, and > 11 analysts according to the analyst coverage reported by the I/B/E/S Summary File. Each month, we independently sort stocks into quintiles based on their average return in the same calendar month in the last five years, and into two groups based on whether the majority of these past seasonal months are associated with information releases. Information releases are defined as quarterly earnings announcements (Compustat item RQD). *info-cycle* is a hedge portfolio with a long position in seasonal winners with prior information releases (*winner & info*) and a short position in seasonal losers with no prior seasonal information releases (*loser & noninfo*). In contrast, *Counter infor-cycle* is a hedge portfolio with a long position in seasonal winners with no prior seasonal information releases (*winner & noninfo*) and a short position in seasonal losers with prior seasonal information releases (*loser & info*). Panel A reports raw portfolio returns, Panel B reports Fama and French (2015) five-factor alphas. t-statistics are in brackets and the portfolios hold stocks from 1985–2017.

Panel A: Raw return (%)

Number of analysts	winner & info	loser & noninfo	winner & noninfo	Loser & info	info-cycle	counter info-cycle	Diff
Q4 (> 11)	1.61	0.99	0.83	1.16	0.62	-0.33	0.94
	[4.16]	[2.59]	[2.06]	[3.00]	[1.81]	[-0.98]	[2.46]
Q3 (≥ 7 & ≤ 11)	1.63	1.18	1.06	1.57	0.45	-0.51	0.96
	[4.24]	[3.2]	[2.52]	[3.90]	[1.57]	[-1.47]	[2.47]
Q2 (≥ 3 & ≤ 6)	1.46	0.83	1.05	1.06	0.63	-0.02	0.64
	[4.12]	[2.51]	[3.01]	[2.81]	[2.99]	[-0.06]	[2.16]
Q1 (≥ 0 & ≤ 2)	1.43	0.63	1.03	0.93	0.80	0.09	0.71
	[4.93]	[2.16]	[3.49]	[3.17]	[3.72]	[0.4]	[2.54]

Panel B: Fama-French five-factor alphas (%)

Number of analysts	winner & info	loser & noninfo	winner & noninfo	loser & info	info-cycle	counter info-cycle	Diff
Q4 (> 11)	0.53	-0.14	-0.21	0.02	0.66	-0.23	0.89
	[2.48]	[-0.61]	[-0.97]	[0.10]	[1.88]	[-0.68]	[2.2]
Q3 (≥ 7 & ≤ 11)	0.14	-0.08	-0.27	0.44	0.23	-0.72	0.94
	[0.64]	[-0.39]	[-1.18]	[1.78]	[0.74]	[-1.95]	[2.25]
Q2 (≥ 3 & ≤ 6)	0.25	-0.48	-0.23	-0.19	0.73	-0.05	0.78
	[1.44]	[-3.76]	[-1.64]	[-1.01]	[3.39]	[-0.19]	[2.5]
Q1 (≥ 0 & ≤ 2)	0.26	-0.57	-0.07	-0.31	0.83	0.24	0.59
	[1.59]	[-4.09]	[-0.47]	[-1.75]	[3.69]	[0.98]	[2.01]

Figure 1: The seasonality anomaly

This figure shows the average monthly returns of seasonality quintiles. At the beginning of each month, we sort stocks into seasonality quintiles based on their average returns in the same calendar month in last five years and hold the portfolios for one month. The y-axis denotes the average monthly returns of these portfolios. 1 in the x-axis denotes the seasonal loser quintile and 5 denotes the seasonal winner quintile, and so on. The sample period is from 1972 to 2017.

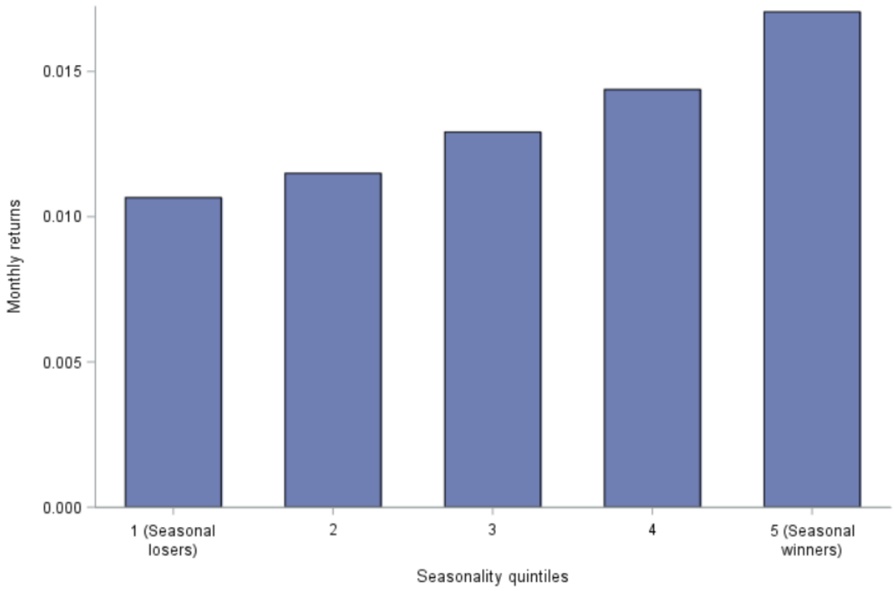
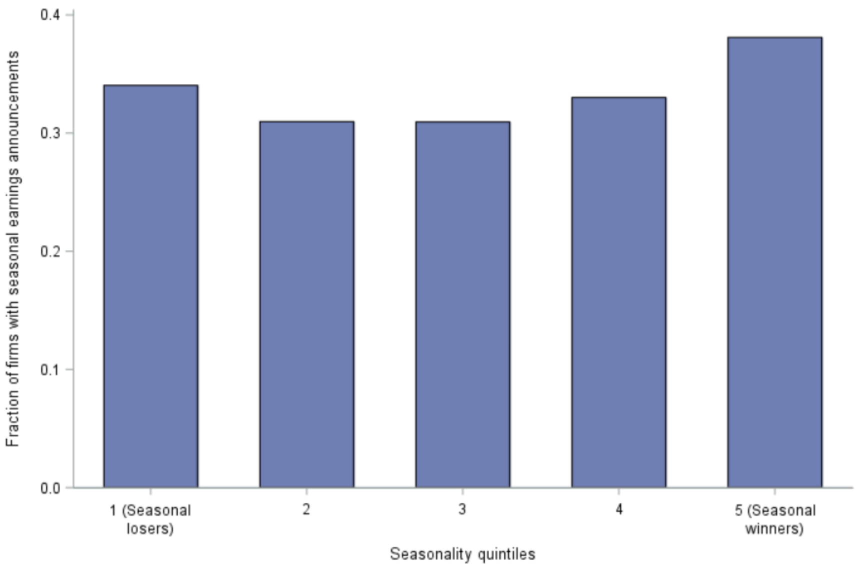


Figure 2: Seasonality quintiles and information release likelihood

This figure shows the information release likelihood in their past seasonal months. At the beginning of each month, we sort stocks into seasonality quintiles based on their average returns in the same calendar month in the last five years. In Panel A, we use the average fraction of firms with earnings announcements as a proxy for the information release likelihood. A dummy variable, "With Seasonal Earnings Announcement", equals to one if the majority of these past seasonal months are associated with an earnings announcement (from the RDQ variable in Compustat). Then the average of this dummy variable is plotted for each seasonality quintile. In Panel B, the information-release proxy is the change in option-implied volatility, which is defined as the month-end to month-end percentage change of daily option-implied volatility (e.g., if the implied volatility changes from 0.5 to 0.45, the percentage change is -10%). The average of this variable is plotted for each seasonality quintile. 1 in the x-axis denotes the seasonal loser quintile and 5 denotes the seasonal winner quintile, and so on. The sample period is from 1972 to 2017 for Panel A and from 1996 to 2017 for Panel B.

Panel A: Using earnings announcements as a proxy for information release



Panel B: Using change in option-implied volatility as a proxy for information release

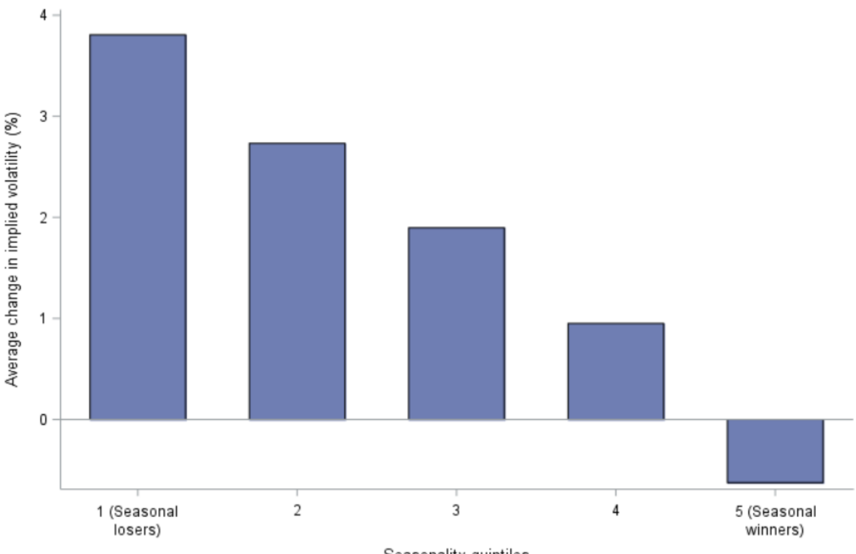
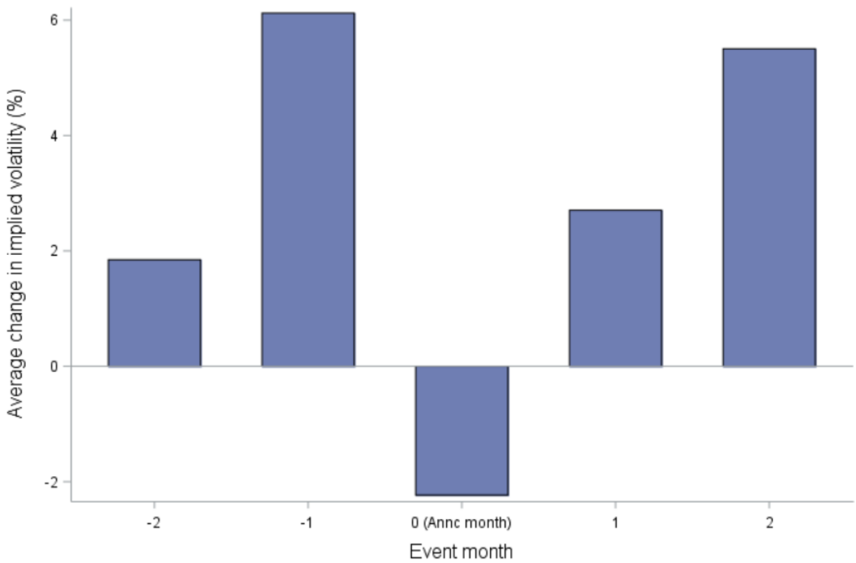


Figure 3: Change in option-implied volatility around earnings announcements

This figure shows the average percentage change in option-implied volatility of firms around earnings announcements at the monthly (Panel A) and weekly (Panel B) horizon. In Panel A, the x-axis denotes the event time where 0 represents the month with an earnings announcement. The y-axis denotes the average percentage change in implied volatility of all firms at the same event month. We define the percentage change as the period-end to period-end percentage change of daily option-implied volatility (e.g., if the implied volatility changes from 0.5 to 0.45, the percentage change is -10%). In Panel B, we calculate the same variables based on weekly horizon. The sample period is from 1996 to 2017.

