ESSAYS ON CORPORATE TEXTUAL DISCLOSURES

By

ANKIT JAIN

A DISSERTATION

Submitted in Partial Fulfilment of the Requirements

for the

Fellow Programme in Management in the Accounting Area

at the Indian School of Business

Indian School of Business

Dissertation Prepared by: Ankit Jain

Title: "Essays on Corporate Textual Disclosures"

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Fellow Programme in Management in the Accounting area by:

Sanjay Kallapur, Chairperson
Krishnamurthy Subramanian, Member
Hariom Manchiraju, Member
Shyam V. Sunder, Outside Member

June 18, 2018

Fellow Programme in Management

Result of Dissertation Defense

STUDENT NAME : Ankit Jain

ACADEMIC PROGRAM : Fellow Programme in Management

STUDENT ID : 51250004

FIELD OF SPECIALIZATION : Accounting

PROGRAMME START DATE: 23 August 2012

DISSERTATION TITLE: ESSAYS ON CORPORATE TEXTUAL DISCLOSURES

DISSERTATION DEFENSE DATE: **04, May 2018**

Results of Defense Pass ☑ Fail ☐

Certification

We, the undersigned members of the defense committee, certify that the candidate has passed the oral defense and the result indicated above was reached after a vote of the committee membership

SIGNATURES OF THE DISSERTATION COMMITTEE MEMBERS

CHAIR Sanjay Kallapur, Professor of Accounting Deputy Dean - Faculty Review & FPM, Academic Director-IEMS	S.Kallapur
COMMITTEE MEMBER Krishnamurthy Subramanian, Associate Professor of Finance, Executive Director, Centre for Analytical Finance, ISB	K Subramanian
COMMITTEE MEMBER Hariom Manchiraju, Assistant Professor of Accounting, ISB	Hariom

COMMITTEE MEMBER	
Shyam V. Sunder Associate Dean, Eller MBA Programs McClelland Professor of Accounting	

An Abstract of the Dissertation of

Ankit Jain for the Fellow Programme in Management

in the Accounting Area

Title: ESSAYS ON CORPORATE TEXTUAL DISCLOSURES

The dissertation consists of two chapters on corporate textual disclosures. In the first chapter, I examine the role of secondary market participants in shaping disclosure tone. Managers use disclosure tone as a strategic tool to manage investors' expectation and demand for information. I provide evidence that investors' actions can in turn exercise a disciplining effect on tone management. Using an exogenous relaxation in the short-selling constraints from Regulation-SHO, I find that short-selling pressure reduces tone management. Greater short selling pressure results in less optimism in tone unrelated to fundamentals, and in disclosures about the past rather than the future. This reduction in tone management is stronger for firms with higher short-selling constraints pre-SHO, overoptimistic and overconfident managers, and lower analyst coverage.

In the second chapter, I and my co-authors construct a weighted word-count based measure to capture the quantity of a firm's 10-K narrative R&D-related disclosures, and document a persistent and significant (statistical and economic) negative association with subsequent firm profitability (ROA) during 1993-2006. These results stand in contrast to prior literature on 10-K narrative disclosures, across disciplines, where such disclosures have been found to convey meaningful information (via positive association) about current and future firm fundamentals. We argue that the unique characteristics of the R&D disclosure environment make it difficult for managers to develop skilled intuitive judgments about the

outcome of their firms’ R&D investments, which in turn could adversely affect the accuracy and credibility of these disclosures. The disclosure-type and features of the environment are thus important considerations in this area.

ACKNOWLEDGEMENTS

I thank the members of my dissertation committee, Sanjay Kallapur (Chair), Shyam V. Sunder, Krishnamurthy Subramanian, and Hariom Manchiraju for their guidance and support throughout the course of this research. I am also grateful to Prachi Deuskar and Ramabhadran S. Thirumalai for insightful discussions and comments. I also thank workshop participants at the Indian School of Business and fellow doctoral students for useful comments. Finally, I would like to thank my parents, wife, and daughter for their love, patience, and support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

DISCLOSURE TONE AND SHORT-SELLING PRESSURE: EVIDENCE FROM REGULATION-SHO

1. INTRODUCTION

Disclosure tone, i.e., optimism or pessimism of qualitative managerial communications, is one of the important characteristics of textual disclosures (Li 2010a). It has a significant impact on stock prices (Henry 2008; Loughran and McDonald 2011). Since accounting numbers and analysts' estimates are either incomplete or biased, managers use disclosure tone to convey a signal of private information (Davis, Piger, and Sedor 2012; Li 2010a). However, managers also use it as a manipulation tool to manage investors and analysts' expectations by employing an overly optimistic or overly pessimistic tone (Huang, Teoh, and Zhang 2014) or by structuring their tone[1] (Allee and DeAngelis 2015). These voluntary narrative disclosures from earnings press releases and earnings conference calls are not subject to explicit rules about the disclosure format, which gives managers flexibility in disclosing tone through these outlets since they can choose what topics to cover, and how to frame specific information (Henry 2008). Unlike accounting numbers where auditors have accounting standards as a benchmark (Lo 2008), there is no benchmark for disclosure tone, which makes its verification harder (Cazier, Merkley, and Treu 2016). Since regulation and verification of disclosure tone is difficult, I examine if sophisticated players from the secondary market could play a disciplinary role in shaping disclosure tone.

Prior accounting and finance research examines how managers use disclosure tone to influence secondary market participants such as its impact on investors' reaction, information environment, and the cost of capital (Kothari, Li, and Short 2009; Loughran and McDonald

[1] Allee and DeAngelis (2015) show that managers deliberately spread optimistic words in the conference calls to influence the perception of analysts and investors.

2011; Feldman, Govindaraj, Livnat, and Segal 2010), whereas the role of the secondary market participants in disclosure tone has not been given attention. Any relationship between disclosure tone and investors' action is subject to endogeneity, and Li (2010b) acknowledges the need for better empirical identification of research employing textual analysis of corporate disclosures. I fill the gap in the literature by examining if participants in the secondary market disciplines managerial tone and I use an exogenous shock to secondary market trading, to make a causal inference.

I choose short-selling pressure from the secondary market for my study as short-sellers are an important group of traders[2] who are sophisticated (Engelberg, Reed, and Ringgenberg 2012). Moreover, short-sellers reveal a signal of private information in the secondary market (Cohen, Diether, and Malloy 2007) and trade on the qualitative news (Beschwitz, Chuprinin, and Massa 2017). It is not clear ex-ante, how short selling pressure would affect the disclosure tone of managers. Managers have incentives to engage in tone management (Huang, Teoh, and Zhang 2014; Arsalan-Ayaydin, Boudt, and Thewissen 2015). However, overly optimistic tone can increase litigation (Rogers, Buskirk, and Zechman 2011) and reputational costs[3]. Therefore, managers face a trade-off between potential benefit and potential cost due to an overly optimistic tone and higher pressure from short-sellers is likely to affect this trade-off.

Prior academic evidence suggests that short-sellers improve stock price informativeness[4]. If short-sellers reveal a signal of negative private information to the market,

[2] Short-sellers accounted for approximately one third of share volume (31 percent) for NASDAQ-listed stocks and one fourth (24 percent) for NYSE-listed stocks in 2005 (Diether, Lee, and Werner 2009).
[3] Graham, Harvey, and Rajgopal (2005) argue that reputation is one of the main concerns for managers when they make disclosures.
[4] Theoretical (Miller 1977; Diamond and Verrecchia 1987; Duffie, Garleanu, and Pedersen 2002; Hong, Scheinkman, and Xiong 2006) and empirical (Bris, Goetzmann, and Zhu 2007; Boehmer, Jones, and Zhang 2008; Saffi and Sigurdsson 2010; Boehmer and Wu 2013) evidence suggests that short-sellers improve informational efficiency. Curtis and Fargher (2014) argue that short-sellers do not push the price below fundamental values.

managers could find it harder to mislead investors by employing an overly optimistic tone due to litigation and reputational concerns (*Disciplining Effect*). Prior studies document the disciplining effect of short-sellers in the context of timely disclosure of bad news earnings guidance (Clinch, Li, and Zhang 2016). On the other hand, regulators and practitioners believe that short-sellers destabilize the financial markets (Lamont 2012) and can also make the stock price overly sensitive to news (Hong, Kubik, and Fishman 2012; Savor and Gamboa-Cavazos 2011). Thus, when short-sellers create a downward pressure on the stock price (Mitchell, Pulvino, and Stafford 2004), managers could increase the optimistic tone strategically by either focusing on positive outcomes or obfuscating negative outcomes (*Stock Price Pressure Effect*). Li and Zhang (2015) argue that managers reduce the precision of bad news earnings guidance when the short-selling pressure is high, which makes the stock price less sensitive to the bad news. Thus, how short-selling pressure affects the disclosure tone of managers is an empirical question. In the current study, I test the above two competing hypothesis (*Disciplining Effect* and *Stock Price Pressure Effect*)[5].

The relationship between short-selling activity and disclosure tone is subject to endogeneity due to reverse causality and omitted variables. The overly optimistic tone could invite the attention of short-sellers (Blau, DeLisle, and Price 2015) and hence the direction of causality instead could be from disclosure tone to the short-selling activity. It is also possible that short-sellers take a position in a firm if they possess some negative private information and hence the relationship could be driven by omitted variables. Another possibility is that firms deliberately change the short-selling constraints in the secondary market (Lamont 2012). To address these potential endogeneity issues, I utilize a regulation-induced

[5] Blau, DeLisle, and Price (2015) find that short-sellers identify inflated talks only when firms have high earnings surprise. However, they do not examine how does ex-ante short selling pressure affects qualitative disclosures.

exogenous shock from Regulation-SHO to the short-selling constraints and employ a difference-in-differences approach.

Regulation-SHO relaxed short sale restrictions by abolishing the uptick rule[6] during the period May 2, 2005 to July 6, 2007, for a set of randomly selected 986 pilot stocks from the Russell 3000 index. While pilot firms are in the treatment group in my study, other firms from the same Russell 3000 index are in the control group. Prior research documents an increase in short sale activity immediately following the implementation date of the pilot program on May 2, 2005 (Diether, Lee, and Werner 2009). I find that monthly short interest position increases by 6.8 percent for pilot firms as compared to control firms. There was no increase in the SEC scrutiny (Hope, Hu, and Zhao 2017) and investors' attention (Fang, Huang, and Karpoff 2016) for pilot firms during Regulation-SHO. Thus, regulation only increased the short-selling pressure for pilot firms and provides an ideal setting to examine the aforementioned hypothesis.

I test my hypothesis using a sample of 4,647 firm-year observations from 1,327 unique non-financial firms from the United States over the period 2002-2007. I analyze the content of earnings conference calls to measure disclosure tone as the conference call is one of the most important venues for company management to communicate its message (Brown, Call, Clement, and Sharp 2017; Frankel, Mayew, and Sun 2010) and also a proxy for voluntary disclosures. I calculate tone as the difference between the count of optimistic and pessimistic words in the presentation section and scale it by the total count of optimistic and pessimistic words. I employ the financial dictionary from Loughran and McDonald (2011) to classify words as optimistic or pessimistic. Since firm fundamentals also affect the tone of managers, I decompose tone into two parts - normal tone and abnormal tone. I follow Huang, Teoh, and Zhang (2014) and run annual cross-sectional regressions of tone on underlying

[6] Uptick rule states that a stock can be sold short only at a price which is above the last traded price of the stock.

firm fundamentals. Normal tone is the expected value of tone from this model and thus it captures a part of tone that can be justified by firm fundamentals. Abnormal tone is the residual from the model and thus captures a part of tone that is discretionary. This is the main variable of interest in my study.

I obtain a negative and significant (statistical and economic) association between short-selling pressure and abnormal tone from the difference-in-differences analysis. The negative association is robust to firm (or industry) and year fixed effects. Specifically, I find that managers reduce abnormal tone by 14.5 percent of its standard deviation and thus the effect is economically significant. I do not find any significant effect of short-selling pressure on the normal tone. Hence, the short-selling pressure only reduces the discretionary part of tone. In another test, I focus on overoptimistic or overconfident managers as these managers are more likely to employ overly optimistic disclosure tone. I find that the negative relationship between short-selling pressure and abnormal tone exists only for firms with overoptimistic or overconfident managers. This evidence rules out the possibility that managers become conservative in tone and engage in downward tone management when the short-selling pressure is high. These findings are consistent with the disciplining effect hypothesis.