Panel A: Change in implied volatility around earnings announcement month



Panel B: Change in implied volatility around earnings announcement week

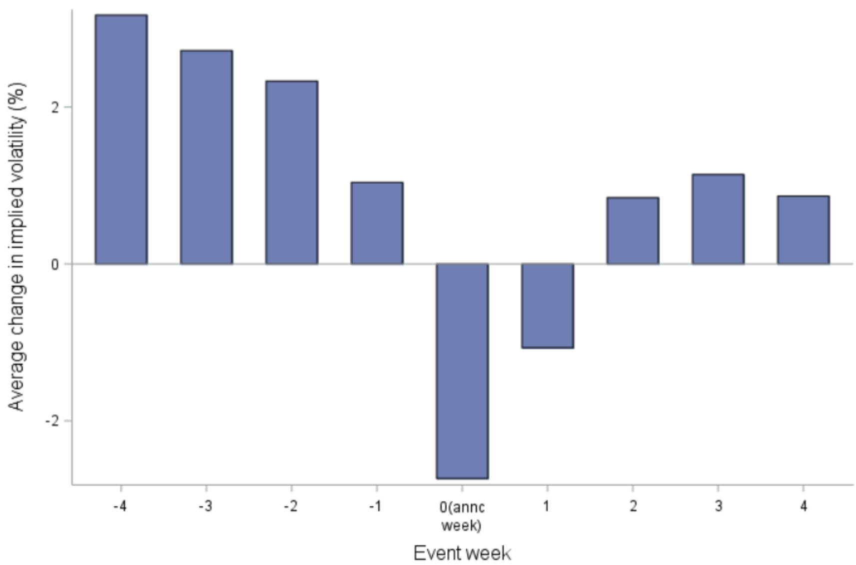
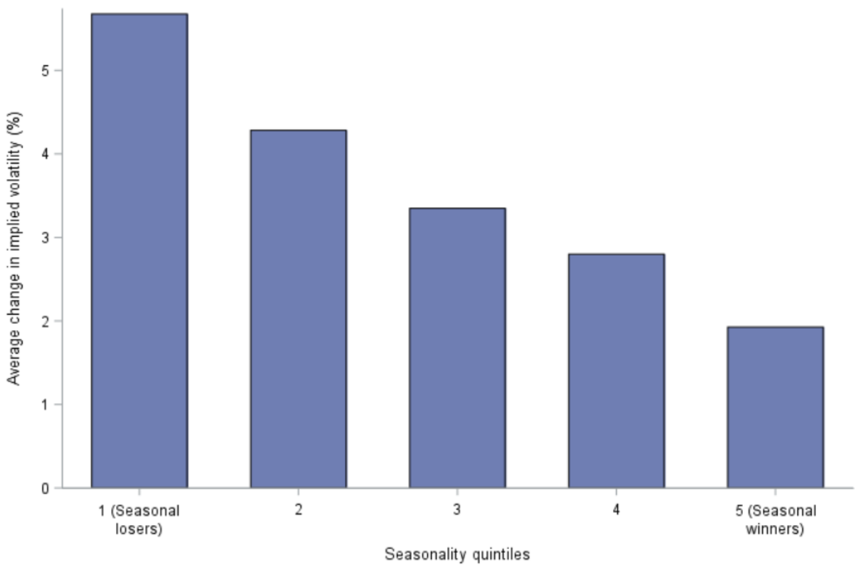


Figure 4: Seasonality quintiles and change in option-implied volatility conditional on earnings announcements

This figure shows the average change in option-implied volatility of firms in their past seasonal months. At the beginning of each month, we sort stocks into seasonality quintiles based on their average returns in the same calendar month in the last five years. We then define a dummy variable “With Seasonal Earnings Announcement” which checks whether the majority of these past seasonal months are associated with an earnings announcement (from the RDQ variable in Compustat). Firms without announcements are plotted in Panel A and firms with announcements are plotted in Panel B. The average percentage change in implied volatility of those past seasonal months for each seasonality quintile is plotted. 1 in the x-axis denotes the seasonal loser quintile and 5 denotes the seasonal winner quintile, and so on. The sample period is from 1996 to 2017.

Panel A: Seasonal returns and change in implied volatility among firms without announcements



Panel B: Seasonal returns and change in implied volatility among firms with announcements

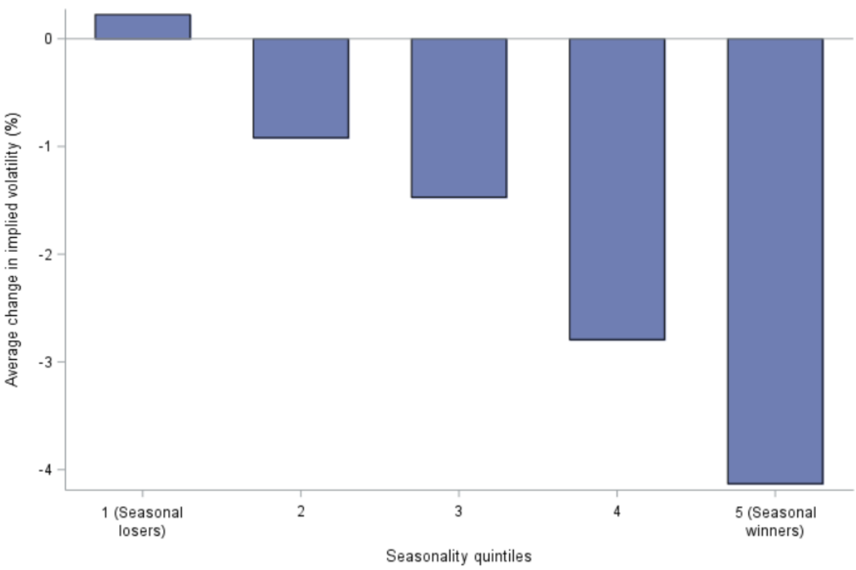
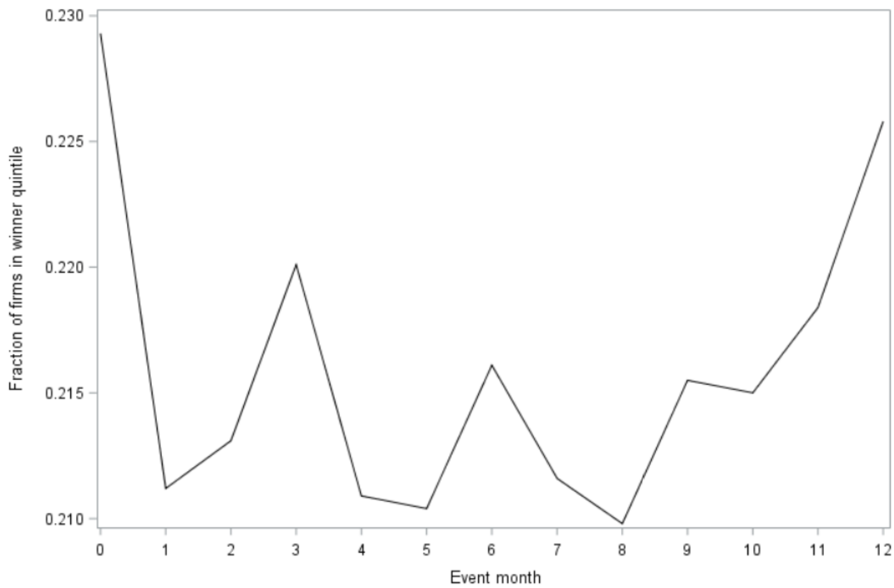


Figure 5: Fraction of month-0 defined seasonal winner firms that remain in winner quintile in future months

Panel A of this figure shows the average fraction of month-0 defined seasonal winner firms that remain in winner quintile for 13 months from month 0 to month 12. In month 0, a seasonal winner firm is defined as the firm in the top quintile based on the average same-calendar month returns in the past five years. We then define a dummy variable "In winner quintile" which checks whether a firm remains in the top quintile based on the return in current month. The average of this dummy variable among all month-0 defined seasonal winners is plotted from month 0 and month 12. In Panel B, we divide month-0 seasonal winners into groups based on whether month-0 is associated with information releases proxied by earnings announcements. The sample period is from 1972 to 2017.

Panel A: Unconditional seasonal winners



Panel B: Info-driven winners versus noninfo-driven winners

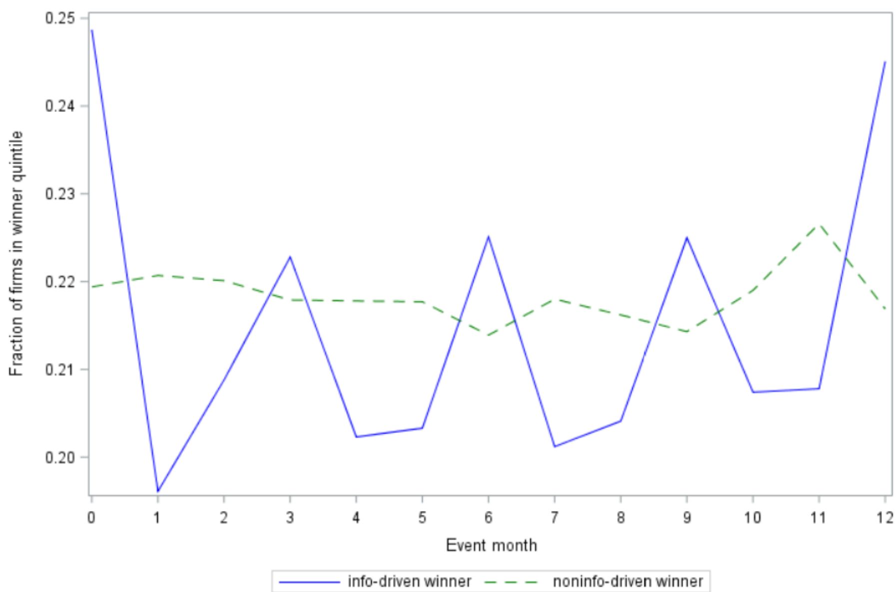
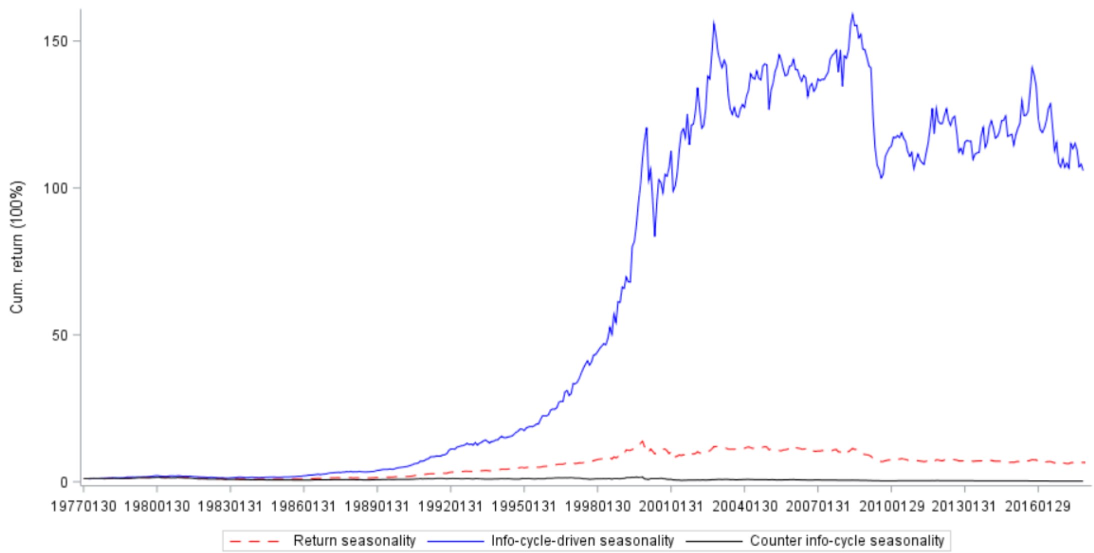


Figure 6: Cumulative returns to return seasonality strategies

This figure presents the value-weighted buy-and-hold cumulative monthly returns of three seasonality strategies for our 1977 to 2017 sample. At the beginning of each month, we sort stocks into seasonality quintiles based on their average returns in the same calendar month in the last five years and hold the portfolios for one month. The top (bottom) quintile is defined as seasonal winners (losers). The cumulative returns of the return seasonality hedge portfolio (denoted by a red-dashed line) come from a long position in seasonal winners and a short position in seasonal losers. We then define a dummy variable "With Seasonal Earnings Announcements" which checks whether the majority of the past seasonal months are associated with an earnings announcement (using the RDQ variable in Compustat). The cumulative returns from the info-cycle-driven seasonality hedge portfolio (denoted by a blue line) come from a long position in seasonal winners with seasonal earnings announcements and a short position in seasonal losers without seasonal earnings announcements. The cumulative returns from the counter info-cycle seasonality hedge portfolio (denoted by a black line) come from a long position in seasonal winners without seasonal earnings announcements and a short position in seasonal losers with seasonal earnings announcements.



Chapter 3

Managerial and analyst horizons during conference calls

By Haoyuan Li

I thank Gennaro Bernile, Weikai Li, Roger Loh, and Rong Wang for comments and suggestions.

Abstract

It is alleged that public-firm managers face short-term pressures from investors. In this paper, I examine managers' tendency to talk about the short versus the long term by analyzing the language in quarterly analyst conference calls. Using the word embedding model, I determine whether conference calls focus on the short or long term. I find that when firms fail to meet analyst expectations, both managers and analysts focus on the short term rather than the long term. However, in macro bad times, analysts question managers about the short term rather than the long term, while managers maintain the same long term-short term balance whether in good or bad macro conditions. Finally, I show that firms whose conference call participants focus more on the long term have negative initial market reactions, but stock prices recover in the subsequent months. subsequent months. The results are consistent with Wall Street exerting excessive short-term pressures on public firm managers.

Keywords: Managerial horizons; Analyst horizons; Conference call; Textual analysis

3.1. Introduction

Managers need to decide on the trade-off between delivering earnings in the short term, and investing in long-term value-maximizing projects which might potentially hurt short-term earnings. The conflict between the short-term and long-term goals has been widely documented (e.g., Bushee (1998), Bhojraj, Hribar, Picconi, and McInnis (2009), and Asker, Farre-Mensa, and Ljungqvist (2014)). This trade-off can be influenced by various factors, such as the capital market pressures (e.g., Graham, Harvey, and Rajgopal (2005) and Bhojraj et al. (2009)), career or compensation incentives (e.g., Jensen and Murphy (1990), Dechow and Sloan (1991), and Edmans, Fang, and Lewellen (2017)), and cultural backgrounds or personal experiences (e.g., Brochet, Li, and Naranjo (2020) and Bernile, Bhagwat, and Rau (2017)). In this paper, I examine whether this trade-off varies over time according to the firm and the economy's recent performance.

There is extant evidence about how the macro economy and firms' recent performance (earnings or stock prices) influence corporate behavior. For example, Campello, Graham, and Harvey (2010) indicate that firms, especially those that are constrained, tend to cancel or postpone planned investments and bypass attractive investment opportunities during financial crises. Dong, Hirshleifer, and Teoh (2020) argue that firms when experiencing overvaluation tend to engage in highly inventive innovations. The above suggests that in bad times managers might focus more on the short term while in good times, managers focus more on the long term.

In this paper, instead of relying on a firm's investment decisions to infer the horizon that managers focus on, I directly proxy for their time-horizon focus by analyzing the language they use in quarterly earnings conference calls. Prior studies typically proxy for managers' time-horizon focus using CEO age, CEO compensation, and analyst coverage. Dechow and Sloan (1991), for example, show that CEOs in their retiring ages spend less on R&D, and Edmans et al. (2017) identify short-term concerns using the amount of stock and options scheduled to vest in a given quarter. However, these are indirect proxies for time horizons and related to many other firm characteristics. My proxy is a more straightforward measure based on the long-term versus short-term keywords which participants use in conference calls. Such a measure also has sufficient time-series variation that can allow me to examine changes conditional on firm or macroeconomic performance.

Specifically, the time horizon that the conference call speaker is focusing on is measured as the ratio of the number of long-term oriented words over the number of short-term oriented words. Based on the word dictionary of short- and long-term oriented keywords from Brochet, Loumioti, and Serafeim (2015), I use the bag-of-words method to measure the frequency of short- and long-oriented words in each segment. I also enlarge the dictionary by employing a Natural Language Processing machine learning approach called word2vec. This method finds the semantic and syntactically similar words based on a seed word. Similar methods have also been used by papers such as Gao, Ren, and Zhang (2020) and Li, Mai, Shen, and Yan (2020).

I analyze conference call transcripts from the Thomson Reuters' StreetEvents (SE) database over the period January 2001 to December 2018. The sample includes earnings conference calls held within two days from quarterly earnings announcement dates (84% of earnings conference calls are held on the day of earnings announcements, and 99% are held within two days (Price, Doran, Peterson, and Bliss (2012))). This setting allows me to test managers' and analysts' horizons immediately after the earnings announcements. Conference call participants have different concerns, which may lead to different time-horizon focus in reaction to bad times. Specifically, I test the focus in three types of "bad times": 1) macro bad times based on NBER business cycles or economic policy uncertainty; 2) firm-specific bad times based on earnings surprises; and 3) down market states based on the recent cumulative market returns or firm-specific returns.

Earnings conference calls represent a unique setting to test the time-horizon focus of managers and analysts and the interactions between the two parties. Conference calls generally consist of two segments. In the first segment, top executives disclose prepared remarks, while in the second segment, managers respond more spontaneously to analysts' questions. These two-segment (presentation and Q&A) and two-party (managers and analysts) settings differentiate conference calls from other disclosure venues, such as annual reports or analyst reports, where only one type of information is disclosed or one party is involved. Also, there are a variety of participants in conference calls. On the company side, the CEO leads the conference, and the CFO, COO, IRO (Investor Relations Officer), or the sales director attend and prepare the presentations (Brown, Call, Clement, and Sharp (2019)). On the analyst side, not only the sell side, but also the buy-side analysts (hedge funds, mutual funds, and registered investment advisers (RIAs)) attend and ask questions (Call, Sharp, and Shohfi (2020b)). Current studies show that the Q&A section is more informative and contains private information from analysts (Mayew, Sharp, and Venkatachalam (2013)). I therefore analyze the presentation and Q&A sections separately.

I first examine how managers and analysts change their time-horizon focus during macro bad times. Macro bad times threaten firms' current revenues and profits in the shorter term, and the new investments, R&D, market expenditure, and employment policy in the longer term (e.g., Campello et al. (2010) and Lins et al. (2017)). During macro bad times, managers are presumably more cautious about short-term threats: promoting current sales, raising short-term funding, and cutting long-term projects. However, managers could be long-term oriented if the crisis or uncertainty does not affect the firm's fundamentals and only exerts a temporary influence on the firm's performance. Also, Loh and Stulz (2018) show that the state of the economy affects the impact of analyst reports. However, there is no evidence on whether this differential impact comes from a shift in the time horizons which analysts focus on. Using the NBER Business cycle and economic policy uncertainty as proxies of macro-bad times, I find that analysts question managers about the short term rather than the long term in bad times compared to good times. However, the macro condition on average does not influence managerial horizons such that they maintain the same long term-short term balance either in good or bad macro conditions.

Next, I investigate the time-horizon focus when the firm itself has poor recent performance. Stakeholders rely on two benchmarks to evaluate firms' quarterly performance—the quarterly earnings in the same quarter last year and the most recent analyst consensus forecast (Graham et al. (2005)). Firm managers have great incentives not to miss the earnings benchmarks, while analysts and investors view failing to meet benchmarks as a signal of unknown problems (e.g., Graham et al. (2005) and Bhojraj et al. (2009)). In this paper, I define bad firm performance using these two proxies. I show that managers and analysts respond differently to the firm missing the two types of earnings benchmarks. When firms fail to meet analysts' expectations, both managers and analysts focus on the short term rather than the long term. On the contrary, when firms perform worse than the same quarter last year, both of them focus on the long term. This result is consistent with Wall Street exerting excessive short-termist pressures on public firm managers (Tian and Wang (2011) and Bhojraj et al. (2009)).

Further, I check whether managers and analysts respond to recent market states, since the pressure for managers to talk about the short-term comes when firms fail to meet analyst expectations rather than from a seasonal decline in performance. Following Cooper, Gutierrez, and Hameed (2004), I define a given month as a "down" macro-market or firm-specific state when the lagged three-month market or firm-specific return, respectively, is negative. Different from the NBER business cycle and high policy uncertainty measures of macro bad times, market states are ex-ante defined variables such that the corporate participants know exactly which state they are in at the time of conference call. I find that managers and analysts react differently to the down market and firm-specific states. When a firm is in the down state, managers are more likely to talk about the long term, while analysts do not change their time-horizon focus. When the market is in the down state, analysts tend to question about the short term, while managers in presentations talk about the long term. These results show that during micro or macro bad times, managers try to maintain a long-term perspective while analysts are more likely to exert short-term pressure as inferred from their conference call language.

Finally, I study how stock prices react to the time-horizon focus in conference calls. I find that the market (measured by the CAR around the earnings conference call) immediately shows negative responses to the time-horizon focus. A one-standard-deviation increase in the average horizon disclosed in the conference calls reduces the CAR around the conference call by 32 basis points from the average CAR. The time horizons disclosed in the presentation section mainly drive this negative relation. In the subsequent months, the stock prices reverse. The magnitude is estimated from regressions that control for various determinants of the earnings announcement CAR, such as earnings surprises, tone in the conference calls, and firm characteristics. I also find that the negative initial market reaction is mainly driven by the managers' time-horizon focus in the presentation section rather than in the QA section. To the extent that managers' prepared comments are done in anticipation of the questions they will receive, this result is consistent with analysts exerting short-term pressures on public firm managers.

This study contributes in several ways to the literature. First, it provides new evidence about the "soft" information conveyed by voluntary disclosures. Many signals from conference calls are documented to provide the incremental information to the market, such as managers' and analysts' tones (Davis, Ge, Matsumoto, and Zhang (2015) and Druz, Petzev, Wagner, and Zeckhauser (2020)), forward-looking statements (Bozanic, Roulstone, and Buskirk (2018)), linguistic complexity (Bushee, Gow, and Taylor (2018)), analyst sequence during the Q&A (Cohen, Lou, and Malloy (2020)), using humor (Call et al. (2020a)), and managerial affective states (Mayew and Venkatachalam (2012)). My study shows that the time-horizon focus in conference calls contains incremental information. And firms whose conference call participants focus more on the long term have better stock return performance in the future, even though they experience a short-term drop in their stock prices. Second, it is related to the literature on the short-termism. Prior studies argue that short-termism distorts investment decisions and hurts innovation (e.g., Bushee (1998) and Asker et al. (2014)). Capital market pressures and managerial compensation are documented as the sources of the short-termism, such as He and Tian (2013), Zhong (2018) and Edmans et al. (2017). My paper provides new evidence to show that pressures from analysts make managers talk about the short term during macro and micro bad times. This paper also contributes to the textual analysis literature. The bag-of-words method is widely used in economics and finance studies, such as Loughran and McDonald (2011), Guiso, Sapienza, and Zingales (2015), and Baker et al. (2016). In this paper, I enlarge the word lists to suit the research purposes, which provides a more comprehensive word dictionary.

The rest of the paper is organized as follows. Section 2 discusses the data, methodology, and descriptive statistics. Section 3–5 examine how managers and analysts react to macro bad times, firm-specific bad times, and down market states in terms of their time-horizon focus. Section 6 investigates the market reactions to the time-horizon focus. Finally, Section 7 concludes.

3.2. Data and methodolodgy

3.2.1. Sample

My primary data is earnings conference calls from the Thomson Reuters' StreetEvents database over the period from January 2001 to April 2019. This dataset includes 414,516 full-text conference call transcripts in XML (Extensible Markup Language) format covering both U.S. and international firms, and different types of calls (e.g., "Earning Conference Call/Presentation", "Sales Conference Call/Presentation" and "M&A Conference Call/Presentation").

In the analysis, I only keep the U.S. firms and the earnings conference calls held on the day or one day after the earnings announcement date. To match the StreetEvent database with CRSP and Compustat, I employ a fuzzy method based on a firm's name and ticker, following Brochet et al. (2015). The matched sample includes 130,422 firm-quarter observations for 4,825 unique firms during 2002-2018. Then I obtain corporate financial statistics from Compustat, stock price information from CRSP, analyst coverage and forecasts from I/B/E/S, and stock institutional ownership from Thomson Reuters Institutional (13f) Holdings database. To avoid the influence of penny stocks, I only include the common stock list in the main exchanges and the firms with stock prices above \$5. The sample size varies in the empirical tests depending on the data availability.

3.2.2. Methods

In this paper, I employ a machine learning method, Word2Vec, to get a larger list of time-horizon words. Word2Vec is a neuro-network method, producing a similar score of two words based on their similarity in location and meaning (Mikolov, Sutskever, Chen, Corrado, and Dean (2013)). A large data sample improves the learning accuracy. So, I use the entire collection of conference calls transcripts from January 2001 to April 2019. Such methods are also used and discussed in Hanley and Hoberg (2019), Cong, Liang, and Zhang (2019), and Li et al. (2020). Following the standard procedures, I conduct the analyzes as below.

1) Text pre-processing step

Data pre-processing is necessary for later training and should be designed based on the analysis purposes. This step usually drops inessential information, remove unclear parts, and get tokenized corps. During the conference calls, besides managers and analysts, operators (or editors, callers in different conference calls) also talk a lot. Their words do not contain much valuable information but frequently repeat, which may cause biases. The first step is to drop the operators' words from the data. Second, I clean the special characteristics (e.g., "$", "@"), and remove punctuations (e.g., ",", "?") and stop words (e.g., "a", "you", "would"). As the conference calls contain many spoken words, I use a larger list of stop words. The third step is to clean numbers. Although numbers are meaningless, I retain the numbers constituting the words, such as FY05 and Q3.

2) Data training process

In the data training process, I follow the standard-setting and set a 150-dimension vector for each word such that each word is represented in 150 dimensions as a token. So, the Word2Vec produces a vocabulary where each item is attached with a vector. Then I can query this vocabulary to detect the relations between words. For instance, I can query the cosine similarity between words based on the word vector. The data training excludes lower-frequency words (less than 100 times).

3) Generating the list of horizon-oriented words

With the training model, I enlarge Brochet et al. (2015)'s long- and short-term word lists based on the cosine similarity. Brochet et al. (2015) employ personal judgments and manual validations to get a dictionary referring to the time horizon of managers' disclosure. Each of the long- and short-word list contains around ten words. In the new list, I include seed words in their lists and words that are sufficiently close to one of the seed words with a cosine similarity larger than 0.5.

I use the following strategies to deal with the words that are related to both long and short-word lists. Some words may appear in both lists. I assign them to either long- or short-word list based on their similarity to both lists. Specifically, I define the "group similarity" as the highest cosine score between a given word and each of the seed words in one list. For example, "bimonthly" has similar scores to "weekly", "monthly", "short_term" at 0.6, 0.50 and 0.40, and hence its "group similarity" to the short-word list is 0.60. Then the word is assigned to the list with a higher"group similarity". For example, "bimonthly" has group similarity scores at 0.6 and 0.5 for short- and long-term lists and therefore is assigned as a short-term word. In addition, few words have close "group similarity" scores to both groups. So, I drop them (for example, "midterm" has similar scores to two lists at 0.60 and 0.55 and is excluded). The list of words is presented in Table 1.