If the negative relationship between investors' activity (short-selling pressure) and abnormal tone is due to the disciplining effect, it should be only for the non-forward-looking disclosures as Private Securities Litigation Reform Act (PSLRA, 1995) provides safe harbor provisions to forward-looking disclosures. In fact, Cazier, Merkley, and Treu (2016) argue that only the tone of non-forward-looking disclosures increases litigation risk. I indeed find that negative association between short selling pressure and abnormal tone exists only for the non-forward-looking disclosures. In another test, I find that the negative association with abnormal tone is significant only for those firms that face higher short-selling constraints ex-ante. This finding suggests that short-selling pressure indeed drives the disciplining effect.

In addition to disclosure tone, I also examine the impact of short selling pressure on the structure of tone. Allee and DeAngelis (2015) show that managers create a positive perception among analysts and investors by deliberately spreading optimistic words (positive tone dispersion) during earnings conference calls and argue that tone dispersion has an additional effect on investors' perception, which is incremental to the level of tone. I calculate tone dispersion measures using their methodology and find that managers reduce the dispersion of positive tone when short-selling pressure is high. I do not find any significant effect on negative tone dispersion. I also examine the immediate market reaction to tone dispersion and find that investors' response to positive tone dispersion decreases in the presence of short-selling pressure. This evidence suggests that investors can see through the tone dispersion in the presence of short-sellers and are careful in their response.

From the cross-sectional analyses, I find that the effect of short-selling pressure is significant only for affected firms with lower analyst coverage. This finding suggests that short-sellers substitute for poor information environment and is consistent with the prior evidence from Pownall and Simko (2005), which shows that investors give more importance to the signal from short-sellers when the analyst coverage is lower. Overall, my findings are consistent with the disciplining effect of short-selling pressure and are robust to alternative explanations, additional controls, and cross-sectional placebo tests.

This study contributes to the literature in several ways. First, I contribute to the literature on disclosure tone. My study is the first one to document the disciplining role of secondary market participants on tone management and provide the causal evidence on the secondary market determinants of disclosure tone. Prior studies have focused on the consequences of disclosure tone such as investors' reaction, the cost of capital, information environment, and competitor behavior (Li 2010b; Loughran and McDonald 2011; Kothari, Li, and Short 2009; Durnev and Mangen 2011). Prior work shows how managers mislead investors by using overly optimistic or overly pessimistic tone (Huang, Teoh, and Zhang 2014); inviting favorable analysts to earnings conference calls who ask positive questions (Cohen, Lou, and Malloy 2014); structuring their tone (Allee and DeAngelis 2015); and blaming external factors when performance is poor (Zhou 2014). I show that trading activity in the secondary market could in turn discipline tone management. While Rogers, Buskirk, and Zechman (2011) argue that litigation is an ex-post disciplining mechanism for disclosure tone, my findings suggest that ex-ante short-selling pressure could increase the perceived litigation concerns for management.

Li and Zhang (2015) show that managers increase the complexity (fog index) of bad news annual reports when short-selling pressure is high and claim that their findings are generalizable to other disclosures. I focus on tone and its structure and find evidence consistent with the disciplining effect rather than obfuscation.

Second, it adds to studies on short-sellers in the accounting and finance literature. Early research shows that short-sellers predict future negative events (Karpoff and Lou 2010; Christophe, Ferri, and Angel 2004; Desai, Krishnamurthy, and Venkataraman 2006; Christophe, Ferri, and Hsieh 2010; Kecskes, Mansi, and Zhang 2012). This literature assumes that short-sellers only affect the information flow into the market. However, recent evidence suggests that short-selling pressure can also directly influence the behavior of firm managers. Massa, Zhang, and Zhang (2015) and Fang, Huang, and Karpoff (2016) document the disciplining effect on earnings management due to short-selling pressure. My findings show that short selling pressure also disciplines tone management. Thus, I contribute to the policy debate on the controversial short-selling activity, by providing additional causal evidence on the benefits of short-selling regulation.

2. BACKGROUND AND HYPOTHESIS DEVELOPMENT

The importance of narrative disclosures has been documented in the literature (Loughran and McDonald 2016; Li 2010b). One of the important attributes of these narrative disclosures is tone that captures the optimistic or pessimistic sentiment (Li 2010a). Prior evidence shows that managers employ disclosure tone to give incremental information to investors regarding future firm performance and tone has positive association with future earnings and short-term market reaction (Henry 2008; Li 2010a; Davis, Piger, and Sedor 2012; Davis and Tama-Sweet 2012; Loughran and McDonald 2011; Price, Doran, Peterson, and Bliss 2012). However, Huang, Teoh, and Zhang (2014) argue that managers mislead investors by employing an overly optimistic tone around important events such as SEOs (Seasoned Equity Offerings) and M&As (Merger and Acquisitions), and overly pessimistic tone around stock option grants. In another study, Allee and DeAngelis (2015) argue that in addition to tone, placement of optimistic and pessimistic words in the conference call also influences the perception of investors and analysts. Thus, managers also use tone as a tool to mislead investors.

There is no formal guideline and benchmark for narrative disclosures which gives managers' flexibility in disclosing tone as they can choose what topics to cover and what topics to avoid. Managers need to follow Generally Accepted Accounting Principles (GAAPs) to report earnings. There are no such guidelines for narrative disclosures. Unlike quantitative disclosures, the verification of these narrative disclosures is also difficult. Although PCAOB (Public Company Accounting Oversight Board) recommends auditors to read the transcripts of earnings conference calls (Auditing Standard No. 12, PCAOB [2010]), narrative disclosures are not formally audited by external auditors (Hobson, Mayew, Peecher, and Venkatachalam 2017). Since disclosure tone is not monitored formally and difficult to be regulated, I examine if sophisticated players from the secondary market could play a disciplinary role in tone management. Prior literature has only focused on the secondary market consequences of disclosure tone (Henry 2008; Kothari, Li, and Short 2009; Loughran and McDonald 2011; Feldman, Govindaraj, Livnat, and Segal 2010; Price, Doran, Peterson, and Bliss 2012), whereas the role of secondary market participants in shaping disclosure tone has not been given attention.

I extend prior research by examining the disciplinary role of secondary market participants in disclosure tone. I specifically focus on short-sellers as they are an important group of traders in the secondary market and do trade on the qualitative news (Beschwitz, Chuprinin, and Massa 2017). They accounted for approximately 31 percent of the trading volume for NASDAQ-listed stocks and 24 percent of the trading volume for NYSE-listed stocks (Diether, Lee, and Werner 2009). Moreover, Boehmer, Jones, and Zhang (2008) find that 75 percent of all short-sales are executed by institutions. Prior evidence suggests that short-sellers improve stock price informativeness[7], have superior ability to process publicly available information (Engelberg, Reed, and Ringgenberg 2012), serve as important information intermediaries (Pownall and Simko 2005), and reveal private information (Coehn, Diether, and Malloy 2007). If short-sellers reveal a signal of private information in the secondary market, investors will be more informed regarding the firm fundamentals. It will also reduce the benefits to managers in communicating the results in a favorable light and hiding unfavorable information. Thus, managers will find it difficult to mislead investors by employing an overly optimistic tone due to litigation (Rogers, Buskirk, and Zechman

[7] Theoretical (Miller 1977; Diamond and Verrecchia 1987; Duffie, Garleanu, and Pedersen 2002; Hong, Scheinkman, and Xiong 2006) and empirical (Bris, Goetzmann, and Zhu 2007; Boehmer, Jones, and Zhang 2008; Saffi and Sigurdsson 2010; Boehmer and Wu 2013) evidence suggests that short-sellers improve informational efficiency.

2011) and reputational concerns. Massa, Zhang, and Zhang (2015) and Fang, Huang, and Karpoff (2016) document the disciplining effect of short-sellers on earnings management. I test if the short-selling activity also disciplines tone management (*Disciplining effect hypothesis*).

It is not obvious that short-sellers could discipline tone management because tone management is subtle and could be difficult to detect. Since short-sellers also make the stock price overly sensitive to news (Hong, Kubik, and Fishman 2012; Savor and Gamboa-Cavazos 2011), managers could increase the tone management in the presence of short-selling activity by either focusing more on positive outcomes or obfuscating negative outcomes. This could help managers in reducing the downward pressure on stock prices created by short-sellers. Li and Zhang (2015) argue that when the short-selling pressure is high, managers obfuscate bad news by selectively reducing the precision of bad news earnings forecasts and increasing the complexity of annual reports. Similarly, short-selling activity could increase the tone management (*Stock Price Pressure Hypothesis*).

Thus, it is not clear ex-ante how short-selling activity would affect the tone management. My hypothesis tests the relationship between short-selling activity and tone management.

H1A (Disciplining Effect Hypothesis): Short-selling activity is negatively associated with tone management.

H1B (Stock Price Pressure Hypothesis): Short-selling activity is positively associated with tone management.

Studying only earnings forecasts and readability of annual reports does not give a complete picture of firms' disclosures. Moreover, the decision to issue earnings forecasts largely differs from the decision to disclose other qualitative disclosures (Bozanic, Roulstone, and Buskirk 2017). I fill the gap in the literature by studying the effect of short-selling activity on another important attribute of disclosures i.e. disclosure tone.

3. EMPIRICAL METHODOLOGY

3.1 Construction of Tone Variables

I analyze the content of earnings conference calls to measure disclosure tone as the conference call is one of the most important venues for company management to communicate its message (Brown, Call, Clement, and Sharp 2017; Frankel, Mayew, and Sun 2010) and also a proxy for voluntary disclosures. I quantify disclosure tone by employing the "bag of words" approach as Henry and Leone (2016) show that the "bag of words" tone measures are as powerful as the Bayesian machine-learning tone measures. Under the "bag of words" approach, every document is represented by the words it contains, ignoring its ordering and grammar. I parse earnings conference call scripts by writing Python programs. Each conference call has two sections – introductory remarks followed by the Q&A (Questions and Answers) section. I focus on the introductory remarks section[8] and remove the list of participants and legal disclaimers from each file. I count the frequency of optimistic and pessimistic words in each document using the financial dictionary from Loughran and McDonald (2011) and control for the negation of an optimistic word if it is accompanied by a negator within a distance of three words (Loughran and McDonald 2016)[9]. Following Allee and DeAngelis (2015), I remove the word "question" from the list of pessimistic words and ignore certain combinations of words that do not capture tone in the conference call[10].

---

[8] My results are robust when I use both sections in the analysis.

[9] I consider following negators: "*no*", "*not*", "*none*", "*neither*", "*never*", "*nobody*", "**n't*".

[10] I do not consider - "*good*" if it is followed by "*morning*", "*afternoon*", "*evening*", or "*day*"; "*effective*" if it is followed by "*income*", "*tax*", or "*rate*"; "*efficiency*" if it is followed by "*ratio*", and "*closing*" if it is followed by "*remark*" or "*remarks*".

I calculate *TONE* as the difference between the count of optimistic and pessimistic words divided by the total count of optimistic and pessimistic words in the conference call[11]. For ease of interpretation, I scale this measure by multiplying by 100. Since *TONE* is jointly determined by firm fundamentals and managerial discretion, following Huang, Teoh, and Zhang (2014), I decompose *TONE* into two parts – *NTONE* and *ABTONE*. I obtain *NTONE* as the predicted value of *TONE* and *ABTONE* as the residual from the following annual cross-sectional regressions:

$$\boldsymbol{TONE_{i,t}} = \alpha + \beta_0 EARNINGS_{i,t} + \beta_1 RETURNS_{i,t} + \beta_2 SIZE_{i,t} + \beta_3 BTM_{i,t} + \beta_4 STD\_RET_{i,t} + \beta_5 STD\_EARN_{i,t} + \beta_6 AGE_{i,t} + \beta_7 BUSINESS\_SEG_{i,t} + \beta_8 GEO\_SEG_{i,t} + \beta_9 LOSS_{i,t} + \beta_{10} \Delta EARNINGS_{i,t} + \beta_{11} AFE_{i,t} + \beta_{12} AF_{i,t} + \boldsymbol{\varepsilon_{it}} \quad \textbf{(1)}$$

The above regression controls for profitability (*EARNINGS*), earnings performance benchmarks (*LOSS*, *ΔEARNINGS*, *AFE* – analyst forecast error), expectation of future performance (*AF* – analyst consensus forecast for one year ahead), market performance (*RETURNS*), growth opportunities (*BTM*), operating and business risk environment (*STD_RET* and *STD_EARN*), age (*AGE*), and operating complexity of a firm (*BUSINESS_SEG* and *GEO_SEG*). Refer to the Appendix A for a detailed definition of these variables. Since *ABTONE* is the residual from the above regressions, it captures that part of *TONE* that cannot be justified by the underlying firm fundamentals and thus at the discretion of managers. Table 1.1(A) presents the results from regression (1). By construction, the average value of *ABTONE* is zero (Table 1.1(B)).