[Table 1 inserts here]

3.2.3. Variables and summary statistics

The variables of interest include the time-horizon focus in conference calls and the proxies for macro and micro bad times. Table 2 shows the summary statistics of the variables. To reduce the influence of outliers, all the continuous variables are winsorized at the 1st and 99th percent levels.

[Table 2 inserts here]

3.2.3.1. Measuring time-horizon focus in conference calls

With the dictionary of time-horizon keywords, I define the *time-horizon focus* as the ratio of long-term oriented words over short-term oriented words talked in conference calls. Besides managers and analysts, conference call transcripts also record the speeches of operators or callers. Their speeches may contain the horizon words without practical meanings, therefore I excluded them from the analysis. Also, conference calls are generally organized in two segments—presentation section by the firm's management and Q&A section in which managers respond to questions from analysts. Prior studies document that managers and analysts disclose different information and thus influence the market differently. In this paper, I measure the time-horizon focus separately in these two sections.

The statistics of time-horizon focus are shown in Table 2. Consistent with Brochet et al. (2015), participants on average focus more on short-term events in conference calls. In terms of different sections, managers use more short-term oriented words, while participants in Q&A section use almost the same number of long-oriented as short-oriented words. In addition, managers' and analysts' horizons have different patterns across years, as shown in Figure 1.

[Figure 1 inserts here]

Figure 1 presents the time-horizon focus of conference call participants from 2002 to 2018. Panel A plots the average time-horizon focus of the whole conference call. The horizon declines during the financial crisis, showing that conference call participants on average focus more on the short term during bad times. Then the horizon increases from 2010 to 2018. Panel B plots the average time-horizon focus grouped by different sections and participants—managers in presentation sections, questions asked by analysts in Q&A sections, and responses by managers in Q&A sections. The time-horizon focus in presentation and Q&A sections show different patterns—managers in presentation sections almost stay at the same level from 2002 to 2018, while managers and analysts in Q&A sections display an upward trend in the same period. Obviously, the time-horizon focus in Q&A sections drive the time-series trend in Panel A.

I further check which keywords drive the average trend of the time-horizon focus. "Year" and "quarter" are the most frequent words among the long- and short-term keywords, respectively. The "year-to-quarter" ratios remain the same level from 2002 to 2018—the ratios are 0.65 in 2002 and 0.62 in 2018, and the average ratio across the sample years is 0.63. However, the "year-to-quarter" ratios in Q&A sections increase from 0.97 (2002) to 1.34 (2008). Also, the results display the same trend if I use the ratio of "top 10 long-term words" over "top 10 short-term words" in both sections. This shows that the trend is mainly driven by the change of the most frequent horizon keywords in Q&A sections.

3.2.3.2. Proxy for macro bad times

In this paper, I define bad times as the recessions marked by the National Bureau of Economic Research (NBER). NBER defined a few periods as recessions based on the significant decline in economic activities spreading across the economies, which represents a situation which everyone should attend to. So, there is only one recession period, December 2007 to June 2009, from 2002 to 2018. As shown in Table 2, around 10% of the observations are in the recession period.

Besides the NBER Business cycle, another widely used measure of macro-economy state is the policy-related economic uncertainty. Baker, Bloom, and Davis (2016) define a monthly continuous measure of policy uncertainty based on three components, including the uncertainty-related words from 10 largest newspapers, the temporary federal tax code provisions, and the

Federal Reserve Bank of Philadelphia's Survey of Professional Forecasters. In the analysis, I define the macro bad times as the months with the uncertainty scores larger than the value of the highest quintile over the period from 2002 to 2018. This measure marks more crisis months than the NBER macro bad months.

3.2.3.3. Proxy for micro bad times

Graham et al. (2005) show that managers fixate two benchmarks when evaluating the firm's performance. One is the earnings in the same quarter last year and the other is analyst consensus. The two benchmarks are important for measuring firm performance. So, I define firms' bad times by checking whether the firms perform better in the current quarter than the same quarter last year, or whether they outperform analysts' forecast or not.

Consistent with prior studies, the earnings surprise is computed as actual earnings minus expected earnings, scaled by stock price. The equation is as below:

$$SUE_{j,t} = \frac{(X_{j,t} - E(X_{j,t}))}{P_{j,t}} \tag{3.1}$$

Following Livnat and Mendenhall (2006), to match the original data reported by the firm and observed by investors, I use the adjusted Earnings Per Share (EPS) by excluding "special items" from the Compustat data. Specifically, I subtract from the primary EPS the amount of special items times 65%, divided by the number of shares used to calculate primary earnings per share. P is the price per share for firm j at the end of quarter t from Compustat.

The other one is the median forecast of analysts. Considering only the most recent forecast for each analyst, I use the median of forecasts reported to I/B/E/S in 90 days prior to the earnings announcement to measure analysts' expectations.

Then I define two dummy variables "Bad SUE" and "Negative Earnings Surp" to check if a firm has a worse performance than last year or earn lower earnings than analysts consensus forecasts. As shown in Table 2, around 37% of the earnings are defined as the bad sue or negative earnings surprise.

3.2.3.4. Proxy for market and firm-specific states

Managers and analysts' time-horizon focus could be influenced by the recent stock market and firm-specific states. Current studies document that stock price is an important consideration for the managers and analysts. For example, Graham et al. (2005) show that managers have strong motivations to maintain or increase their firm's stock price. Also, Clement et al. (2011) argue that analysts observe the actions of investors and revise their earnings forecasts referring to the stock returns. Therefore, the up or down state of the market could influence managers' and analysts' time-horizon focus. In addition, managers and analysts may have different sensitivity to the market and firm-specific states—such that analysts are more sensitive about the market states while managers are more sensitive about the firm-specific states.

The recent up or down states are defined as whether the market or firm cumulative returns are positive or not. Cooper et al. (2004) measure the market states based on the returns during 1 to 3 years. Considering that the target of analysis is to investigate the market states when the conference calls are holding, I define the states based on the recent 3 months cumulative returns. The three-month interval could capture the market states from the last conference call to the current one. Specifically, I measure the firm-specific state using firms' monthly returns and the market state using the CRSP valued weighted returns. In Table 2, 31% and 42% of the observations are in market and firm-specific down states.

3.2.3.5. Control variables

In the regression, I control for a series of variables which are related to the time-horizon focus or market reactions. Prior studies show that managerial tone in conference calls contains value-relevant information and engenders stock market reactions (Mayew and Venkatachalam (2012) and Druz et al. (2020)). Following Druz et al. (2020), I use the sentiment word lists complied by Loughran and McDonald (2011) and define the tone as the ratio of the net number of positive words over the total number of sentiment words, i.e.,

$$Tone_{j,t} = \frac{positive\ words_{j,t} - negative\ words_{j,t}}{positive\ words_{j,t} + negative\ words_{j,t}} \tag{3.2}$$

The other controls are institution ownership, number of analysts, size, and book-to-market. All the above variables are calculated at the end of fiscal quarter t.

Panel B of Table 2 reports the correlation coefficients between the key variables. The coefficient of the time-horizon focuses in presentation and Q&A sections is 0.44, which shows that managers and analysts have different focuses in these two sections. There is a low correlation between the tone and time-horizon focus of conference calls with a value equals to 0.15. Also, the macro and firm-specific bad times are not very correlated.

Figure 2 plots the average tone and time-horizon focus of conference calls by calendar years. Although these two have a low correlation, they display a close time trend from 2002 to 2018, which decreases during the financial crisis and increases for a long time later.

[Figure 2 inserts here]

3.3. Time-horizon focus and the state of economy

Macro conditions are key contexts for managers and investors' decision making. Literature has documented the impact of bad conditions on managers and analysts' decisions. For example, Gulen and Ion (2016) shows that managers tend to delay the capital investment during the high policy uncertainty period, and such influence can last up to two years. Jens (2017) provides empirical evidence showing that political uncertainty could depress firms' financing activities and then the investment decisions. Also, Loh and Stulz (2018) show that the state of economy makes analysts' revisions and recommendations more valuable.

However, the influence of macro conditions on managers' and analysts' horizons is unexplored. During the recession periods, they may focus on the long-term or the short-term. In this section, I use the following framework to estimate the effects of macro conditions on managers' and analysts' time horizons.

$$Time\_horizon\ focus_{j,t} = \beta_0 + \beta_1 \times Macro\ Bad\ Condition_{j,t} + \delta \times Controls_{j,t} + \mu \tag{3.3}$$

During the conference calls, managers and analysts disclose different information (prepared vs. impromptu). In this section, I investigate the time-horizon focus in the whole conference call, presentation, and Q&A, respectively. I use two bad-condition proxies: the recession periods marked by NBER and the high uncertainty periods in the Baker et al. (2016) policy uncertainty index. In addition, I control variables that are documented to influence horizons, such as the number of analysts (Tian and Wang (2011) and He and Tian (2013)), institutional ownership (Hartzell and Starks (2003)), as well as firm characteristics. I also include year, calendar month and industry (Fama and French (1997) 10-industry classification) fixed effects to control for persistent effects across industries and months. Table 3 presents the regression results.

[Table 3 inserts here]

As shown in the Column (1) and (2), managers and analysts tend to talk more about the short term during the recession periods. But after checking the horizons in different segments separately, I find that managers and analysts have different horizon focus during the macro-bad times. Column (3) and Column (4) show that managers tend to maintain the same long term-short term balance whether in good or bad macro conditions, while analysts mainly drive the short-oriented tendency in the conference calls.

In addition, control variables predict the time-horizon focus as in the same direction documented in prior studies. Graham et al. (2005) documented that most CFOs believe that institutional investors set the stock price on the buy-side in the long run, and analysts affect the short-term prices. In Table 3, conference call participants tend to focus on the longer term if a firm has a high institutional ownership percentage, while the participants tend to focus on the shorter term if a firm has a greater analyst coverage. In addition, the size of firm is positively related to the time-horizon focus, and the book-to-market ratio has a negative relation.

3.4. Time-horizon focus and firm-specific performance

Managers care about two earnings benchmarks: the seasonally-lagged quarterly earnings and the analyst consensus forecast. Meeting or exceeding the earnings benchmarks sends a positive signal to the market and affects the stock prices. While failing to meet the benchmarks could cause a series of long and short-term problems such as creating uncertainty about a firm's future prospects and causing investors to suspect undisclosed problems in the firm (Graham et al. (2005)). Therefore, after the earnings announcement, managers have to spend considerable time explaining why they missed the benchmarks. Failing to hit the benchmarks will affect managers' horizons. In this section, I examine the relation between the failure to meet earnings benchmarks and time-horizon focus.

$$Time\_horizon\ focus_{j,t} = \beta_0 + \beta_1 \times Firm\ Bad\ Performance_{j,t} + \delta \times Controls_{j,t} + \mu \quad (3.4)$$

I define the firm's bad performance based on whether they hit the earnings benchmarks or not. $Bad\ SUE_{j,t}$ is a dummy variable which equals to one if the firm has worse earnings in quarter t comparing to the same quarter last year. $Negative\ Surp_{j,t}$ denotes the quarter when the firm fails to meet analysts' consensus forecast in recent 90 days. The controls include the macro conditions, institutional ownership, number of analysts and firm characteristics. Regression results are shown in Table 4.

[Table 4 inserts here]

Both managers and analysts react strongly to the failure of meeting earnings benchmarks. However, their time-horizon focus changes differently between the failures to two earnings benchmarks. If firms fail to meet analysts' expectations, both managers and analysts focus on the short term rather than the long term. It is consistent with studies on the extreme pressures from the analysts. While if the firms earn less earnings than the same quarter last year, both managers and analysts focus on the longer period.

3.5. Time-horizon focus and recent market state

Besides the macro and firm-specific bad times, the recent market and stock down states affect managers' and analysts' time-horizon focus. Here I define the market down state as the negative cumulative returns during three months prior to the month of conference calls and the firm down state as the negative cumulative firm-specific returns. Different from the macro recessions and firms' under-benchmark earnings, the market and stock down states represent investors' views rather than firms' real losses.

Maintaining and supporting the stock price is an important motivation for managers' earnings management and voluntary disclosure (Graham et al. (2005)). If a firm's stock price is performing bad, the firm is likely to receive more unfavorable comments from analysts and the media. Also, Clement et al. (2011) shows that analysts tend to observe the actions of investors and revise their earnings forecasts in response to recent stock returns. To test the relation between the states of market and stock and time-horizon focus, I estimate the following regression.

$$Time\_horizon\ focus_{j,t} = \beta_0 + \beta_1 \times market\ states_{j,t} + \delta \times Controls_{j,t} + \mu \tag{3.5}$$

[Table 5 inserts here]

As shown in the first three columns of Table 5, the overall time-horizon focus in conference calls is not influenced by the recent market state. But the market state has different influences on the time-horizon focus in the two segments of conference calls. With the controls, managers focus on the longer term if the market is in a down state in the recent three months, while analysts shorten their time horizon. In the firm-specific down states, managers tend to focus more on the long term, while analysts have no significant reactions.