[INSERT TABLE 1.1 HERE]

Using specification (1), I also obtain *ABTONE* of forward-looking disclosures (*FLD ABTONE*) and non-forward-looking disclosures (*Non-FLD ABTONE*) separately. I use the

[11] My results are qualitatively similar when I scale *TONE* by the count of total words in the entire document.

dictionary from Muslu, Radhakrishnan, Subramanyam, and Lim (2014) to classify each sentence in the conference call, into a forward-looking and non-forward-looking statement. Then, I count the frequency of optimistic and pessimistic words separately in forward-looking statements and non-forward-looking statements.

In addition to *TONE*, which captures the word choices, I capture the structure of tone by calculating tone dispersion of optimistic (*POSITIVE_ARF*) and pessimistic (*NEGATIVE_ARF*) words using Allee and DeAngelis (2015). *POSITIVE_ARF* (*NEGATIVE_ARF*) tone dispersion captures the degree to which optimistic (pessimistic) words are evenly distributed throughout the introductory remarks session of the conference call. Allee and DeAngelis (2015) argue that managers create a positive influence among analysts and investors by increasing the dispersion of positive tone or by decreasing the dispersion of negative tone.

3.2 Research Design

The relationship between disclosure tone and short-selling pressure is subject to endogeneity due to reverse causality or omitted variables. The overly optimistic tone could also invite the attention of short-sellers (Blau, DeLisle, and Price 2015) and hence the direction of causality instead could be from disclosure tone to the short-selling activity. Moreover, as Lamont (2012) argues that the number of short positions depends upon both demand and supply of shorting. Therefore empirically estimating the cost and benefits of short-selling is tricky. It is possible that short-sellers take a position in a firm if they possess some negative private information and hence the relationship could be driven by omitted variables. Another possibility is that firms deliberately change the short-selling constraints in the secondary market (Lamont 2012). To address these potential endogeneity issues and make a causal inference, I utilize a regulation induced (Regulation-SHO) exogenous shock to short-selling constraints and employ a difference-in-differences approach.

Regulation-SHO, which was passed by the SEC, removed the uptick rule criteria for a set of randomly selected firms of the Russell 3000 index during the period May 2, 2005 to July 6, 2007. The uptick rule states that a stock can be sold short only at a price that is above the last traded price of the stock. (i.e. during an uptick). Prior evidence (Angel 1997; Alexander and Peterson 1999) shows that the uptick rule is binding and prevents the execution of short-sales. In an NYSE survey (Opinion Research Corporation 2008), 85 percent of surveyed managers were in favor of reinstitution of the uptick-rule "as soon as practical". This evidence shows that managers are concerned about short-selling activities and the uptick rule in particular. The SEC abolished the uptick rule for 986 stocks (called pilot stocks). While the list of pilot stocks was made public on July 28, 2004, the effective date was from May 2, 2005. Refer to Figure 1.1 for the timeline of Regulation-SHO. Pilot stocks constitute the treatment group in my study and the remaining stocks in the Russell 3000 index during the year 2004 and 2005 constitute the control group.

[INSERT FIGURE 1.1 HERE]

Prior evidence shows that short-sales and share volume increased around the implementation date of Regulation-SHO (Diether, Lee, and Werner 2009; SEC 2007). However, Grullon, Michenaud, and Weston (2015) show that anticipated future removal of short-selling constraints increases the short-selling activity in the current period. Since the list of pilot firms was announced well in advance of the effective date (Figure 1), short-selling activity increased well before the effective date once the list of firms was made public. Therefore for the difference-in-differences analysis, I consider the period before the announcement date as the pre-event period and period after the effective date as the post-event period. I use the fiscal year-end date of a firm to classify each observation into pre-event or post-event period and consider observations starting from the fiscal year 2002[12] till the ending date of Regulation-SHO (July 6, 2007). I exclude observations between the announcement and effective dates (July 28, 2004 to May 2, 2005, which I call the announcement period) as these observations cannot be classified into pre-event or post-event period. Additionally, I also exclude financial firms as calculation of tone for these firms poses a problem. Some words are perceived as pessimistic for non-financial firms, but they are not necessarily pessimistic for financial firms (Jegadeesh and Wu 2013).

Regulation-SHO increased the short-selling pressure only for pilot firms without affecting the scrutiny by the regulator or the attention of investors. Hope, Hu, and Zhao (2017) find no significant increase in the number of SEC comment letters and the numbers of topics covered in those comment letters and thus argue that SEC did not increase the scrutiny for pilot firms. In another study, Fang, Huang, and Karpoff (2016) do not find any increase in investors' attention. They use three proxies for investors' attention – the frequency with which a stock is searched on Google, total trading volume, and the number of forecasts by analysts. They do not find any evidence of lobbying for the Regulation-SHO pilot program. Moreover, I find that monthly short interest position increases by 6.8 percent for pilot firms as compared to control firms (Refer to Table A1). Thus, Regulation-SHO provides an ideal setting to examine the causal effect of short-selling pressure on disclosure tone.

I run the following empirical specification for the difference-in-differences analysis:

$$\boldsymbol{ABTONE_{i,t}} = \alpha + \beta_1 POST_{i,t} + \beta_2 POST_{i,t} * PILOT_{i,t} + \beta_3 PILOT_{i,t} + Controls + \varepsilon_{i,t} \qquad (\mathbf{2})$$

where *POST* is a dummy variable which is equal to 1 during the post-event period and 0 otherwise; and *PILOT* is a dummy variable which is equal to 1 for pilot firms and 0 otherwise. Controls include firms (or industry) and year fixed effects. I do not include

[12] Conference call scripts are only available starting from the year 2002 (after SEC passed Regulation-FD).

dummy variable *PILOT* separately when I include firm fixed effects as firm fixed effects will absorb *PILOT* dummy.

The main coefficient of interest is β_2, which captures the causal impact of short selling pressure on *ABTONE* for pilot firms as compared to control firms. A positive and significant value of β_2 will be consistent with the *Stock Price Pressure Effect* (H1B), while a negative and significant value of β_2 will be consistent with the *Disciplining Effect* (H1A). An insignificant value of β_2 will imply that short-selling pressure does not affect disclosure tone.

4. DATA

I collect data from several sources. I hand collect annual conference call scripts from LEXISNEXIS database using Fair-Disclosure wire and extract the introductory remarks from each script by writing Python programs. I obtain accounting and short-interest data from COMPUSTAT, and stock return data from CRSP and then match observations from LEXISNEXIS and COMPUSTAT using the name of a company and earnings announcement date. I ensure that the difference between the conference call date in LEXISNEXIS and the earnings announcement date from COMPUSTAT is not more than three days. I remove firms in the finance, insurance, and real estate sectors (SIC codes between 6000 and 6999). Finally, I obtain institutional ownership from THOMSON REUTERS, analyst forecasts from I/B/E/S, and index constituents for Russell 3000 from BLOOMBERG.

[INSERT FIGURE 1.2 HERE]

Figure 1.2 describes the sample selection process. I start with 7,571 observations after merging LEXISNEXIS and COMPUSTAT data for non-financial firms. I exclude 1,471 observations during the announcement period and consider only those firms that exist in pre-event as well as post-event periods. The final sample contains 4,647 observations from 1,327 unique firms. 386 unique firms are in the treatment group, and 941 firms are in the control group.

[INSERT TABLE 1.2 HERE]

Table 1.2 presents summary statistics of my sample. I winsorize all continuous variables at 1 percent to mitigate the effect of outliers. As shown in Panel A, the average length of conference call transcript is 3,040 words, out of which 57 words are optimistic and 26 words are pessimistic. These values are very similar to prior studies employing earnings conference call data (Blau, DeLisle, and Price 2015; Matsumoto, Pronk, Roelofsen 2011). Summary statistics of tone dispersion measures *POSITIVE_ARF* and *ADJ_POSITIVE_ARF* are also very similar to that of Allee and DeAngelis (2015). I decompose total variation in *TONE* and *ABTONE* and find that about 53 percent variation in these variables is time-series (Panel B). This variation makes these tone variables a good proxy for discretionary tone and rules out the possibility that tone in the conference call could be boilerplate.

Panel C in Table 1.2 shows the univariate comparison between pilot and control firms during the *PRE* period. *PRE* period is the one-year period ending on the Regulation-SHO announcement date. Although the selection into the Regulation-SHO program was random, there are some differences between pilot and control firms. Pilot firms have larger market capitalization, lower return and earnings volatility, and are older as compared to control firms. Therefore, I control for all of these variables in my difference-in-differences regressions[13].

5. RESULTS AND DISCUSSION

5.1 Short-selling Pressure and Tone

[INSERT FIGURE 1.3 HERE]

[13] Prior studies also document some differences in pilot and control firms (Li and Zhang 2015; Clinch, Li, and Zhang 2016).

I provide visual evidence of the impact of short-selling pressure on *TONE* in Figure 1.3. As can be seen, the level of *TONE* is comparable for pilot and control firms before Regulation-SHO. *TONE* increases during Regulation-SHO period for the pilot as well as control firms. However, the increase is smaller in magnitude for pilot firms. This suggests a disciplinary role of short-sellers on pilot firms.

[INSERT TABLE 1.3 HERE]

I provide formal evidence in Panel A of Table 1.3, which presents results of the difference-in-differences analysis from regression model (2). The coefficient of *PILOT * POST* captures the relationship between short-selling pressure and *TONE.* This coefficient is negative and statistically significant at the 5 percent level across all specifications. This negative relationship is robust to various fixed effects and control variables. I control for industry and year fixed effects in column (1), industry-year fixed effects in column (2), and firm and year fixed effects in columns (3)-(4), and cluster the standard errors at the firm level in all specifications. In column (4), I additionally control for all the determinants of *TONE* from the expected tone model (1) and find similar results. For brevity, I do not show coefficients for these control variables in column (4). I find that managers of pilot firms decrease their *TONE* in earnings conference calls by 3.91 (column 4) during Regulation-SHO as compared to control firms. This reduction in *TONE* corresponds to 14.3 percent of its standard deviation[14]. Thus, this effect is economically significant as well. This result is consistent with the *Disciplining Effect* hypothesis (H1A).

Next, I separately examine the effect of short-selling pressure on *NTONE* and *ABTONE*. While *NTONE* is the expected or predicted value of *TONE*, *ABTONE* is the residual from regression model (1). If the relationship between short-selling pressure and

[14] Standard deviation of *TONE* is 27.37 (Table 1.2(A)). The coefficient of *PILOT * POST* in column (4) of Table 1.3 is 3.91 which is 14.3 percent of 27.37.

TONE is due to the disciplining effect, it should affect *ABTONE* without affecting *NTONE*, as *ABTONE* captures the discretionary component of *TONE*. I report regression results of *NTONE* in Table B2 (Appendix B). The coefficient of *PILOT * POST* is insignificant across all specifications. Thus, *NTONE* is not affected by short-selling pressure.

Panel B of Table 1.3 reports the relationship between short-selling pressure and *ABTONE*. I find the main coefficient to be negative, which is statistically significant as well as economically meaningful, and robust to various fixed effects specifications. Specifically, I find that managers of pilot firms decrease *ABTONE* by 14.5 percent of its standard deviation (based on coefficient in column 5). The difference-in-differences design requires the parallel trend assumption. I test this assumption by adding an interaction term of *PILOT* and *PRE* variables in specifications (3) and (5), where *PRE* is a dummy variable which is equal to 1 for the one-year period ending on the announcement date of Regulation-SHO. Thus, *PILOT * PRE* captures trends in *ABTONE* between pilot and control firms just before Regulation-SHO came into effect. I find the coefficient of *PILOT * PRE* to be statistically insignificant (column (3) and (5)). Thus, there is no pre-event trend in *ABTONE* between pilot and control firms[15]. These results provide causal evidence that short-sellers discipline managers and that they help reduce tone management

5.2 Overoptimistic and Overconfident Managers

If the negative relationship between *ABTONE* and short-selling pressure is due to the disciplining effect, then the short-selling pressure would only affect managers who are either overoptimistic or overconfident. I test this hypothesis by conducting a sub-sample analysis and separately running regression model (2) for firms with overoptimistic/overconfident

[15] In another unreported test, I split *POST* period into two parts – *POST1* and *POST2* and then run similar specification. I find that the effect of short-selling pressure is stronger in the *POST2* period i.e. the second half of the post Regulation-SHO The coefficient of *PILOT * POST2* is higher in magnitude and statistically different from that of *PILOT * POST1*. Li and Zhang (2015) also find stronger effect of Regulation-SHO on management forecasts in latter half of the treatment period.

managers and firms without overoptimistic/overconfident managers. I identify overoptimistic managers during the *PRE* period by employing positive *ABTONE* ($ABTONE \geq 0$) as a proxy for overoptimistic tone. I use three different proxies from prior literature to identify overconfident managers during the *PRE* period. My first proxy is *RETAINER* from Sen and Tumarkin (2015) and the other two proxies *OC_FIRM4* and *OC_FIRM5* are from Schrand and Zechman (2012).

[INSERT TABLE 1.4 HERE]

Column (1) of Table 1.4 reports results from the difference-in-differences analysis for firms with overoptimistic managers and column (5) reports results for firms without overoptimistic managers. I find that the negative relationship between short-selling pressure and *ABTONE* is statistically significant only for firms with overoptimistic managers. Similarly, I find that short-selling pressure disciplines overconfident managers (Columns (2) – (4)) and there is no effect on managers who are not overconfident (Columns (6)-(8)). *PILOT * POST* is significant at 5 percent level in columns (1) to (4) and is insignificant in columns (5) to (8). This evidence is consistent with the disciplining role of short-sellers and rules out the possible alternative explanation that in the presence of short-selling pressure, managers engage in downward tone management or increase the conservatism in their tone.

5.3 Type of Disclosure Tone

Optimistic tone of managers increases the subsequent litigation risk. Rogers, Buskirk, and Zechman (2011) document a positive association between disclosure tone and subsequent litigation from shareholders. However, Private Securities Litigation Reform Act (PSLRA, 1995) provides safe harbor provisions to forward-looking disclosures, and therefore the positive relationship between disclosure tone and subsequent litigation exists only for tone from non-forward-looking disclosures (Cazier, Merkley, and Treu 2016). Thus, if the negative relationship between short-selling pressure and *ABTONE* is due to the disciplining effect, it should be only for the non-forward-looking disclosures.

I examine the association of short-selling pressure with *ABTONE* for forward-looking disclosures and non-forward-looking disclosures, separately. I calculate abnormal tone from forward-looking disclosures (*FLD ABTONE*) and non-forward looking disclosures (*Non-FLD ABTONE*) by employing the dictionary of forward-looking phrases from Muslu, Radhakrishnan, Subramanyam, and Lim (2014) and classifying each statement of the conference call into a forward-looking or non-forward statement. Table 1.5 presents the regression results[16]. I find that short-selling pressure affects only *Non-FLD ABTONE*. It has no significant impact on *FLD ABTONE*. These findings are consistent with the disciplinary role of short-sellers[17] and suggest that ex-ante short-selling pressure could increase the perceived litigation concerns for management.

[INSERT TABLE 1.5 HERE]

5.4 Sub-Sample Analyses

I create sub-samples based on short-selling constraints at the time of Regulation-SHO announcement (i.e. during the *PRE* period). Low institutional ownership is a proxy for short-selling constraints as many institutional investors lend shares to short-sellers (D'Avolio 2002). Smaller firms are also harder to short (Grullon, Michenaud, and Weston 2015). I create two sub-samples based on the median values of these two proxies of short-selling constraints during the *PRE* period and run difference-in-differences regressions. As shown in Panel A of Table 1.6, I find that the relationship between short-selling pressure and *ABTONE* is statistically significant only for those firms that faced higher short-selling constraints i.e. firms with low institutional ownership and those smaller in size. Although results for firms

[16] I remove 53 observations out of total 4,599 observations based on the highest values of |*DFBETA*| to ensure robustness of my results (top 1.1 percent observations). All results presented in this paper are robust to |*DFBETA*| sensitivity test using multiple cutoff values.

[17] In another unreported test, I find that the disciplining role of short-sellers exists only for firms in the high-litigation industries. I use industry based proxy from prior literature to capture litigation risk.

with low institutional holdings are statistically weaker and significant only at 10 percent (column 2), results for smaller firms are significant at 5 percent (column 4). Overall, the findings in Panel A are consistent with the fact that the impact of Regulation-SHO on short-selling activity is stronger for firms facing higher short-selling constraints ex-ante.

[INSERT TABLE 1.6 HERE]

Next, I create sub-samples based on the level of analyst coverage as prior evidence suggests that analysts play a significant monitoring role (Chen, Harford, and Lin 2015) and reduce the information asymmetry between investors and managers (Kelly and Ljungqvist 2012). I find that short-sellers discipline managers of those firms that have a lower analyst coverage (Panel B of Table 1.6). Thus, short-sellers substitute for a poor information environment. This finding is consistent with the prior evidence on the information intermediary role of short-sellers in Pownall and Simko (2005). They argue that investors give more importance to the signal from short-sellers when analyst coverage is lower.

5.5 Structure of Tone

In addition to tone, managers use other subtle techniques to amplify the effect of positive or negative news. They deliberately spread good news during the conference calls and bunch together bad news. This has an additional effect on investors' perception, which is incremental to the level of tone (Allee and DeAngelis 2015). Allee and DeAngelis (2015) capture positive tone dispersion (negative tone dispersion) by calculating the degree to which optimistic (pessimistic) words are evenly distributed throughout the entire conference call and show that tone dispersion influences the perception of investors and analysts.

[INSERT TABLE 1.7 HERE]

I examine how short-selling pressure affects the structure of tone. Using the difference-in-differences specification (Table 1.7), I find that managers reduce positive tone dispersion (*POSITIVE_ARF*). This relationship is robust to an alternative measure of positive tone dispersion (*ADJ_POSITIVE_ARF*). Since total words, as well as optimistic or pessimistic words, could also influence tone dispersion measures, I control for the length of conference calls (*Total Words*), the total count of optimistic words (*Optimistic Words*), and the total count of pessimistic words (*Pessimistic Words*) in the regression model. These coefficients load up with significant values. I also include all determinants of tone from expected tone model (1) as control variables. For brevity, I do not show coefficient estimates for these control variables in Table 1.7. Although the primary coefficient (*PILOT * POST*) is statistically significant only at 10 percent level, it is economically significant. Pilot firms decrease the dispersion of optimistic words (*POSITIVE_ARF*) by 12 percent of its standard deviation as compared to control firms. In untabulated results, I find that the effect of short-selling pressure on negative tone dispersion (*NEGATIVE_ARF*) is not significant.

[INSERT TABLE 1.8 HERE]

Next, I examine investors' reaction to tone dispersion (*POSITIVE_ARF*) in the presence of short-selling pressure. I measure investors' response by calculating cumulative abnormal returns around the three days window of earnings announcement date (*CAR[-1,+1]*). If short-sellers reveal a signal of private information to investors, then managers will find it difficult to mislead investors. For ease of interpretation, I use annual decile rankings (*D_POSITIVE_ARF*) of *POSITIVE_ARF* as an independent variable. Table 1.8 presents investors' reaction to *POSITIVE_ARF*. The main coefficient of interest is the triple interaction term – *PILOT * POST * D_POSITIVE_ARF*, which captures the incremental effect of investors' reaction for pilot firms after Regulation-SHO, due to positive tone dispersion. I additionally control for other double interaction terms and standalone terms as well. I add *EARNINGS* and *D_SUE* (annual decile ranking of standardized unexpected earnings) variables to control for other contemporaneous quantitative disclosures. I find the main coefficient to be negative (column 2) and statistically significant at 10 percent. I interpret this as an evidence of investors' being more informed in the presence of short-selling pressure and thus exercising caution when responding to positive tone dispersion. This result is robust to industry and year fixed effects and firm characteristics such as size (*SIZE*), growth opportunities (*BTM*), stock performance (*RETURNS*), stock return volatility (*STD_RET*) and earnings volatility (*STD_EARN*)[18].

5.6 Additional Analyses and Robustness Checks

I conduct additional tests to rule out alternative explanations of my results. Grullon, Michenaud, and Weston (2015) argue that short-selling pressure reduces the level of equity and debt financing. I control for these variables in the difference-in-differences specification. Additionally, I control for discretionary accruals because Massa, Zhang, and Zhang (2015) and Fang, Huang, and Karpoff (2016) show that managers reduce discretionary accruals in the presence of short-selling pressure. As shown in Table 1.9, I still obtain similar results.

[INSERT TABLE 1.9 HERE]

In addition to disclosure tone, I examine other characteristics of conference calls such as length and readability. I use total count of words and sentences in the introduction section as a proxy for length. I calculate readability using three proxies for business communication proposed by Loughran and McDonald (2014) – *Common Words*, *Financial Terminology*, and *Vocabulary*. Refer to the Appendix for a detailed definition of these variables. As shown in Table 1.10, length and readability of conference calls are not affected by the presence of short-selling pressure.

[18] In untabulated test, I also include another triple interaction term - *PILOT * POST * D_NEGATIVE_ARF* in the same specification. I still find the coefficient *PILOT * POST * D_POSITIVE_ARF* to be negative and significant. But, the coefficient of *PILOT * POST * D_NEGATIVE_ARF* is insignificant.

[INSERT TABLE 1.10 HERE]

As a further robustness check, I perform a cross-sectional placebo test. I randomly categorize firms in my sample as treatment and control firms and run the difference-in-differences regression model (2). I repeat the randomization process 1,000 times and perform a Monte-Carlo analysis. I obtain distribution of the main coefficient (β_2) from a Monte-Carlo analysis and still find the main coefficient to be negative and significant at 5 percent level in two-tailed tests. Finally, I employ the *DFBETA* sensitivity check using multiple cutoff values for all results presented in this paper and find similar results. This test rules out the concern that few influential observations drive my results.

To provide external validity to my results, I additionally examine narrative disclosures from 10-Ks and calculate *ABTONE* using a similar procedure. I examine the association between short-selling pressure and *ABTONE* over the period 1993-2010. I measure short-selling pressure (*SHORT*) by calculating the annual average of monthly short-interest. I scale this measure by the total number of shares outstanding. I obtain a negative and significant association between *SHORT* and *ABTONE* (Panel A in Table B3 – Appendix B)[19]. I take care of endogeneity issues by employing fixed effect specifications. I control for CEO, firm, and year fixed effects. In another test (Panel B in Table B3 – Appendix B), I find that the negative relationship between *SHORT* and *ABTONE* exists only for overoptimistic managers. These findings are consistent with the disciplinary role of short-sellers.

6. CONCLUSION

This paper provides causal evidence of secondary market determinants of disclosure tone. Prior studies only examine the consequences of disclosure tone such as its impact on

[19] I download 10-Ks from the SEC EDGAR database and collect CEO level information from EXECUCOMP for running this test.

investors' reaction, information environment, and the cost of capital. This is the first study to argue that trading in the secondary market plays a disciplinary role in tone management. I choose short-selling pressure as a setting because short-sellers are sophisticated players in the secondary market. I resolve endogeneity issues by exploiting a regulation-induced exogenous shock to short-selling constraints (Regulation-SHO). Since short-selling pressure only affects firms with overoptimistic or overconfident managers, my results show evidence of the disciplinary role of short-sellers and rules out the concern that managers could become conservative in their tone in the presence of short-sellers. While Rogers, Buskirk, and Zechman (2011) argue that litigation is an ex-post disciplining mechanism for disclosure tone, my findings suggest that ex-ante short-selling pressure could increase the perceived litigation concerns for management.

In addition to tone, I examine the structure of tone and argue that investors are more informed in the presence of short-sellers. Overall, my findings show the benefits of the controversial short-selling activity.

CHAPTER 2

DOES GREATER R&D QUALITATIVE DISCLOSURE PROVIDE INFORMATION ABOUT FIRM PROFITABILITY?

(With Pratik Goel and Sanjay Kallapur)

1. INTRODUCTION

Narrative or qualitative information disclosed by firms in outlets such as the Management Discussion and Analysis (MD&A) section in 10-K and 10-Q filings, earnings press releases, and conference call transcripts, helps them convey information about their economic reality beyond the financial statement numbers. The disclosures' quantity and tone have been found to be associated with contemporaneous and future performance (Li 2010a; Li et al. 2013; Law and Mills 2015).

Firms' narrative disclosures should be even more relevant in the context of their economically important[20] research and development (R&D) activities: information about the future benefit of R&D is highly value-relevant to investors, and asymmetrically vested with managers (Aboody and Lev 2000; Chan et al. 2001; Kothari et al. 2002) who are unable to communicate it through financial statements because all R&D must be expensed. Narrative disclosures are the only tool for managers to communicate this information to investors; one would accordingly assume that the quantity and tone of R&D-related disclosures are informative about the future profitability of R&D.