3.6. The market reaction to time-horizon focus

In this section, I examine how the market reacts to the time-horizon focus in conference calls. Prior studies have documented that investors respond to the qualitative or "soft" information in conference calls (e.g., Davis et al. (2015), Bushee et al. (2018), and Mayew and Venkatachalam (2012)). However, it is less evident that how the time horizons would be incorporated into the prices. To test the relation between time-horizon focus and stock prices, I estimate the following regression.

$$CAR_{j,t} = \beta_0 + \beta_1 \times Time\_horizon\ focus_{j,t} + \delta \times Controls_{j,t} + \mu \tag{3.6}$$

In Equation (6), I regress CAR[0,1] and CAR[2,63] on the time-horizon focus in the whole conference call and the horizons in different sections. CAR[0,1] and CAR[0,63] are cumulative Fama and French (1993) adjusted returns in the first two days and next three months from the conference call date. The control variables include various determinants of the earnings announcement CAR, such as earnings surprise magnitude, tone of the conference calls, and firm characteristics. Year, calendar month and industry (Fama and French (1997) 10-industry classification) fixed effects are included. Regression results for immediate and medium-term stock price reactions are shown in Table 6 and Table 7.

3.6.1. Immediate stock market reactions

Managers' long time-horizon focus induces uncertainty and makes the market harder to predict future prospects. Graham et al. (2005) shows that uncertainty about earnings could induce a perceived estimation risk and lead to a higher risk premium. If so, the market should negatively react to the time-horizon focus disclosed in the conference calls.

[Table 6 inserts here]

Table 6 reports how the market reacts to the time-horizon focus on the conference call date and the next day. Column (1) and (2) shows that the market has a negative initial reactions to the time horizons. A one-standard-deviation increase in the average horizon exhibited in the conference calls will lead to 14 basis points lower than the average value. In Column (2), the relation still holds with controls. Consistent with prior studies, earnings surprises, conference call tones, and book-to-market ratio strongly predict the initial market reactions, while the firm size is negatively related to the initial returns.

However, market reacts differently to the time-horizon focus in the presentation and Q&A sections. In Column (3), a one-standard-deviation increase in the managers' horizons will reduce the CAR by 14 basis points. However, in Column (5), the horizons in Q&A section do not drive the initial market reactions. Comparing with Column (3) and Column (5), it shows that the time-horizon focus in presentation conveys a more straightforward signal to investors, and mainly drive the initial market reactions to the whole conference calls.

3.6.2. Event-time returns beyond the initial market reactions

The current results show that the market will initially react negatively to the longer horizons in the presentation section. However, it is not clear whether and how the market reacts to the time-horizon focus in weeks or months after the calls. The market could rational-price, over-price, or under-price the horizons. If the market can price horizons rationally, there will be no relation between the time-horizon focus and market returns in the post-call period. In the case of under-reaction or over-reaction, there will be a drift or a reversal after the initial reaction.

[Table 7 inserts here]

Table 7 reports the significance of the post-call reversal in 2 to 60 trading days after the conference call. The first two columns of Table 7 show that investors over-react to the time-horizon focus in the whole section of conference calls. In Column (1), a one-standard-deviation increase in the time-horizon focus will lead to 42 basis points higher than the average value. Different from investors only reacting to the horizons in presentation sections, markets positively react to horizons in presentation and Q&A sections. It shows that investors need time to digest the time-horizon focus information. Controlling for the earnings surprises, conference call tones, and firm characteristics, I find that both presentations and Q&A sections can predict the returns in subsequent months. Also, earnings surprises can predict the earnings announcement CARs positively, which is consistent with the post earnings announcement drift. The regression coefficient of Tone is not significant, showing that investors can digest the tone information rationally.

In conclusion, combining Table 6 and Table 7, investors exert pressures to the managers when they have the long-term focus, which is consistent with Asker et al. (2014).

3.7. Conclusion

Managers need to balance the short-term need of delivering earnings and the long-term objective of maximizing firm value. In this paper, I investigate how macro and micro bad times influence managers' and analysts' focus on the long term versus short term. I use a novel measure to analyze the time-horizon focus of managers and analysts by studying the transcripts of earnings conference calls. I show that the state of macro economy and firm-specific performance affects managers' and analysts' horizons differently. In general, analysts tend to focus more on the short term in macro and micro bad times, while managers tend to maintain their long-term and short-term balance or focus more on the long term in bad times compared with the good times. Also, I document that firms focusing on longer term have negative initial market reactions but have better stock performance in the subsequent months. My results are consistent with capital markets exerting excessive short-term pressures on public firm managers.

Table 1: Time-horizon seed words and dictionary

This table presents a subset of the time-horizon dictionary. Based on the seed words of time horizons taken from Brochet, Loumioti, and Serafeim (2015), I enlarge the long- and short-word lists using a machine learning (Word2Vec) method. The training sample includes all conference call transcripts from Thomson Reuters' StreetEvents database over the period from January 2001 to April 2019. The full long-term and short-term oriented lists contain 291 and 91 words respectively. For the sake of space, Table 1 presents a subset of words which account for 99% of the words in terms of frequency.

Short-term oriented words	Long-term oriented words
seed words from Brochet, Loumioti, and Serafeim (2015)	
day, days, daily, week, weeks, weekly, month, months, monthly, quarter, quarters, quarterly, short_term	long_term, long_run, year, years, annual, annually, looking_ahead, look_ahead, outlook
The dictionary (these words account for 99% cumulatively)	
quarter, fourth_quarter, yearoveryear, months, quarters, q1, day, q2, q4, q3, days, sequentially, quarterly, yeartodate, month, june, december, march, september, week, january, april, firstquarter, fourthquarter, sequential, july, thirdquarter, secondquarter, october, weeks, february, periods, november, august, term, shortterm, 12_months, prioryear, summer, minutes, yearago, nearterm, yearoveryear_basis, yesterday, quarteroverquarter, yearonyear, sequential_basis, daily, yearago_period, monthly, hours, seasonally, short_term, hour, regular, quarterend, weeks_ago, couple_quarters, consecutive_quarters, 90_days, yeartoyear, quarterly_basis, average_daily, friday, weekend, monday, quartertoquarter, weekly, 60_days, q1_q2, tuesday, evening, yearoveryear_comparison, everyday, q2_q3, 2q, 4q, ninemonth_period, thursday, couple_weeks, january_february, 3q, sixmonth_period, wednesday, 12_18_months, officially, 1q, frequent, daytoday, yesterday_afternoon, 45_days, ninemonth	year, years, guidance, future, half, fiscal, outlook, expectations, longterm, fiscal_year, trends, annual, roughly, forecast, yearend, expectation, calendar, long_term, annualized, projections, sustainable, years_ago, ultimately, looking_ahead, fiscal_2008, longerterm, fiscal_2007, fy16, fy15, longer_term, remain_committed, annually, decade, forecasts, views, sustain, trajectory, fy14, fy17, projection, guided, fiscal_2003, foreseeable_future, threeyear, look_ahead, absolutely, backdrop, guide, eventually

Table 2: Summary statistics

This table reports summary statistics of the main variables used in the study. The sample consists of 109,543 firm-quarter observations with earnings conference calls over the period of 2002-2018. Panel A provides the summary statistics. Panel B presents the correlations between time-horizon focus and other main variables. Definitions of the variables are provided in Appendix A. All continuous variables are winsorized at the 1st and 99th percentiles.

Panel A: Summary statistics for time-horizon focus and firm characteristics

Variable	N	Mean	Std dev	P10	P25	P50	P75	P90
Time-horizon focus (Total)	109543	0.90	0.46	0.42	0.58	0.81	1.11	1.49
Time-horizon focus (Presentation)	109444	0.79	0.48	0.31	0.46	0.68	0.98	1.38
Time-horizon focus (Q&A)	107320	1.20	0.77	0.48	0.68	1.00	1.48	2.13
Time-horizon focus (Response)	106792	1.27	0.92	0.46	0.68	1.03	1.56	2.33
Time-horizon focus (Analyst)	104934	1.32	1.19	0.38	0.60	1.00	1.60	2.57
Tone	109552	-0.07	0.16	-0.28	-0.18	-0.07	0.04	0.13
NBER Recession	109555	0.09	0.28	0.00	0.00	0.00	0.00	0.00
High Uncertainty	109555	0.19	0.39	0.00	0.00	0.00	0.00	1.00
Earnings Surp	109376	0.07	0.73	-0.30	-0.03	0.06	0.21	0.52
SUE	105445	0.06	2.86	-1.44	-0.30	0.13	0.50	1.44
Negative Earnings Surp	109376	0.36	0.48	0.00	0.00	0.00	1.00	1.00
Bad SUE	105445	0.37	0.48	0.00	0.00	0.00	1.00	1.00
Market Down State	109555	0.31	0.46	0.00	0.00	0.00	1.00	1.00
Firm-specific Down State	109513	0.42	0.49	0.00	0.00	0.00	1.00	1.00
High Inst. Ownership	109555	0.66	0.47	0.00	0.00	1.00	1.00	1.00
Number of Analysts	109555	7.36	6.18	1.00	3.00	6.00	10.00	16.00
Lg(Size)	109555	12.25	1.60	10.30	11.07	12.10	13.28	14.45
Book-to-Market	107084	0.48	0.37	0.12	0.24	0.41	0.65	0.94
CAR[0,1]	108862	0.35	7.60	-8.32	-3.49	0.28	4.31	9.29
CAR[2,63]	108862	-0.92	18.38	-22.37	-10.32	-0.38	9.14	19.89

Table 2 (continued)

Panel B: The correlation matrix

	(1)	(2)	(3)	(4)	5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Time-horizon focus (Total) (1)	1.00											
Time-horizon focus (Presentation) (2)	0.87	1.00										
Time-horizon focus (Q&A) (3)	0.76	0.44	1.00									
Time-horizon focus (Response) (4)	0.66	0.39	0.90	1.00								
Time-horizon focus (Analyst) (5)	0.54	0.32	0.71	0.43	1.00							
Tone (6)	0.15	0.14	0.10	0.09	0.06	1.00						
NBER Recession (7)	-0.04	-0.01	-0.06	-0.05	-0.05	-0.15	1.00					
High Uncertainty (8)	-0.05	-0.04	-0.05	-0.04	-0.04	-0.05	0.10	1.00				
Negative Earnings Surp (9)	-0.05	-0.03	-0.04	-0.04	-0.01	-0.16	0.02	0.00	1.00			
Bad SUE (10)	-0.01	-0.01	0.00	-0.01	0.02	-0.15	0.08	0.01	0.23	1.00		
Market Down State (11)	-0.03	-0.02	-0.03	-0.03	-0.03	-0.09	0.34	0.20	0.01	0.02	1.00	
Firm-specific Down State (12)	-0.04	-0.02	-0.03	-0.03	-0.02	-0.10	0.14	0.06	0.07	0.10	0.32	1.00

Table 3: Time-horizon focus in macro bad times

Table 3 estimates the relationship between the macro bad times and time-horizon focus of conference call participants from 2002 to 2018. Time-horizon focus in Column (1) and (2) refer to the ratio of long-term oriented over short-term oriented keywords in the whole conference calls. Column (3) and (4) refer to the ratio in presentation sections, while Column (5) and (6) refer to the ratio in Q&A sections. Definitions of the variables are provided in Appendix A. And all continuous variables are winsorized at the 1st and 99th percentiles. t-statistics based on standard errors clustered by firm are in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% two-tailed levels, respectively. Industry fixed effects (FE) use the Fama and French (1997) 10-industry classification.

	Time-horizon focus (Total)		Time-horizon focus (Presentation)		Time-horizon focus (Q&A)	
Variable	(1)	(2)	(3)	(4)	(5)	(6)
NBER Recession	-0.033***		-0.007		-0.074***	
	(-4.368)		(-0.898)		(-5.222)	
High Policy Uncertainty		-0.008**		-0.003		-0.018***
		(-2.396)		(-0.849)		(-2.641)
Bad SUE	0.019***	0.019***	0.018***	0.018***	0.035***	0.035***
	(5.644)	(5.618)	(4.971)	(4.967)	(6.333)	(6.299)
Negative Surp	-0.026***	-0.026***	-0.021***	-0.021***	-0.038***	-0.038***
	(-6.655)	(-6.676)	(-5.098)	(-5.103)	(-6.214)	(-6.241)
High Inst. Ownership	0.042***	0.042***	0.030***	0.030***	0.053***	0.053***
	(4.730)	(4.729)	(3.305)	(3.305)	(4.263)	(4.261)
Num of Analysts	-0.006***	-0.006***	-0.007***	-0.007***	-0.009***	-0.009***
	(-7.464)	(-7.461)	(-9.358)	(-9.356)	(-7.558)	(-7.553)
Size	0.069***	0.070***	0.055***	0.055***	0.096***	0.096***
	(17.120)	(17.138)	(13.421)	(13.428)	(17.192)	(17.215)
Book-to-market	-0.064***	-0.063***	-0.056***	-0.056***	-0.053***	-0.052***
	(-5.749)	(-5.723)	(-4.855)	(-4.862)	(-3.558)	(-3.499)
Constant	0.375***	0.374***	0.509***	0.508***	0.220***	0.216***
	(6.623)	(6.602)	(8.747)	(8.751)	(2.787)	(2.740)
Observations	107,075	107,075	106,978	106,978	104,890	104,890
R-squared	0.196	0.196	0.166	0.166	0.126	0.125
Ind, Year, Month FE	Yes	Yes	Yes	Yes	Yes	Yes

Table 4: Time-horizon focus in firm-specific bad times

Table 4 estimates the relationship between the firm-specific bad times and time-horizon focus of conference call participants from 2002 to 2018. Time-horizon focus in Column (1) and (2) refers to the ratio of long-term oriented over short-term oriented keywords in the whole conference calls. Column (3) and (4) refer to the ratio in presentation sections, while Column (5) and (6) refer to the ratio in Q&A sections. Definitions of the variables are provided in Appendix A. And all continuous variables are winsorized at the 1st and 99th percentiles. t-statistics based on standard errors clustered by firm are in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% two-tailed levels, respectively. Industry fixed effects (FE) use the Fama and French (1997) 10-industry classification.

	Time-horizon focus (Total)			Time-horizon focus (Presentation)			Time-horizon focus (Q&A)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Negative Surp	-0.021***		-0.026***	-0.017***		-0.021***	-0.030***		-0.038***
	(-5.619)		(-6.654)	(-4.190)		(-5.098)	(-4.986)		(-6.213)
Bad SUE		0.013***	0.019***		0.013***	0.018***		0.027***	0.035***
		(3.959)	(5.648)		(3.725)	(4.973)		(4.873)	(6.338)
High Inst. Ownership	0.042***	0.043***	0.042***	0.030***	0.031***	0.030***	0.053***	0.054***	0.053***
	(4.745)	(4.794)	(4.731)	(3.319)	(3.357)	(3.305)	(4.285)	(4.334)	(4.265)
NBER Recession	-0.030***	-0.032***	-0.031***	-0.005	-0.007	-0.006	-0.067***	-0.071***	-0.069***
	(-3.920)	(-4.184)	(-4.039)	(-0.657)	(-0.875)	(-0.764)	(-4.702)	(-4.926)	(-4.820)
High Policy Uncertainty	-0.006*	-0.006*	-0.006*	-0.002	-0.003	-0.003	-0.013*	-0.013*	-0.013*
	(-1.698)	(-1.740)	(-1.730)	(-0.688)	(-0.725)	(-0.716)	(-1.837)	(-1.881)	(-1.869)
Num of Analysts	-0.006***	-0.006***	-0.006***	-0.007***	-0.007***	-0.007***	-0.008***	-0.009***	-0.009***
	(-7.317)	(-7.381)	(-7.458)	(-9.227)	(-9.290)	(-9.355)	(-7.353)	(-7.465)	(-7.550)
Size	0.069***	0.070***	0.069***	0.055***	0.056***	0.055***	0.095***	0.097***	0.096***
	(17.013)	(17.227)	(17.117)	(13.320)	(13.517)	(13.420)	(17.033)	(17.294)	(17.187)
Book-to-market	-0.062***	-0.063***	-0.064***	-0.055***	-0.056***	-0.056***	-0.050***	-0.053***	-0.054***
	(-5.592)	(-5.690)	(-5.758)	(-4.713)	(-4.810)	(-4.859)	(-3.350)	(-3.504)	(-3.575)
Constant	0.388***	0.360***	0.376***	0.520***	0.497***	0.509***	0.244***	0.198**	0.221***
	(6.865)	(6.346)	(6.629)	(8.970)	(8.537)	(8.750)	(3.090)	(2.498)	(2.794)
Observations	107,075	107,075	107,075	106,978	106,978	106,978	104,890	104,890	104,890
R-squared	0.196	0.195	0.196	0.165	0.165	0.166	0.125	0.125	0.126
Ind, Year, Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 5: Time-horizon focus in market and firm-specific down states

Table 5 estimates the relationship between the down states and time-horizon focus of conference call participants from 2002 to 2018. Time-horizon focus in Column (1) and (2) refers to the ratio of long-term oriented over short-term oriented keywords in the whole conference calls. Column (3) and (4) refer to the ratio in presentation sections, while Column (5) and (6) refer to the ratio in Q&A sections. Definitions of the variables are provided in Appendix A. And all continuous variables are winsorized at the 1st and 99th percentiles. t-statistics based on standard errors clustered by firm are in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% two-tailed levels, respectively. Industry fixed effects (FE) use the Fama and French (1997) 10-industry classification.

	Time-horizon focus (Total)			Time-horizon focus (Presentation)			Time-horizon focus (Q&A)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Market Down state	0.004		0.002	0.010***		0.007**	-0.011**		-0.011**
	(1.401)		(0.717)	(3.717)		(2.510)	(-2.195)		(-2.111)
Firm Down state		0.005*	0.005*		0.010***	0.008***		-0.002	0.001
		(1.911)	(1.661)		(3.454)	(2.729)		(-0.366)	(0.161)
Bad SUE	0.019***	0.018***	0.018***	0.017***	0.017***	0.017***	0.035***	0.035***	0.035***
	(5.605)	(5.518)	(5.529)	(4.955)	(4.756)	(4.789)	(6.296)	(6.350)	(6.323)
Negative Surp	-0.026***	-0.026***	-0.026***	-0.021***	-0.021***	-0.021***	-0.038***	-0.038***	-0.038***
	(-6.681)	(-6.737)	(-6.735)	(-5.110)	(-5.204)	(-5.193)	(-6.242)	(-6.230)	(-6.244)
High Inst. Ownership	0.042***	0.042***	0.042***	0.030***	0.030***	0.030***	0.053***	0.053***	0.053***
	(4.727)	(4.730)	(4.730)	(3.305)	(3.310)	(3.309)	(4.258)	(4.259)	(4.259)
Num of Analysts	-0.006***	-0.006***	-0.006***	-0.007***	-0.007***	-0.007***	-0.009***	-0.009***	-0.009***
	(-7.473)	(-7.493)	(-7.493)	(-9.374)	(-9.411)	(-9.410)	(-7.552)	(-7.540)	(-7.541)
Size	0.070***	0.070***	0.070***	0.055***	0.056***	0.056***	0.096***	0.096***	0.096***
	(17.142)	(17.121)	(17.122)	(13.440)	(13.450)	(13.448)	(17.207)	(17.152)	(17.154)
Book-to-market	-0.063***	-0.063***	-0.063***	-0.056***	-0.056***	-0.055***	-0.052***	-0.052***	-0.052***
	(-5.680)	(-5.652)	(-5.642)	(-4.809)	(-4.781)	(-4.759)	(-3.496)	(-3.464)	(-3.486)
Constant	0.371***	0.368***	0.367***	0.502***	0.499***	0.496***	0.223***	0.217***	0.222***
	(6.531)	(6.461)	(6.440)	(8.619)	(8.531)	(8.475)	(2.810)	(2.723)	(2.788)
Observations	107,075	107,075	107,075	106,978	106,978	106,978	104,890	104,890	104,890
R-squared	0.196	0.196	0.196	0.166	0.166	0.166	0.125	0.125	0.125
R-squared	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 6: Initial market reactions to time-horizon focus

Table 6 estimates the effect of time-horizon focus on earnings conference call two-day CARs in percent from 2002 to 2018. Time-horizon focus in Column (1) and (2) refers to the ratio of long-term oriented over short-term oriented keywords in the whole conference calls. Column (3) and (4) refer to the ratio in presentation sections, while Column (5) and (6) refer to the ratio in Q&A sections. Definitions of the variables are provided in Appendix A. And all continuous variables are winsorized at the 1st and 99th percentiles. t-statistics based on standard errors clustered by firm are in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% two-tailed levels, respectively. Industry fixed effects (FE) use the Fama and French (1997) 10-industry classification.

				CAR[0,1]			
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Time-horizon focus (Total)	-0.141**	-0.316***					
	(-2.471)	(-5.490)					
Time-horizon focus (Presentation)			-0.141***	-0.303***			-0.281***
			(-2.606)	(-5.650)			(-5.004)
Time-horizon focus (Q&A)					-0.014	-0.098***	-0.035
					(-0.433)	(-3.143)	(-1.076)
Earnings Surp		2.378***		2.376***		2.378***	2.377***
		(44.574)		(44.554)		(44.552)	(44.555)
SUE		0.043***		0.043***		0.043***	0.043***
		(3.778)		(3.793)		(3.804)	(3.783)
Tone		7.482***		7.478***		7.423***	7.485***
		(42.198)		(42.239)		(42.048)	(42.183)
High Inst. Ownership		-0.002		-0.006		-0.009	-0.005
		(-0.039)		(-0.112)		(-0.173)	(-0.090)
Num of Analysts		0.017***		0.017***		0.018***	0.017***
		(3.399)		(3.321)		(3.602)	(3.293)
Size		-0.218***		-0.223***		-0.230***	-0.221***
		(-10.958)		(-11.338)		(-11.656)	(-11.156)
Book-to-market		0.375***		0.378***		0.388***	0.377***
		(4.724)		(4.759)		(4.896)	(4.754)
Constant	0.489	3.588***	0.479	3.622***	0.335	3.485***	3.619***
	(1.393)	(8.480)	(1.374)	(8.549)	(0.961)	(8.253)	(8.546)
Observations	102,734	102,734	102,734	102,734	102,734	102,734	102,734
R-squared	0.001	0.079	0.001	0.079	0.001	0.079	0.079
Ind, Year, Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 7: Event-time returns beyond the initial market response

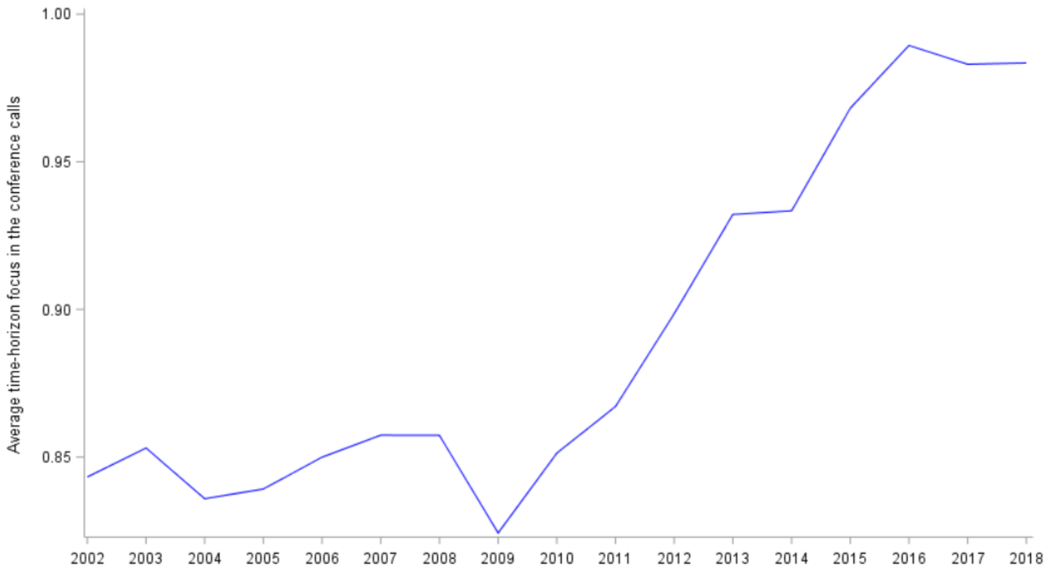
Table 7 estimates the effect of time-horizon focus on earnings conference call CARs in percent from 2 days after the conference call date through the 63rd day after that date, from 2002 to 2018. Time-horizon focus in Column (1) and (2) refers to the ratio of long-term oriented over short-term oriented keywords in the whole conference calls. Column (3) and (4) refer to the ratio in presentation sections, while Column (5) and (6) refer to the ratio in Q&A sections. Definitions of the variables are provided in Appendix A. And all continuous variables are winsorized at the 1st and 99th percentiles. t-statistics based on standard errors clustered by firm are in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% two-tailed levels, respectively. Industry fixed effects (FE) use the Fama and French (1997) 10-industry classification.

				CAR[2,63]			
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Time-horizon focus (Total)	0.445***	0.880***					
	(3.513)	(6.781)					
Time-horizon focus (Presentation)			0.246**	0.574***			0.373***
			(2.037)	(4.669)			(2.835)
Time-horizon focus (Q&A)					0.243***	0.405***	0.321***
					(3.174)	(5.216)	(3.863)
Earnings Surp		0.412***		0.416***		0.411***	0.412***
		(3.556)		(3.588)		(3.541)	(3.550)
SUE		-0.311***		-0.312***		-0.311***	-0.311***
		(-10.071)		(-10.103)		(-10.082)	(-10.070)
Tone		0.169		0.264		0.279	0.197
		(0.370)		(0.579)		(0.614)	(0.432)
High Inst. Ownership		-0.225*		-0.207		-0.212*	-0.218*
		(-1.752)		(-1.614)		(-1.651)	(-1.695)
Num of Analysts		0.115***		0.114***		0.113***	0.115***
		(8.917)		(8.826)		(8.792)	(8.929)
Size		-0.403***		-0.373***		-0.381***	-0.393***
		(-7.558)		(-7.037)		(-7.198)	(-7.386)
Book-to-market		2.681***		2.661***		2.650***	2.664***
		(12.213)		(12.135)		(12.085)	(12.145)
Constant	-6.881***	-3.879***	-6.619***	-3.829***	-6.677***	-3.626***	-3.804***
	(-5.803)	(-2.783)	(-5.593)	(-2.750)	(-5.642)	(-2.600)	(-2.730)
Observations	102,734	102,734	102,734	102,734	102,734	102,734	102,734
R-squared	0.033	0.040	0.033	0.039	0.033	0.039	0.040
Ind, Year, Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Figure 1: Average time-horizon focus by calendar years

This figure shows the average time-horizon focus by calendar years from 2002 to 2018. In Panel A, the time-horizon focus refers to the ratio of long-term oriented over short-term oriented keywords in the whole conference calls. In Panel B, the time-horizon focus as shown in the red solid line, blue medium dash line and green dash-dot line refers to the ratio in presentation sections, questions asked by analysts in Q&A sections, and responses answered by managers in Q&A sections.