However, there are other reasons to expect that this might not hold. First, it is possible that although managers possess superior information about the outcome of R&D investments, they might strategically withhold disclosures for competitive reasons. Second, Kahneman and Klein (2009) argue that a high-validity environment (one that has sufficient regularity, which

[20] In 2012, R&D expenditures in the U.S. were estimated at USD 420 billion, contributing approximately USD 1.2 trillion to the U.S. economy (Battelle, *R&D Magazine*, 2012).

provides valid causal cues), and adequate opportunities for learning through rapid and unequivocal feedback, are necessary conditions for the development of skilled judgments. A typical firm's R&D environment, on the other hand, is characterized by low validity and delayed outcome feedback. Such environments could make it difficult for managers to develop skilled intuitive judgments about the future success of R&D investments, reducing the informativeness of their disclosures. In view of these opposing factors, the information content of R&D narrative disclosures about its future profitability is an empirical question; it has not been studied and this paper fills the gap.

Specifically, we examine the association of the quantity and tone of R&D disclosure with future profitability, and empirically test for some of the above alternative explanations. Using a *bag of words* approach and the financial sentiment dictionary of Loughran and McDonald (2011), we capture the quantity and tone of a firm's 10-K narrative R&D-related disclosure by constructing a *weighted* word-count-based quantity measure[21] and a weighted tonal measure, and then examine the association of our measures with future firm profitability (measured by return on assets (ROA)) using 15,579 firm-year observations from publicly-listed US firms during the 1993-2006 period.[22] Besides including R&D intensity, capital intensity, and advertising intensity (Lev and Sougiannis 1996; Ciftci and Cready 2011), we control for innovative efficiency (patents and citations), R&D growth, and current and past ROA following Hirshleifer et al. (2013). To ensure that our results for R&D disclosure and tone are not confounded by overall disclosure and tone, we control for 10-K length (Li 2008) and tone of forward-looking disclosures (Li 2010a). Finally, we include a Big N dummy along with industry (or firm) and year fixed effects in all our empirical specifications.

[21] We also construct and use alternative measures of R&D disclosure quantity, including a sentence-count-based measure (as in Merkley (2014), an unweighted word-count-based measure, and a unique sentence-count-based measure (to allay concerns about the potential repetition of R&D disclosures).

[22] Our sample period ends in 2006 because it is the last year of availability of patents data in the NBER database, needed to compute innovative efficiency—see below. However our results also hold for an alternative sample using Kogan et al.'s (2017) patents data that ends in 2010.

We find a statistically and economically significant negative association between our measure of R&D disclosure quantity and subsequent period ROA (also, *adjusted* ROA – computed by adding back R&D expenditures). Specifically, an increase from the 25th to the 75th percentile in R&D disclosure quantity is associated with 40 per cent decline in ROA for an average firm in our sample. The association between R&D disclosure tone and future ROA, on the other hand, is not significant. Thus, while tone does not seem to matter, a greater R&D disclosure quantity is associated with lower future profitability.[23] Furthermore, we find that R&D disclosure quantity in the current period is also negatively associated with future profitability beyond the subsequent year (that is, ROA up to five years ahead) and thereby conclude that this negative association is persistent.

This evidence of a strong negative association is surprising, since R&D disclosures have been found to be positively correlated with future firm fundamentals (Gu and Li 2003), and the literature suggests that managers adjust them in response to earnings performance for the purpose of providing more relevant information to investors (Gu and Li, 2003; Merkley, 2014). To rule out any measurement error concerns that the association may be driven by the amount of negative or uncertainty-related words contained in the R&D narrative disclosures, we additionally show that the association persists when we include two measures of negative R&D sentiment (*R&D Pessimism* and *R&D Uncertainty*) along with their corresponding interaction terms, where both interactions were found to load insignificantly.[24]

A priori, the negative association is unlikely to be explained by competitive concerns because higher competition should result in lower R&D disclosure as well as lower future profits, a positive relationship. Nevertheless, we examine this explanation empirically using

[23] We perform a host of empirical tests to document the robustness of our results to alternative firm profitability measures (adjusted ROA; cash flow from operating activities (CFO)), inclusion of only the R&D-intensive firms (identified from Hirshleifer et al. (2013)), and choice of sample period (using 1993-2010 instead of 1993-2006).
[24] Using several empirical tests, we also rule out the possibility that this association is statistically mechanical on account of high earnings persistence and the negative association between current ROA and subsequent R&D-related disclosures documented in Merkley (2014).

three different measures of competition, viz., Herfindahl-Hirschman Index (HHI) and two measures from Karuna (2007). We find that our main (negative) association remains highly significant when we include these measures along with their interaction with R&D quantity.

The argument about managers' unskilled and intuitive judgments in the R&D disclosure environment predicts an insignificant association between R&D disclosure quantity and future ROA. Hence, our finding of a significant negative association suggests that it is the worst R&D performers who are also the most biased. This seems consistent with Kruger and Dunning's (1999) argument that the poorest performers hold the least accurate assessments of their skills and thus end up overestimating their performance relative to that of their peers. Empirically we decompose the R&D disclosures (identified at the sentence level) into forward-looking and non-forward-looking disclosures, and obtain a significant negative association with future ROA for only the forward-looking R&D disclosure quantity; this is consistent with difficulty in predicting future profitability.

Our study contributes to the literature in several ways. First, we contribute to the extant literature on R&D. Prior studies have documented that future earnings depend on R&D expenditures (Eberhart et al. (2004, 2008)), innovative efficiency (Hirshleifer et al. 2013), patent citations (Gu 2005; Pandit et al. 2011), and industry-level technological progress (Matolcsy and Wyatt 2008). In this context, our paper shows that narrative R&D disclosure quantity contributes to the information mix albeit through a negative association with future profitability.

Second, we contribute to the literature on R&D disclosures by providing the first large-sample empirical evidence on its association with future earnings. Prior literature on the determinants and consequences of R&D disclosure quantity is limited, comprising mainly of small-sample studies that focus only on specific industries or time periods (such as, Gu and Li (2003), and Jones (2007)). Merkley (2014) is a large-sample empirical study on R&D

narrative disclosures; but he examines the determinants of such disclosures. He also provides evidence on their stock price consequences, but not their future earnings consequences. It is important to provide evidence on the latter because previous studies such as Huang et al. (2014) have found a divergence between the two: in their study an optimistic tone in the earnings press release is associated with positive short-term stock price response, but negative future earnings performance.

Third and finally, we contribute more broadly to the extant literature on narrative disclosures that has so far categorized such disclosures as being either informative (see studies cited in Li (2010b), p. 149-150; and Li et al. 2013) or strategic (Li 2008; Huang et al. 2014) -- an issue that is as yet unsettled (Loughran and McDonald 2016). We argue in this study that depending on the specific features of the disclosure environment a third possibility warrants consideration, namely, that managers may not be well-informed about the future outcomes, which in turn reflects in the quality of their disclosures. In our setting, the high outcome uncertainty of R&D investments together with delayed feedback could make managers' disclosures uninformative, or even biased because the worst performers might have the most optimistic assessments. Thus, we show that the disclosure type and features of its environment matter, and must be considered while drawing conclusions about the information content of narrative disclosures.

2. EMPIRICAL METHODOLOGY

2.1 Measuring the Quantity and Tone of 10-K Narrative R&D Disclosures

We capture the quantity and tone of a firm's R&D-related narrative disclosures by constructing a textual measure from the annual 10-K filing. Textual measures have the advantage of capturing information from many different sources that are hard to identify empirically (Li 2010b). Our choice of 10-Ks is motivated by prior research suggesting that 10-K filings are an important source of information (Previts et al. 1994; Leder 2003; Brown and Tucker 2010; Lehavy et al. 2011; Merkley 2014). Furthermore, the information on R&D activities is largely descriptive in nature; a significant portion of which is contained in the qualitative 10-K disclosures (Entwistle 1999; Jones 2007). We define R&D disclosure quantity as the proportion of R&D related words in the 10-K and R&D tone as the difference between the count of optimistic and pessimistic words in R&D related disclosures scaled by the total number of words in R&D related disclosures in the 10-K.

We employ the bag of words approach to represent the 10-K text numerically. Under this approach, each document is represented by the words it contains, ignoring any punctuation and ordering. Every word is identified and counted the number of times it appears in the document. We also use an algorithm to reduce each word to its 'stem', so that different forms of the same word are considered as a single word (for example, the words "develop", "developed", "developing", and "development" are stemmed to "develop"). Although textual measures based on a simple word-count have been widely used in prior literature,[25] raw word count gives high weightage to common words (Loughran and McDonald 2011). It is critical, therefore, to weight the word counts properly, since the adoption of an appropriate term-weighting scheme can have a significant impact on the effectiveness of the information thus retrieved (Buckley 1993; Jurafsky and Martin 2009; Loughran and McDonald 2011).

We adopt the following term-weighting scheme from prior literature, but modify it to account for the variation over time [similar to Loughran and McDonald (2011)]:

$$\mathrm{w_{i,j,t}} = \begin{cases} \dfrac{(1 + \log(\mathrm{tf_{i,j,t}})}{(1 + \log(\mathrm{a_{j,t}})} * \log(\dfrac{\mathrm{N_t}}{\mathrm{df_{i,t}}}), & \text{if } tf_{i,j} \geq 1 \\ 0, & \text{Otherwise} \end{cases}$$

[25] Li et al. (2013) construct a measure of competition; Bodnaruk et al. (2015) construct a measure of financial constraints based on the frequency of negative words; Audi et al. (2016) construct a measure of trust in a firm's corporate culture.

where $a_{j,t}$ denotes the average word count of documents in year t, $tf_{i,j,t}$ is the raw count of the i^{th} word in the j^{th} document in year t, $df_{i,t}$ represents the number of documents containing at least one occurrence of the i^{th} word in year t, and finally N_t is the total number of 10-K documents in year t. The above weighting scheme offers several benefits. The term frequency (*tf)* helps attenuate the impact of high-frequency words, which is accomplished by the logarithmic transformation[26]. Furthermore, stop words such as *"the"*, *"of"*, *"all"*, *"for"* etc. are suitably assigned a weight of zero, since these words appear frequently in all documents and do not provide any information. The document frequency (*df*) in the weighting scheme accounts for the commonality of words, that is, it assigns lesser weight to words that are commonly used across the 10-K documents[27]. Finally, the weighting scheme also accounts for the average length of 10-K document (*a*) and its variation across our sample period, since the 10-Ks have become significantly lengthier over time and it is more likely for a word appearing in 1994 to have a different impact than a word appearing in 2006.

Our textual measure, $R\&D\ QTY_{it}$, is defined as the ratio of the weighted count of R&D-related words in firm *i*'s 10-K document in year *t* to the weighted count of the total number of words in that document.[28] For ease of interpretation, we multiply this measure by a scaling factor of 1,000. We use Python scripts to search for R&D-related words in the entire document on each 10-K filing.[29] To identify R&D-related words, we refer to the dictionary of commonly-used R&D keywords and phrases developed by Merkley (2014) and modify it to

[26] As an example, consider the word *"gain"* which appears 92,286 times across our 10-K sample in the year 2006, and another word *"supportable"* that appears only 326 times. The weighting scheme will assign a lesser weight to *"gain"*, since it is unlikely that the collective impact of *"gain"* is more than 283 (=92286/326) times that of *"supportable"*.

[27] For example, in the year 2006, the word *"risk"* appears in 7,323 documents across our 10-K sample, while the word *"insurgency"* appears in only 13 documents. The second term in the weighting scheme will adjust the first term (increase for *"insurgency"*, and decrease for *"risk"*) appropriately to reflect this feature.

[28] We additionally scale our measure by a constant (1000) for ease of interpretability of its coefficient estimate.

[29] We do not extract R&D-related words from only a specific 10-K section since the amount of R&D-relevant information is spread over various sections throughout the entire 10-K (Entwistle 1999), including MD&A (22.7%), Corporate Overview (50.8%), signed letters section (18.9%), audited financial statements (1%), and a separate R&D section (6.6%).

include only words.[30] Examples of R&D-related words from our (modified) dictionary include "research", "innovate", "breakthrough", "development", "clinical" etc. The dictionary in Merkley (2014) was compiled after consultation with industry personnel on R&D disclosure topics, and even compares to those of James and Shaver (2016) and Entwistle (1999), thus assuring us of its credibility.

We define another textual measure *R&D Tone$_{it}$* as the difference between the weighted count of optimistic words and the weighted count of pessimistic words in R&D related sentences, divided by the weighted count of the total number of words in R&D related sentences. Similar to *R&D QTY$_{it}$*, we use a scaling factor of 1,000 in *R&D Tone$_{it}$* as well. We identify R&D related sentences using the Dictionary from Merkley (2014) and employ the financial dictionary from Loughran and McDonald (2011) to identify optimistic and pessimistic words.

2.2 The Association of Narrative R&D Disclosures with Firm Profitability

To examine the association of our textual measure of R&D disclosure quantity (*R&D QTY$_{it}$*), with subsequent firm profitability, we estimate a (OLS) model similar to Hirshleifer et al. (2013), stated as follows:

$$\mathrm{ROA}_{i,\,t+1} = \alpha_0 + \alpha_1 \ln\,(\mathrm{R\&D\ QTY})_{\mathrm{i,t}} + \alpha_2 \mathrm{R\&D\ Tone}_{\mathrm{i,t}} + \alpha_3 \mathrm{ROA}_{i,t} + \alpha_4 \Delta\mathrm{ROA}_{i,t} + \gamma \mathbf{Z}_t + \varepsilon \quad \text{-- (1)},$$

where $\mathrm{ROA}_{\mathrm{i,t+1}}$ ($\mathrm{ROA}_{i,t}$) is firm *i*'s Return on Assets in year *t+1 (t)*; $\Delta\mathrm{ROA}_{\mathrm{i,t}}$ is the change in ROA between year *t* and year *t-1*; and $\mathbf{Z}_t$ is a vector of controls, including the innovative efficiency (IE) measure(s) from Hirshleifer et al. (2013), R&D expenditure, advertising and capital expenditures, other innovation-related variables (such as, R&D growth and change in

[30] Since our measure is based on a word count, unlike sentence count as in Merkley (2014), we modified this dictionary of R&D phrases to include only words. We validated the modified dictionary through a manual examination of a random selection of 100 10-K filings.

adjusted patent citations), firm size, and auditors' quality.[31] We additionally control for 10-K length and the tone of forward-looking disclosures. We follow Muslu et al. (2015) to identify forward-looking disclosures in the 10-K, and calculate tone of these disclosures using the dictionary of positive and negative words from Loughran and McDonald (2011). All variable definitions and measurement have been outlined in Appendix C. We include two-digit SIC industry dummies[32] in the regression to account for differences in industry characteristics or environments, along with year dummies and cluster all standard errors by both firm and year.

We include current ROA in the above model to capture the persistence in operating performance (Gu 2005; Pandit et al. 2011). Change in ROA serves as a control to account for the mean reversion in profitability (Fama and French 2000). More importantly, we control for the IE measure based on patents,[33] where IE is defined as the ability of the firm to generate patents per dollar of R&D investment (Hirshleifer et al. 2013). Hirshleifer et al. (2013) find a positive relation between this IE measure and future ROA, suggesting that IE measures contain incremental information about subsequent operating performance of the firms. We control for 10-K length and the tone of forward-looking disclosures to capture narrative disclosures other than R&D in the annual report, as prior literature (Li 2008, Li 2010a) has shown that these variables are informative about firms' profitability beyond accounting numbers. We also control for auditors' quality by including a dummy variable Big N that is equal to 1 for firms which are audited by Big N auditors. Finally, we control for advertising and capital expenditures as prior studies have found that they explain operating

[31] As the distributions of our textual measure and some of the controls (including IE) are highly skewed and(or) are often zero, we use a logarithmic transformation of these variables, similar to Hirshleifer et al. (2013) and Lerner (1994).

[32] We do not include firm fixed effects in our main model to avoid inducing a Nickell's bias in the coefficient estimates. The estimates of a model having individual fixed effects and lagged value of the dependent variable as an independent variable (as in our case) are biased when estimated using an OLS (Nickell 1981). As a robustness test, we also include firm fixed effect (See table D1 in Appendix D) instead of industry fixed effects after dropping $ROA_{i,t}$ and Δ $ROA_{i,t}$ and get qualitatively similar results.

[33] We also run all our main regressions using the other IE measure in Hirshleifer et al. (2013) – one based on citations, but do not tabulate those results in the paper.

performance of the firm (Lev and Sougiannis 1996; Pandit et al. 2011; Hirshleifer et al. 2013).

We are interested in the coefficients of *ln (R&D QTY$_{it}$)* and *(R&D Tone$_{it}$)*, which will inform us about the association between the quantity and tone of narrative R&D-related disclosures contained in the firms' 10-K filing and future firm profitability (ROA).

3. DATA

Our sample includes firm-year observations from 1993 to 2006, and consists of firms in the intersection of the Compustat and SEC EDGAR databases, matched using the Central Index Key (CIK). We match remaining unmatched observations using the IRS tax identification number (Nini et al. 2012). We obtain relevant accounting data from Compustat files and the patents and citations data from the National Bureau of Economic Research (NBER) database. We remove firms in the finance, insurance and real estate sectors (SIC codes between 6000 and 6999), and those with negative book value of equity. Figure 2.1 describes our sample selection procedure in detail. Our final sample comprises of 15,579 firm-year observations from 3,069 unique firms.

[INSERT FIGURE 2.1 HERE]

We download all 10-Ks from the SEC EDGAR database. Following Loughran and McDonald (2011), we remove 10-Ks that contain less than 2,000 words, and only include one filing per firm per year by removing the filings that were filed within 180 days from a prior filing. In case there were multiple 10-Ks filed within a year, we consider only the first filing.

Table 2.1 presents the summary statistics for our sample. All variables have been winsorized at the 1 per cent and 99 per cent levels. As shown in Panel A, the average size of a firm in our sample in terms of sales, total assets, and market value is \$ 2,005 million, \$ 2,046 million, and \$ 3,002 million, respectively. Furthermore, a firm expends about \$ 51 million on its R&D activities on average per year. In terms of the textual characteristics, the 10-K filing for an average firm in our sample consists of a total of about 37,583 words, of which nearly 192 are R&D-related.[34]

[INSERT TABLE 2.1 HERE]

4. RESULTS

4.1 Narrative R&D Disclosure Quantity and Future ROA

We want to examine the association of *ln* (*R&D* QTY_{it}) and *R&D* $Tone_{it}$ with future firm profitability (ROA). Panel A of Table 2.2 presents our main results. We find that *ln* (*R&D* QTY_{it}) correlates negatively with subsequent ROA.[35] This association is statistically significant even after controlling for other narrative disclosures in the 10-K, R&D expenditure, the patents-based IE measure in Hirshleifer et al. (2013), and industry and time effects.[36] It is also economically significant – specifically, an increase from the 25th to the 75th percentile in R&D disclosure quantity results in a decrease in ROA of the magnitude of 0.012, which corresponds to a 40 per cent decline in ROA for an average firm in our sample (from Column 5, Table 2.2). We do not, however, obtain any significant association between *R&D* $Tone_{it}$ and ROA.

[INSERT TABLE 2.2 HERE]

---

[34] To validate our measure of R&D disclosure quantity (*ln* (*R&D* QTY_{it})) and ensure that the textual methodology is identifying the R&D-related words accurately, we manually read the10-K filings of the ten firms with the most R&D-related disclosures and the ten with the least amount of such disclosures, as identified by the *ln* (*R&D* QTY_{it}) measure, and noticed significant differences between them with regards to the number of R&D-related words contained in them.

[35] When we decompose ROA into margin and turnover (DuPont decomposition), we find that the negative effect of R&D disclosure quantity holds for both.

[36] The association remains negative and significant when we deflate the relevant control variables by the market value of equity (similar to the model of operating performance in Hirshleifer et al. (2013)) instead of average total assets.

Consistent with the literature, we find that the IE measures[37] [ln(1+*Patents/RDC*)] significantly predict higher ROA. The significantly positive slopes on ROA and the significantly negative slopes on change in ROA confirm both persistence and mean reversion in profitability. Firm size (measured as the natural logarithm of total assets) correlates positively with future ROA. Also, the coefficient on advertising expenditure is positive and that on R&D-growth is insignificant, similar to Hirshleifer et al. (2013). The coefficient on capital expenditure, however, loads insignificantly for our sample.[38]

[INSERT TABLE 2.3 HERE]

Next, we check whether the negative association thus obtained is also persistent. As shown in Table 2.3, the association between *ln* (*R&D QTY*) and future ROA remains significantly negative till five years ahead. In other words, narrative R&D disclosures in a year can predict negative ROA for up to five years ahead[39]. Furthermore, in Table 2.4, we examine the possibility that this association is statistically mechanical on account of high earnings persistence and the negative association between current ROA and subsequent R&D-related disclosures documented in Merkley (2014). To do this, first, we use change in ROA as a dependent variable in column (1) and still obtain a significantly negative association. Second, we partition our sample into gain and loss firms based on current period ROA and separately estimate model specification (1) for the two subsamples. As observed in columns (2) and (3), we again obtain a significant and negative association for both. We can thus rule out the possibility that our main association is statistically mechanical.

[37] In untabulated results, we find that the other IE measure [ln(1+Citations/RD)] also significantly predicts higher ROA and our main association remains significantly negative when it is included as a control (instead of the patents-based IE measure).
[38] Hirshleifer et al. (2013) and Pandit et al. (2011) obtain a positive association between capital expenditures (intensity) and future ROA. For our sample, this association is positive and significant when we use a model specification with firm fixed effects.
[39] As an alternative measure of future profitability, we use average future ROA (cumulated) instead of future ROA and find a significant and negative association between current *ln (R&D QTY)* and future ROA for up to three years ahead. We find the association to be insignificant at time t+4 and t+5.

[INSERT TABLE 2.4 HERE]

An important concern is also that the observed negative association may be driven by the amount of negative words (pessimism) or uncertainty contained in the R&D disclosures. To check for this possible measurement error, we construct two tonal measures from the R&D narrative disclosures[40] (namely, *R&D pessimism* and *R&D uncertainty*[41]), and interact them with *ln* (*R&D* QTY_{it}) in two separate OLS regressions similar to model (1). Specifically, we define R&D pessimism (uncertainty) as the ratio of a weighted count of negative (uncertain) words contained in the narrative R&D-related disclosures to a weighted count of the total words in these disclosures. We employ the financial sentiment dictionary by Loughran and McDonald (2011) to identify the negative and uncertainty-denoting words in the R&D disclosures. This domain-specific dictionary is widely used by researchers to gauge the linguistic tone of text (see, for example, Feldman et al. 2010; Chen et al. 2014; Kearney and Liu 2014 etc.). Table 2.5 presents the results. For both *R&D pessimism* (column 1) and *R&D uncertainty* (column 2), we still obtain a significant negative association between *ln* (*R&D* QTY_{it}) and future ROA and the corresponding interaction terms load insignificantly[42]. Hence, the negative association between *ln* (*R&D* QTY_{it}) and future ROA is not driven by the amount of negative or uncertainty-denoting words contained in R&D disclosures, thereby allaying any concerns with regards to measurement.

[INSERT TABLE 2.5 HERE]

This evidence of a strong negative association is surprising, since R&D disclosures have been found to be positively correlated with future firm fundamentals (Gu and Li 2003),

[40] We construct these measures from the narrative R&D disclosures, as opposed to the entire 10-K. Narrative R&D disclosures comprise of the sentences containing R&D-related phrases.

[41] We compute *R&D Uncertainty*, in addition to *R&D Pessimism*, to capture the uncertainty component of R&D disclosures. Some word examples from uncertainty dictionary include “ambiguous”, “cautious”, “confusion”, “doubt”, “unexpected” etc.

[42] Additionally, in an unreported test, we examine our main association after adding an *R&D optimism* tonal measure and obtain similar results.

and the current consensus is that managers adjust them in response to earnings performance in order to provide more relevant information to investors (Merkley 2014; Gu and Li 2003). Next, we examine some possible explanations in a bid to understand the aforementioned negative association.

4.2 Explaining the Negative Association between R&D Disclosure Quantity and Future ROA

Prior studies have shown that competition affects the firms' voluntary disclosure quantity in their SEC filings (Scott 1994; Harris 1998; Botosan and Stanford 2005), with mixed evidence with regards to the direction of the association. In the case of R&D disclosures, managers' strategic disclosure behavior in response to greater competition could help explain our findings. However, a priori, the negative association is unlikely to be explained by competitive concerns because higher competition should result in lower R&D disclosure as well as lower future profits, a positive relationship. Nevertheless, we examine this explanation empirically using three different measures of competition, viz., Herfindahl-Hirschman Index (HHI) and two measures from Karuna (2007).

The most popular and widely used proxy for competition is industry concentration, measured using the Herfindahl-Hirschman Index (HHI). Higher (lower) values of HHI indicate greater (lesser) industry concentration and thus lesser (greater) competition. (The definition and construction of HHI is outlined in Appendix A.) However, industry concentration (or HHI) as a measure of competition has been criticized by prior studies since the relation between concentration and competition is not clear (especially when market structure is not exogenous), and this measure fails to capture several important dimensions of competition, including product substitutability, market size, and entry costs (Raith 2003; Karuna 2007). In view of this, we also examine the effect of competition by including two measures of competition from Karuna (2007) in two separate regressions (after excluding industry effects) – the price-cost margin (*PC_MARGIN*) which captures product substitutability, and market size (*MKT_SIZE*). Lower price-cost margin (or greater product substitutability), and greater market size reflect greater price competition (Karuna 2007). The definition and measurement of these variables is outlined in Appendix A.