Panel A: Time-horizon focus in the whole conference calls



Panel B: Time-horizon focus by different sections and participants in the conference calls

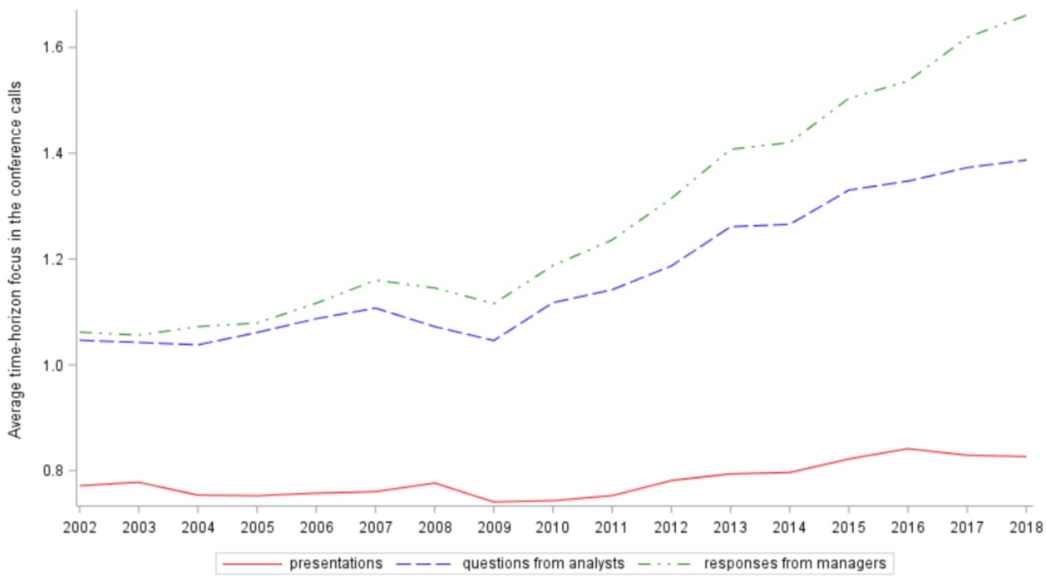


Figure 2: Average time-horizon focus and tone in conference calls by calendar years

This figure shows the average time-horizon focus and tone in conference calls by calendar years from 2002 to 2018. The time-horizon focus as shown in the blue solid line refers to the ratio of long-term oriented over short-term oriented keywords in the whole conference calls. And the tone as shown in the green dash-dot line refers to the ratio $(n-p)/(n+p)$, where n and p are the numbers of negative and positive words used in the conference calls, respectively.

Appendix A: Variable definition

This appendix describes the variables used in the main analysis. All continuous variables are winsorized at the 1st and 99th percentiles.

Variable	Definition
Time-horizon focus	
Total	Ratio of long-term oriented over short-term oriented keywords disclosed in the whole conference calls. See Table 1 for the dictionary of keywords
Presentation	Ratio of long-term oriented over short-term oriented keywords disclosed in the presentation section of conference calls.
Q&A	Ratio of long-term oriented over short-term oriented keywords disclosed in the Q&A section of conference calls.
Analyst	Ratio of long-term oriented over short-term oriented keywords used by the analysts in the Q&A section of conference calls.
Response	Ratio of long-term oriented over short-term oriented keywords responded by the analysts in the Q&A section of conference calls.
Macro bad times	
NBER Recession	Months marked by the National Bureau of Economic Research (NBER), which refers to the months between December 2007 to June 2009 in the sample period.
High Policy Uncertainty	Months with the policy-related uncertainty scores (Baker, Bloom, and Davis (2016)) larger than the value of highest quintile over the period 2002 to 2018. The uncertainty index is constructed from three types of underlying components: news coverage, federal tax code provisions, dispersion in economic forecasts of the CPI and government spending.
Firm-specific bad times	
Bad SUE	Dummy variable which equals to one if the firm earns less in quarter t than the same quarter last year.
Negative Earnings Surp	Dummy variable which equals to one if the firm fails to meet analysts' consensus forecast in quarter t. Analysts' consensus forecast is based on the estimates issued within 90 days before the report date.
Down states	
Market Down State	Dummy variable which equals to one if a firm's cumulative return is negative during the prior three months.

Appendix A (continued)

Variable	Definition
Firm Down State	Dummy variable which equals to one if the market cumulative return is negative during the prior three months.
Other variables	
Book-to-market	The ratio of book value of equity divided by market capitalization as at the end of month t
CAR[0,1]	The two-day, [0,1], cumulative Fama and French (1993) adjusted stock return in percent on and after the date of earnings conference call.
CAR[2,21]	The 20 trading days, [2,21], cumulative Fama and French (1993) adjusted stock return in percent from 2 days after the conference call date through the 21st day after that date.
High Inst. Ownership	Dummy variable which equals to one if a firm's institutional ownership is larger than the average value of all firms over period from 2002 to 2018. Institutional ownership is defined as the percentage of shares held by institutional investors at the end of the most recent reporting period.
Number of Analysts	The number of analysts who cover a firm, based on the the number of valid estimates for the consensus forecast within 90 days before the quarterly earnings announcement.
Size	Stock's market capitalization as at the end of quarter t

Appendix B: Document examples for time-horizon focus

I provide the representative examples for the time-horizon focus.

An example of short time-horizon focus in presentation section

Good morning, everyone, and thanks for joining us. This morning, we'll recap our fourth quarter 2017 results, review the asset management area and discuss our outlook for 2018. Then we'll open up the call for questions. With that, let's turn to an overview of our fourth quarter results. Comparable earnings per share from continuing operations were $1.37, up 28% from $1.07 in the prior year driven primarily by improved results in used vehicle sales and commercial rental and continued strong ChoiceLease performance. Comparable results were just above the midpoint of our forecast range of $1.31 to $1.41 as performance in our Dedicated—outperformance in our Dedicated and benefits from a favorable tax rate, unrelated to tax reform, were offset by lower-than-expected FMS results. (Q4 2017 Ryder System Inc Earnings Call)

An example of long time-horizon focus in Q&A section

Question: Okay, thanks for that. Could you talk a little bit about the rental conversions? It looks like rentals in the used units continues to decline and new unit sales are doing better than used unit sales. What are the trends as far as converting from used ones to new ones? Are they typically buying the same used unit that they are currently renting? Or are you seeing some buys, some of these units you're referring to that are being sold by other residents in the community? (Q4 2015 Equity LifeStyle Properties Inc Earnings Call)

Response: Well, let me start with what unit they're buying. When we reference conversions, when it's 10% for the year, those are current renters either buying the home that they're in or another home across the portfolio that's owned by ELS. [...] Those home sale prices that we touch in resales and used home sales are up 8% to 10% annually alongside of those homeowner sales in number up 8%. So we see some favorable trends in transactions and pricing, and I do think that has a favorable effect on our rental conversion program. (Q4 2015 Equity LifeStyle Properties Inc Earnings Call)

Bibliography

Ai, Hengjie, and Ravi Bansal, 2018, Risk preferences and the macroeconomic announcement premium, *Econometrica* 86, 1383–1430.

Ai, Hengjie, Leyla Jianyu Han, Xuhui Pan, and Lai Xu, 2020, The cross section of monetary policy announcement premium, Working paper, University of Minnesota.

An, Byeong-Je, Andrew Ang, Turan G. Bali, and Nusret Cakici, 2014, The joint cross section of stocks and options, *Journal of Finance* 69, 2279–2337.

Asker, John, Joan Farre-Mensa, and Alexander Ljungqvist, 2014, Corporate investment and stock market listing: A puzzle?, *Review of Financial Studies* 28, 342–390.

Asparouhova, Elena, Hendrik Bessembinder, and Ivalina Kalcheva, 2010, Liquidity biases in asset pricing tests, *Journal of Financial Economics* 96, 215–237.

Asquith, Paul, Michael B. Mikhail, and Andrea S. Au, 2005, Information content of equity analyst reports, *Journal of Financial Economics* 75, 245–282.

Bagnoli, Mark, Susan G. Watts, and Yong Zhang, 2008, Reg-FD and the competitiveness of all-star analysts, *Journal of Accounting and Public Policy* 27, 295–316.

Baker, Scott R., Nicholas Bloom, and Steven J. Davis, 2016, Measuring economic policy uncertainty, *Quarterly Journal of Economics* 131, 1593–1636.

Barber, Brad, Reuven Lehavy, Maureen McNichols, and Brett Trueman, 2001, Can investors profit from the prophets? Security analyst recommendations and stock returns, *Journal of Finance* 56, 531–563.

Barber, Brad M., Emmanuel T. De George, Reuven Lehavy, and Brett Trueman, 2013, The earnings announcement premium around the globe, *Journal of Financial Economics* 108, 118–138.

Bernile, Gennaro, Vineet Bhagwat, and P. Raghavendra Rau, 2017, What doesn't kill you will only make you more risk-loving: Early-life disasters and CEO behavior, *Journal of Finance* 72, 167–206.

Bernile, Gennaro, Jianfeng Hu, and Yuehua Tang, 2016, Can information be locked up? Informed trading ahead of macro-news announcements, *Journal of Financial Economics* 121, 496–520.

Bessembinder, Hendrik, 2018, Do stocks outperform treasury bills?, *Journal of Financial Economics* 129, 440–457.

Bhojraj, Sanjeev, Paul Hribar, Marc Picconi, and John McInnis, 2009, Making sense of cents: An examination of firms that marginally miss or beat analyst forecasts, *Journal of Finance* 64, 2361–2388.

Birru, Justin, Sinan Gokkaya, and Xi Liu, 2019, Capital market anomalies and quantitative research, Fisher College of Business working paper, Ohio State University.

Blei, David M., Andrew Y. Ng, and Michael I. Jordan, 2003, Latent dirichlet allocation, *Journal of Machine Learning Research* 3, 993–1022.

Bogousslavsky, Vincent, 2016, Infrequent rebalancing, return autocorrelation, and seasonality, *Journal of Finance* 71, 2967–3006.

Bondt, Werner F. M. De, and Richard H. Thaler, 1990, Do security analysts overreact?, *American Economic Review* 80, 52–57.

Bozanic, Zahn, Darren T. Roulstone, and Andrew Van Buskirk, 2018, Management earnings forecasts and other forward-looking statements, *Journal of Accounting and Economics* 65, 1–20.

Brochet, Francois, Heather Li, and Patricia L. Naranjo, 2020, Is myopia contagious? The effect of investor culture on corporate disclosure time orientation, Working paper, Boston University.

Brochet, Francois, Maria Loumioti, and George Serafeim, 2015, Speaking of the short-term: disclosure horizon and managerial myopia., *Review of Accounting Studies* 20, 1122–1163.

Brown, Lawrence D., Andrew C. Call, Michael B. Clement, and Nathan Y. Sharp, 2019, Managing the narrative: Investor relations officers and corporate disclosure, *Journal of Accounting and Economics* 67, 58–79.

Bushee, Brian J., 1998, The influence of institutional investors on myopic R&D investment behavior, *Accounting Review* 73, 305–333.

Bushee, Brian J., Ian D. Gow, and Daniel J. Taylor, 2018, Linguistic complexity in firm disclosures: Obfuscation or information?, *Journal of Accounting Research* 56, 85–121.

Call, Andrew C., Rachel Flam, Joshua A. Lee, and Nathan Y. Sharp, 2020a, Analysts' and managers' use of humor on public earnings conference calls, Working paper, Arizona State University.

Call, Andrew C., Nathan Y. Sharp, and Thomas Shohfi, 2020b, Which buy-side institutions participate in public earnings conference calls? Implications for capital markets and sell-side coverage, Working paper, Arizona State University.

Campbell, John Y., Jens Hilscher, and Jan Szilagyi, 2008, In search of distress risk, *Journal of Finance* 63, 2899–2939.

Campello, Murillo, John R. Graham, and Campbell R. Harvey, 2010, The real effects of financial constraints: Evidence from a financial crisis, *Journal of Financial Economics* 97, 470–487.

Chang, Tom Y., Samuel M. Hartzmark, David H. Solomon, and Eugene F. Soltes, 2017, Being surprised by the unsurprising: earnings seasonality and stock returns, *Review of Financial Studies* 30, 281–323.