[INSERT TABLE 2.6 HERE]

Table 2.1 presents the summary statistics for the competition measures, while Table 2.6 contains the regression results. Since any measure of competition is likely to be highly correlated with the industry dummies, we ran all regressions in this table after omitting the industry effects. The standard errors were clustered at both firm and year levels. As shown in Table 2.6, all the three interaction terms involving the three measures of competition are statistically insignificant. Therefore, it appears that competitive pressures facing the firm do not lead managers to disclose R&D strategically in our sample. But more importantly, the negative association between *ln (R&D QTY)* and future ROA remains highly significant across all the specifications.

We now extend a psychology-based explanation in our attempt to unravel the negative association. Kahneman and Klein (2009) argue that a high-validity environment (one that has sufficient regularity, which provides valid causal cues), and adequate opportunities for learning through rapid and unequivocal feedback, are necessary conditions for the development of skilled judgments. A typical firm's R&D environment, on the other hand, is characterized by low validity and delayed outcome feedback. Such environments could make it difficult for managers to develop skilled intuitive judgments about the future success of R&D investments, reducing the informativeness of their disclosures. This argument about managers' unskilled and intuitive judgments in the R&D disclosure environment predicts an insignificant association between R&D disclosure quantity and future ROA. Hence, our finding of a significant negative association suggests that it is the worst R&D performers who are also the most biased. This seems consistent with Kruger and Dunning's (1999) argument that the poorest performers hold the least accurate assessments of their skills and thus end up overestimating their performance relative to that of their peers. This line of reasoning could hold even in the context of R&D, where the worst performers may not be capable to accurately judge the future scope and viability of an R&D investment decision.

Empirically, we decompose the R&D disclosures (identified at the sentence level) into forward-looking and non-forward-looking disclosures and then examine their association with future profitability. Following Merkley (2014), we first count the number of R&D-related sentences in each firm's 10-K filing, identified using the R&D dictionary described earlier in the paper. Next, we classify these R&D sentences into FLS and non-FLS using the dictionary of future-oriented phrases and keywords from Muslu et al. (2015). Finally, we compute a measure of forward-looking (non-forward-looking) R&D disclosure quantity by taking a logarithmic transformation of the count of R&D-related FLS (non-FLS) in the 10-K. All variable definitions and measurement are in Appendix C. As shown in Table 2.1, the average firm in our sample discloses 14 R&D-related sentences, of which 2 are FLS and 12 are non-FLS. Table 2.7 contains the regression results. Interestingly, we obtain a significant negative association with future ROA for only the forward-looking R&D disclosure quantity; this seems consistent with difficulty in predicting future profitability

[INSERT TABLE 2.7 HERE]

4.3 Additional Analysis and Robustness Checks

As additional analysis, first, we rerun our model specification (1) by focusing only on firms from the six largest and most R&D-intensive industries (Gu and Li 2003; Hirshleifer et al. 2013), with two-digit SIC of 28 (chemicals, biotech and pharmaceuticals), 35 (computer hardware and machinery), 36 (electrical and electronics), 37 (transportation equipment), 38 (medical and scientific instruments), and 73 (computer software and data services). More than 80 per cent of the total R&D expenditure comes from these six industries, which justifies our choice of the six R&D-intensive industries. Moreover, we adopt an industry-wise term-weighting scheme for R&D intensive industries where the words could be strongly linked to the language of specific industry segments (Loughran and McDonald 2011). The regression results are reported in Panel A of Table 2.8. It can be observed that the negative association between R&D disclosure quantity and future ROA is even stronger for the R&D-intensive firms. In terms of economic significance, a one standard deviation change in R&D disclosure quantity leads to a 53.3 per cent decline in ROA for an average R&D-intensive firm.

[INSERT TABLE 2.8 HERE]

Second, we test the alternative explanation that as narrative R&D disclosures comprise mainly of the manager's projection of future R&D spending, the ROA in the subsequent period declines once the firm actually incurs these expenditures in that period. As shown in Panel B of Table 2.8, when we adjust our ROA measure by adding back R&D expenditure and then rerun model (1) using this alternative measure, the significant negative association between *ln (R&D QTY)* and subsequent-period adjusted ROA still holds, thereby ruling out the above possible explanation.

Moreover, in an unreported test,[43] we additionally control for other variables from prior literature (summarized and used by Merkley (2014)) which could potentially affect a firm's disclosure choices, such as firm age, outside monitoring (captured by analyst coverage and institutional holding), information uncertainty (captured by standard deviation of monthly returns and standard deviation of ROA), leverage, book-to-market ratio, tangible

[43] See Table D3 in Appendix D

assets (PP&E and inventories), and stock issuance. We again obtain a significant and negative association between *ln (R&D QTY)* and future ROA.[44]

[INSERT TABLE 2.9 HERE]

Finally, we run a host of robustness and sensitivity checks. First, in Panel A of Table 2.9, we show that our main results are robust to using cash flow from operating activities (CFO) as an alternative firm profitability measure. Specifically, an increase from the 25th to the 75th percentile in R&D disclosure quantity results in a decrease in CFO (scaled by average total assets) of the magnitude of 0.003, which corresponds to a 3.75 per cent decline in CFO for an average firm in our sample[45]. Second, in Panel B of the same table, we show the robustness of our main results to using alternative measures of R&D disclosure quantity, namely, an unweighted word-count-based measure (Column 1), a sentence-count-based measure used by Merkley (2014) (Column 2), and a measure based on unique sentence count (Column 3). Third, we find that the negative association between R&D disclosure quantity and future profitability exists for not only for firms with low innovation efficiency but for firms with high innovation efficiency as well (Refer to Table D2 in Appendix D). This result rules out a possible explanation that firms with low innovation efficiency requires more explanation regarding their R&D activities. Fourth, to allay concerns with regard to our chosen sample period ending in 2006 due to non-availability of patents data in the NBER database, useful for computing innovative efficiency, in an unreported test[46], we show that our results also hold for an alternative sample using data from Kogan et al. (2017) that ends in 2010.

---

[44] Furthermore, we classify the firms into introductory and growth phase using Dickinson (2011) methodology and still obtain a negative and significant association between R&D disclosure quantity and future ROA for firms which are not in the introductory or growth phase.

[45] Additionally, we find the association of *ln (R&D QTY)* and future innovation efficiency to be insignificant. Thus, current *ln(R&D QTY)* does not predict future innovation efficiency. In another test, we find that firms disclosing more about R&D do not generate higher stock returns up to five years ahead.

[46] See Table D4 in Appendix D

5. CONCLUSION

In this paper, we examine the association between the quantity and tone of R&D-related narrative disclosures in a firm's 10-K filing and future performance (ROA), and offer new insights on the role, importance, and credibility of narrative disclosures and the R&D disclosure process. We chose the R&D setting due to its typical characteristics and the important role of R&D in the creation of future firm value and growth.

The empirical findings of the paper can be easily summarized. We obtain a persistent and significant (statistical and economic) negative association between only R&D disclosure quantity and future profitability. This association is robust to alternative measures, specifications, and explanations examined in the study. We then offer a psychology-based explanation for the negative association. The unique characteristics of the R&D disclosure environment could make it difficult for managers to develop skilled intuitive judgments about the outcome of their firms' R&D investments.

Taken together, the evidence in this paper suggests that a firm's narrative disclosures may not always be meaningful for analyzing current and future firm fundamentals, and the type of a disclosure and features of its environment are important considerations in this regard. Future research on narrative disclosures should take cognizance of this result.

APPENDIX A: VARIABLE DEFINITION

Variable	Definition/Measurement
EARNINGS	Earnings before extra-ordinary items/total assets
RETURNS	Annual stock return over the fiscal year
SIZE	ln (market capitalization)
BTM	Book-to-market ratio
STD_RET	Standard deviation of monthly stock returns over the fiscal year
STD_EARN	Standard deviation of *EARNINGS* over the last five years with at least three non-missing values
AGE	ln (1+ Age) where Age is no of years since a firm appears in CRSP
BUSINESS_SEG	ln (1+number of business segments)
GEO_SEG	ln (1+number of geographical segments)
LOSS	A dummy variable which is equal to 1 when *EARNINGS* < 0 and zero otherwise
Δ EARNINGS	Change in *EARNINGS*
AFE	(Actual EPS - median of most recent consensus analyst forecasts)/stock price at fiscal year ending
AF	Analyst consensus forecast for one-year-ahead EPS/stock price at the fiscal year ending
TONE	(optimistic words - pessimistic words) / (optimistic words + pessimistic words) * 100
ABTONE	Residual from the annual cross-sectional regressions **(1)**
NTONE	Predicted value from the annual cross-sectional regressions **(1)**
ACCRUALS	(Earnings before extra-ordinary items - operating cash flows)/total assets
SHORT	Annual average of monthly short interest scaled by the total no of shares outstanding
DA	Discretionary accruals from Dechow, Sloan, and Sweeney's (1995) modification of Jones's (1991) model

PILOT	A dummy variable for pilot firms during Regulation-SHO and zero otherwise
POST	A dummy variable which is equal to 1 for observations during the Regulation-SHO pilot program (2-May-2005 to 6-Jul-2007) and 0 for observations before the announcement date (28-Jul-2004)
PRE	A dummy variable which is equal to 1 for observations during the one-year period before the Regulation-SHO was announced (29-Jul-2003 to 28-Jul-2004) and 0 otherwise
NASDAQ	A dummy variable which is equal to 1 for firms which are listed on NASDAQ, and zero otherwise
SUE	Change in *EARNINGS* scaled by its standard deviations, calculated over previous 20 quarters data (with at least ten non-missing observations to calculate standard deviations)
Common Words	Average across all words in a particular document of the percent of documents in which each word appears, multiplied by hundred (Loughran and McDonald 2014)
Financial Terminology	Count of unique financial words divided by the total number of unique words in a particular document multiplied by hundred. I use Campbell Harvey's Hypertextual Finance Glossary to count financial words (Loughran and McDonald 2014)
Vocabulary	The proportion of unique words in a particular document from Loughran-McDonald's (2011) master dictionary, multiplied by hundred (Loughran and McDonald 2014)
POSITIVE_ARF	Average reduced frequency (*ARF*) of optimistic words calculated using the methodology from Allee and DeAngelis (2015)
ADJ_POSITIVE_ARF	Adjusted average reduced frequency (*ARF*) of optimistic words calculated using the methodology from Allee and DeAngelis (2015)
FLD ABTONE	*ABTONE* of forward-looking disclosures. Classification into forward-looking and non-forward-looking disclosures is done using the dictionary of phrases from Muslu, Radhakrishnan, Subramanyam, and Lim (2014)
Non-FLD ABTONE	*ABTONE* of non-forward-looking disclosures. Classification into forward-looking and non-forward-looking disclosures is done using the dictionary of phrases from Muslu, Radhakrishnan, Subramanyam, and Lim (2014)
RETAINER	A measure of CEO optimism from Sen and Tumarkin (2015)

OC_FIRM4	A measure of CEO overconfidence calculated using Schrand and Zechman (2012)
OC_FIRM5	A measure of CEO overconfidence calculated using Schrand and Zechman (2012)
Institutional Ownership	Percentage of institutional ownership
ANALYST	No of analysts following a firm during the fiscal year
D_POSITIVE_ARF	Annual decile ranking of *POSITIVE_ARF*
D_SUE	Annual decile ranking of *SUE*
EQUITY_ISSUANCE	Sale of common and preferred stock/total assets
DEBT_ISSUANCE	Long-term debt issues/total assets
Overoptimistic	A dummy variable which is equal to one when $ABTONE \geq 0$ and zero otherwise
Overconfident	A dummy variable which is equal to one for firms with overconfident managers and zero otherwise. I identify overconfident managers using any of these three proxies - *RETAINER*, *OC_FIRM4*, and *OC_FIRM5*.