Chen, Shuping, and Dawn A. Matsumoto, 2006, Favorable versus unfavorable recommendations: The impact on analyst access to management-provided information, *Journal of Accounting Research* 44, 657–689.

Chowdhry, Bhagwan, and Vikram Nanda, 1991, Multimarket trading and market liquidity, *Review of Financial Studies* 4, 483–511.

Cieslak, Anna, Adair Morse, and Annette Vissing-Jorgensen, 2019, Stock returns over the FOMC cycle, *Journal of Finance* 74, 2201–2248.

Clement, Michael B., Jeffrey Hales, and Yanfeng Xue, 2011, Understanding analysts' use of stock returns and other analysts' revisions when forecasting earnings, *Journal of Accounting and Economics* 51, 279–299.

Cohen, Lauren, Dong Lou, and Christopher J. Malloy, 2020, Casting conference calls, *Management Science,* forthcoming.

Cong, Lin, Tengyuan Liang, and Xiao Zhang, 2019, Textual factors: A scalable, interpretable, and data-driven approach to analyzing unstructured information, Working paper, Cornell University.

Cooper, Michael J., Huseyin Gulen, and Michael J. Schill, 2008, Asset growth and the cross-section of stock returns, *Journal of Finance* 63, 1609–1651.

Cooper, Michael J., Jr. Gutierrez, Roberto C., and Allaudeen Hameed, 2004, Market states and momentum, *Journal of Finance* 59, 1345–1365.

Cowen, Amanda, Boris Groysberg, and Paul Healy, 2006, Which types of analyst firms are more optimistic?, *Journal of Accounting and Economics* 41, 119–146.

Daniel, Kent, and Tobias J. Moskowitz, 2016, Momentum crashes, *Journal of Financial Economics* 122, 221–247.

Daniel, Kent, and Sheridan Titman, 2006, Market reactions to tangible and intangible information, *Journal of Finance* 61, 1605–1643.

Davis, Angela K., Weili Ge, Dawn Matsumoto, and Jenny Li Zhang, 2015, The effect of manager-specific optimism on the tone of earnings conference calls, *Review of Accounting Studies* 20, 639–673.

Dechow, Patricia M., and Richard G. Sloan, 1991, Executive incentives and the horizon problem: An empirical investigation, *Journal of Accounting and Economics* 14, 51–89.

Dong, Ming, David A. Hirshleifer, and Siew Hong Teoh, 2020, Misvaluation and corporate inventiveness, Working paper, University of California, Irvine.

Druz, Marina, Ivan Petzev, Alexander F. Wagner, and Richard J. Zeckhauser, 2020, When managers change their tone, analysts and investors change their tune, *Financial Analysts Journal* 76, 47–69.

Edmans, Alex, Vivian W. Fang, and Katharina A. Lewellen, 2017, Equity vesting and investment, *Review of Financial Studies* 30, 2229–2271.

Engelberg, Joseph, R. David McLean, and Jeffrey Pontiff, 2019, Analysts and anomalies, *Journal of Accounting and Economics,* forthcoming.

Fama, Eugene F., and Kenneth French, 2006, Profitability, investment and average returns, *Journal of Financial Economics* 82, 491–518.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.

Fama, Eugene F., and Kenneth R. French, 1997, Industry costs of equity, *Journal of Financial Economics* 43, 153–193.

Fama, Eugene F., and Kenneth R. French, 2015, A five-factor asset pricing model, *Journal of Financial Economics* 116, 1–22.

Fama, Eugene F., and James D. MacBeth, 1973, Risk, return, and equilibrium: empirical tests, *Journal of Political Economy* 81, 607–636.

Frazzini, Andrea, and Owen A. Lamont, 2007, The earnings announcement premium and trading volume, NBER working paper, Yale School of Management.

Gao, Zhenyu, Haohan Ren, and Bohui Zhang, 2020, Googling investor sentiment around the world, *Journal of Financial and Quantitative Analysis* 55, 549–580.

Graham, John R., Campbell R. Harvey, and Shiva Rajgopal, 2005, The economic implications of corporate financial reporting, *Journal of Accounting and Economics* 40, 3–73.

Grinblatt, Mark, Gergana Jostova, and Alexander Philipov, 2018, Analyst bias and mispricing, Working paper, University of California, Los Angeles.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales, 2015, The value of corporate culture, *Journal of Financial Economics* 117, 60–76.

Gulen, Huseyin, and Mihai Ion, 2016, Policy uncertainty and corporate investment, *Review of Financial Studies* 29, 523–564.

Guo, Li, Frank Weikai Li, and K.C. John Wei, 2019, Security analysts and capital market anomalies, *Journal of Financial Economics,* forthcoming.

Hanley, Kathleen Weiss, and Gerard Hoberg, 2019, Dynamic interpretation of emerging risks in the financial sector, *Review of Financial Studies* 32, 4543–4603.

Hartzell, Jay C., and Laura T. Starks, 2003, Institutional investors and executive compensation, *Journal of Finance* 58, 2351–2374.

Hartzmark, Samuel M., and David H. Solomon, 2018, Recurring firm events and predictable returns: The within-firm time series, *Annual Review of Financial Economics* 10, 499–517.

He, Jie (Jack), and Xuan Tian, 2013, The dark side of analyst coverage: The case of innovation, *Journal of Financial Economics* 109, 856–878.

Henry, Elaine, and Andrew J Leone, 2016, Measuring qualitative information in capital markets research: Comparison of alternative methodologies to measure disclosure tone, *Accounting Review* 91, 153–178.

Heston, Steven L., and Ronnie Sadka, 2008, Seasonality in the cross-section of stock returns, *Journal of Financial Economics* 87, 418–445.

Hirshleifer, David, Kewei Hou, Siew Hong Teoh, and Yinglei Zhang, 2004, Do investors overvalue firms with bloated balance sheets?, *Journal of Accounting and Economics* 38, 297–331.

Hirshleifer, David A., Danling Jiang, and Yuting Meng, 2019, Mood beta and seasonalities in stock returns, *Journal of Financial Economics,* forthcoming.

Hong, Harrison, and Jeffrey D. Kubik, 2003, Analyzing the analysts: Career concerns and biased earnings forecasts, *Journal of Finance* 58, 313–351.

Hou, Kewei, Haitao Mo, Chen Xue, and Lu Zhang, 2020, An augmented q-factor model with expected growth, *Review of Finance,* forthcoming.

Hou, Kewei, Chen Xue, and Lu Zhang, 2018, Replicating anomalies, *Review of Financial Studies,* forthcoming.

Hu, Grace Xing, Jun Pan, Jiang Wang, and Haoxiang Zhu, 2019, Premium for heightened uncertainty: Solving the FOMC puzzle, Working paper, Massachusetts Institute of Technology.

Hu, Minqing, and Bing Liu, 2004, Mining and summarizing customer reviews, in *Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '04, 168–177.

Huang, Allen H., Reuven Lehavy, Amy Y. Zang, and Rong Zheng, 2018, Analyst information discovery and interpretation roles: A topic modeling approach, *Management Science* 64, 2833–2855.

Huang, Allen H., Amy Y. Zang, and Rong Zheng, 2014, Evidence on the information content of text in analyst reports, *Accounting Review* 89, 2151–2180.

Jegadeesh, Narasimhan, Joonghyuk Kim, Susan D. Krische, and Charles M. C. Lee, 2004, Analyzing the analysts: When do recommendations add value?, *Journal of Finance* 59, 1083–1124.

Jegadeesh, Narasimhan, and Sheridan Titman, 1990, Evidence of predictable behavior of security returns, *Journal of Finance* 45, 881–898.

Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *Journal of Finance* 48, 65–91.

Jens, Candace E., 2017, Political uncertainty and investment: Causal evidence from U.S. gubernatorial elections, *Journal of Financial Economics* 124, 563–579.

Jensen, Michael, and Kevin Murphy, 1990, Performance pay and top-management incentives, *Journal of Political Economy* 98, 225–64.

Johnson, Travis L., Jinhwan Kim, and Eric C. So, 2019, Expectations management and stock returns, *Review of Financial Studies,* forthcoming.

Joos, Peter, Joseph D. Piotroski, and Suraj Srinivasan, 2016, Can analysts assess fundamental risk and valuation uncertainty? An empirical analysis of scenario-based value estimates, *Journal of Financial Economics* 121, 645–663.

Keloharju, Matti, Juhani T. Linnainmaa, and Peter Nyberg, 2016, Return seasonalities, *Journal of Finance* 71, 1557–1590.

Keloharju, Matti, Juhani T. Linnainmaa, and Peter Mikael Nyberg, 2019, Return seasonalities due to risk or mispricing? Evidence from seasonal reversals, *Journal of Financial Economics,* forthcoming.

Kirk, Marcus P., and Stanimir Markov, 2016, Come on over: Analyst/investor days as a disclosure medium, *Accounting Review* 91, 1725–1750.

Li, Feng, Russell Lundholm, and Michael Minnis, 2013, A measure of competition based on 10-K filings, *Journal of Accounting Research* 51, 399–436.

Li, Kai, Feng Mai, Rui Shen, and Xinyan Yan, 2020, Measuring corporate culture using machine learning, Working paper, University of British Columbia.

Linnainmaa, Juhani T., and Yingguang Zhang, 2019, The earnings announcement return cycle, Working paper, Dartmouth College.

Lins, Karl V., Henri Servaes, and Ane Tamayo, 2017, Social capital, trust, and firm performance: The value of corporate social responsibility during the financial crisis, *Journal of Finance* 72, 1785–1824.

Livnat, Joshua, and Richard R. Mendenhall, 2006, Comparing the post–earnings announcement drift for surprises calculated from analyst and time series forecasts, *Journal of Accounting Research* 44, 177–205.

Loh, Roger K., and Rene M. Stulz, 2018, Is sell-side research more valuable in bad times?, *Journal of Finance* 73, 959–1013.

Loughran, Tim, and Bill McDonald, 2011, When is a liability not a liability? textual analysis, dictionaries, and 10-Ks, *Journal of Finance* 66, 35–65.

Loughran, Tim, and Jay R. Ritter, 1995, The new issues puzzle, *Journal of Finance* 50, 23–51.

Lucca, David O., and Emanuel Moench, 2015, The pre-FOMC announcement drift, *Journal of Finance* 70, 329–371.

Malmendier, Ulrike, and Devin Shanthikumar, 2014, Do security analysts speak in two tongues?, *Review of Financial Studies* 27, 1287–1322.

Mayew, William J., Nathan Y. Sharp, and Mohan Venkatachalam, 2013, Using earnings conference calls to identify analysts with superior private information, *Review of Accounting Studies* 18, 386–413.

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

McLean, R. David, and Jeffrey Pontiff, 2016, Does academic research destroy stock return predictability?, *Journal of Finance* 71, 5–32.

Mikolov, Tomas, Ilya Sutskever, Kai Chen, Greg Corrado, and Jeffrey Dean, 2013, Distributed representations of words and phrases and their compositionality, in *Proceedings of the 26th International Conference on Neural Information Processing Systems–Volume 2*, NIPS'13, 3111–3119.

Newey, Whitney K., and Kenneth D. West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703–708.

Ohlson, James A., 1980, Financial ratios and the probabilistic prediction of bankruptcy, *Journal of Accounting Research* 18, 109–131.

Piotroski, Joseph D., and Eric C. So, 2012, Identifying expectation errors in value/glamour strategies: A fundamental analysis approach, *Review of Financial Studies* 25, 2841–2875.

Price, S. McKay, James S. Doran, David R. Peterson, and Barbara A. Bliss, 2012, Earnings conference calls and stock returns: The incremental informativeness of textual tone, *Journal of Banking and Finance* 36, 992–1011.

Ritter, Jay R., 1991, The long-run performance of initial public offerings, *Journal of Finance* 46, 3–27.

Robert, Novy-Marx, 2013, The other side of value: The gross profitability premium, *Journal of Financial Economics* 108, 1–28.

Sloan, Richard G., 1996, Do stock prices fully reflect information in accruals and cash flows about future earnings?, *The Accounting Review* 71, 289–315.

Stambaugh, Robert F., and Yu Yuan, 2016, Mispricing factors, *Review of Financial Studies* 30, 1270–1315.

Tetlock, Paul C., 2007, Giving content to investor sentiment: The role of media in the stock market, *Journal of Finance* 62, 1139–1168.

Tetlock, Paul C., Maytal Saar-Tsechansky, and Sofus Macskassy, 2008, More than words: Quantifying language to measure firms' fundamentals, *Journal of Finance* 63, 1437–1467.

Tian, Xuan, and Tracy Yue Wang, 2011, Tolerance for failure and corporate innovation, *Review of Financial Studies* 27, 211–255.

Titman, Sheridan, K. C. John Wei, and Feixue Xie, 2004, Capital investments and stock returns, *Journal of Financial and Quantitative Analysis* 39, 677–700.

Womack, Kent L., 1996, Do brokerage analysts' recommendations have investment value?, *Journal of Finance* 51, 137–167.

Xing, Yuhang, 2008, Interpreting the value effect through the q-theory: An empirical investigation, *Review of Financial Studies* 21, 1767–1795.

Zhong, Rong (Irene), 2018, Transparency and firm innovation, *Journal of Accounting and Economics* 66, 67–93.