APPENDIX B: ADDITIONAL RESULTS

TABLE B1: Impact of Regulation-SHO on Short Positions

This table presents the impact of Regulation-SHO on monthly short-selling activity. Panel A presents summary statistics and Panel B presents results from the difference-in-differences analysis. The dependent variable is monthly short-interest scaled by the total no of shares outstanding (*SHORT*). *PILOT* is a dummy variable which is equal to 1 for the pilot firms in Regulation-SHO and zero otherwise. *POST* is dummy variable which is equal to 1 for observations during the Regulation-SHO program and zero otherwise. All specifications include firm and month fixed effects. All variable definitions are outlined in the Appendix A. Standard errors have been clustered at the firm level in column (1), month level in column (2), and at the firm and month level in column (3)-(4). t-statistics reported in parentheses are based on heteroscedasticity-robust standard errors. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Summary Statistics of *SHORT*

	N	**Mean**	**S.D.**	**0.25Q**	**Median**	**0.75Q**
SHORT	89,876	0.05	0.05	0.02	0.03	0.07

Panel B: Regression Results

	(1)	**(2)**	**(3)**	**(4)**
		SHORT		
*PILOT * POST*	0.004***	0.004*	0.004*	0.003*
	[10.423]	[1.936]	[1.939]	[1.796]
SIZE				0.008***
				[4.500]
BTM				-0.001
				[-0.364]
EARNINGS				0.002
				[0.339]
RETURNS				-0.008***
				[-8.152]
STD_RET				0.028***
				[8.816]
NASDAQ				0.030***
				[5.887]
Firm FE	Yes	Yes	Yes	Yes
Months FE	Yes	Yes	Yes	Yes
Observations	89,876	89,876	89,876	85,275
Adjusted R-squared	0.694	0.695	0.694	0.709

TABLE B2: Impact of Short-Selling Pressure on *NTONE*

This table presents the impact of short-selling pressure on *NTONE* from the difference-in-differences analysis. *PILOT* is a dummy variable which is equal to 1 for the pilot firms in Regulation-SHO and zero otherwise. *POST* is dummy variable which is equal to 1 for observations during the Regulation-SHO program and zero otherwise. All variable definitions are outlined in the Appendix A. t-statistics (in brackets) are based on heteroscedasticity-robust standard errors that are clustered at the firm level. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	**(2)**	**(3)**
		NTONE	
POST	-1.732	-2.043	-3.645***
	[-1.001]	[-0.942]	[-2.823]
*PILOT * POST*	0.147	-0.068	0.396
	[0.309]	[-0.143]	[0.831]
PILOT	0.791*	0.884**	
	[1.866]	[2.078]	
Firm FE	No	No	Yes
Industry FE	Yes	No	No
Industry * Year FE	No	Yes	No
Year FE	Yes	No	Yes
Observations	4,647	4,647	4,647
Adjusted R-squared	0.167	0.181	0.464

TABLE B3: Impact of Short-Selling Pressure on 10-K *ABTONE*

Panel A: This table reports the coefficient estimates of a regression of 10-K *ABTONE* on *SHORT* (annual average of monthly short interest scaled by the total no of shares outstanding). Standard errors have been clustered at the firm level in specifications (1) and (6), and at the CEO level in specifications (2)-(5). All variable definitions are outlined in the Appendix A. t-statistics reported in parentheses are based on heteroscedasticity-robust standard errors. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

	ABTONE				***Δ ABTONE***	***ABTONE***
	(1)	(2)	(3)	(4)	(5)	(6)
	Fixed Effect Specification				*Change Specification*	
SHORT	-9.872*	-13.159*	-15.433**	-14.642**		-5.655**
	[-1.818]	[-1.960]	[-2.291]	[-2.008]		[-2.079]
Δ SHORT					-11.185*	
					[-1.773]	
Lagged ABTONE						0.644***
						[69.868]
CEO FE	No	Yes	Yes	Yes	No	No
Firm FE	Yes	No	No	Yes	No	No
Industry FE	No	No	Yes	No	No	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	28,686	14,659	14,659	14,659	11,911	24,168
Adjusted R-squared	0.522	0.530	0.535	0.431	0.001	0.454

Panel B: This table reports the coefficient estimates of a regression of 10-K *ABTONE* on *SHORT* (annual average of monthly short interest scaled by the total no of shares outstanding) on subsamples. *ABTONE* ≥ 0 (Column 1) refers to firms with positive abnormal tone, whereas *ABTONE* < 0 (Column 2) refers to firms with negative abnormal tone. All specification includes firm and year fixed effects. All variable definitions are outlined in the Appendix A. t-statistics (in brackets) are based on heteroscedasticity-robust standard errors that are clustered at the firm level. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

	ABTONE	
	(1)	(2)
	ABTONE ≥ 0	*ABTONE < 0*
SHORT	-11.370***	4.416
	[-3.279]	[0.628]
Firm FE	Yes	Yes
Year FE	Yes	Yes
Observations	15,895	12,791
Adjusted R-squared	0.477	0.402

APPENDIX C: VARIABLE DEFINITION

Variable	Notation	Definition/Measurement
R&D disclosure quantity	ln (R&D QTY)	$\ln(1 + (\frac{\text{Weighted count of R\&D related words in } 10 - \text{K}}{\text{Weighted count of all words in } 10 - \text{K}}) * 1000)$
Sentiment of R&D disclosures	R&D Tone	(Difference between weighted count of optimistic and pessimistic words in R&D related disclosures scaled by weighted count of all words in R&D related disclosures) * 1000
R&D Capital	RDC	$\text{RD}_{it} + 0.8 * \text{RD}_{it-1} + 0.6 * \text{RD}_{it-2} + 0.4 * \text{RD}_{it-3} + 0.2 * \text{RD}_{it-4}$
Innovative Efficiency	ln (1+ Patents/RDC)	Patents granted in year t / RDC(t-2)
Length of the 10-K document	10K Length	ln (total words in 10-K document)
Tone of the forward looking disclosures	FLS Tone	Weighted count of optimistic words in 10-K forward-looking disclosures subtracted by the weighted count of pessimistic words in it, divided by weighted count of all words in 10-K forward-looking disclosures
Advertising Expenditures	ln (1+AD/Asset)	ln (1+Advertising Expenditures/Average Total Assets)
Capital Expenditures	ln (1+Capex/Asset)	ln (1+Capital Expenditures/Average Total Assets)
Size	ln (Asset)	ln (Asset)
Operating Performance	ROA	Return on Asset (ROA): Income before extra-ordinary items plus interest expenses divided by average total assets
R&D Intensity	ln (1+ R&D Exp/Asset)	ln (1+R&D Expenditures/Average Total Assets)

R&D Growth	R&D Growth Dummy	For firms that have (as of the beginning of their R&D increase year) an R&D intensity (i.e., the ratios of R&D to assets and R&D to sales) of at least 5 percent, it is equal to 1 when firm increase its dollar R&D by at least 5 percent, and increase its ratio of R&D to assets by at least 5 percent (e.g., from 10 percent to 10.5 percent). Otherwise, it is equal to 0.
Adjusted Patent Citation	APC	Citations in year t scaled by total assets averaged over years t-1 and t-2
R&D Optimism	R&D Optimism	(Weighted count of optimistic words in R&D related disclosure scaled by weighted count of all words in R&D related disclosure) * 1000.
R&D Pessimism	R&D Pessimism	(Weighted count of pessimistic words in R&D related disclosure scaled by weighted count of all words in R&D related disclosure) * 1000.
R&D Uncertainty	R&D Uncertainty	(Weighted count of words involving uncertainty in R&D related disclosure scaled by weighted count of all words in R&D related disclosure) * 1000.
Herfindahl-Hirschman Index	HHI	Sum of squared market shares, where market share of an individual firm is calculated by using firm's net sales divided by the total sales value of the whole industry
Price-cost margin	PC_MARGIN	Sales/operating costs, for each industrial segment; where operating costs include cost of goods sold, selling, general, and administrative expense, and depreciation, depletion, and amortization.
Market size	MKT_SIZE	Natural log of industry sales
R&D related forward looking disclosure	ln (FLS)	ln (1+ R&D-related forward-looking sentences)
R&D related non-forward looking disclosure	ln (Non-FLS)	ln (1+ R&D-related non-forward-looking sentences)
Auditor Quality	Big N	A dummy variable which is equal to 1 for firms which were audited by Big N, 0 otherwise

APPENDIX D: ADDITIONAL RESULTS

TABLE D1: Main Specification with Firm Fixed Effects

This table reports the coefficient estimates of a regression of subsequent-period ROA on current R&D disclosure quantity using firm fixed effect specification. All variable definitions are outlined in Appendix C. t-statistics (in brackets) are based on standard errors that are clustered at the firm and year level. ***, **, * represents statistical significance at the 1%, 5% and 10% levels.

VARIABLES	**ROA (t+1)**
ln (R&D QTY)	-0.009**
	[-2.730]
R&D Tone	0.000
	[0.993]
ln (1+ Patents/RDC)	0.004
	[0.295]
ln (1+ R&D Exp/Asset)	-0.063
	[-0.735]
R&D Growth Dummy	0.004
	[0.961]
Δ APC	-0.073
	[-0.404]
ln (1+AD/Asset)	0.066
	[1.288]
ln(1+Capex/Asset)	0.112**
	[2.413]
ln (Asset)	-0.030***
	[-4.241]
FLS Tone	-0.004
	[-0.190]
10K Length	-0.011***
	[-4.672]
Big N	-0.005
	[-0.975]
Fixed Effects:	
Year	Yes
Firm	Yes
Observations	15,579
Adjusted R-squared	0.487

TABLE D2: Sub-Sample Analysis based on Innovation Efficiency

This table reports the results from sub-sample analysis based on innovation efficiency. Column (1) is based on a sample of firms with low innovation efficiency and column (2) is based on firms with high innovation efficiency. All variable definitions are outlined in Appendix C. t-statistics (in brackets) are based on standard errors that are clustered at the firm and year level. ***, **, * represents statistical significance at the 1%, 5% and 10% levels.

VARIABLES	**ROA (t+1)**	
	(1)	**(2)**
	Low IE	**High IE**
ln (R&D QTY)	-0.008***	-0.009***
	[-4.885]	[-4.204]
R&D Tone	-0.000	-0.000
	[-0.763]	[-0.119]
ROA	0.615***	0.607***
	[25.917]	[28.248]
Δ ROA	-0.097***	-0.125***
	[-5.502]	[-3.113]
ln (1+ R&D Exp/Asset)	-0.017	-0.016
	[-0.520]	[-0.210]
R&D Growth Dummy	-0.000	0.005
	[-0.040]	[1.031]
Δ APC	0.014	0.062
	[0.023]	[0.462]
ln (1+AD/Asset)	0.030	0.060
	[1.501]	[1.511]
ln(1+Capex/Asset)	-0.007	-0.005
	[-0.249]	[-0.071]
ln (Asset)	0.005***	0.006***
	[4.363]	[4.643]
FLS Tone	-0.052*	-0.028
	[-1.934]	[-1.305]
10K Length	-0.012***	-0.012***
	[-6.265]	[-3.605]
Big N	0.001	-0.004
	[0.442]	[-1.035]
Fixed Effects:		
Year	Yes	Yes
Industry	Yes	Yes
Observations	11,063	4,516
Adjusted R-squared	0.419	0.439

TABLE D3: Additional Controls

This table reports the coefficient estimates of a regression of subsequent-period ROA on current R&D disclosure quantity after adding additional control variables in the main specification. All variable definitions are outlined in Appendix C. Specification includes year and industry fixed effects. t-statistics (in brackets) are based on standard errors that are clustered at the firm and year level. ***, **, * represents statistical significance at the 1%, 5% and 10% levels.

VARIABLES	ROA (t+1)
ln (R&D QTY)	-0.007**
	[-2.612]
R&D Tone	-0.000
	[-0.641]
ROA	0.589***
	[20.504]
∆ ROA	-0.082**
	[-2.496]
ln (1+ Patents Filed/RDC)	0.017
	[1.656]
ln (1+ R&D Exp/Asset)	-0.005
	[-0.124]
R&D Growth Dummy	0.004
	[0.549]
∆ APC	0.244
	[1.606]
ln (1+AD/Asset)	0.003
	[0.076]
ln(1+Capex/Asset)	-0.095*
	[-1.917]
ln (Asset)	-0.000
	[-0.327]
FLS Tone	-0.049**
	[-2.687]
10K Length	-0.010**
	[-2.897]
Big N	-0.005
	[-1.279]
No Analysts	0.001**
	[2.880]
Inst Own	-0.001
	[-0.175]
Age	0.000
	[0.644]
Age * Age	0.000
	[0.047]
BTM	-0.028***

	[-6.265]
Capital Intensity	0.029***
	[3.255]
Std Dev Returns	-0.014
	[-1.056]
Std Dev ROA	-0.070
	[-1.363]
Leverage	-0.012
	[-0.826]
Net Stock Issued	-0.009**
	[-2.842]
Fixed Effects:	
Year	Yes
Industry	Yes
Observations	5,346
Adjusted R-squared	0.439

TABLE D4: Examining the Association between R&D Disclosure Quantity and ROA using the Extended Patent Database

This table report the coefficient estimates of a regression of subsequent-period ROA on current R&D disclosure quantity using Kogan et al. (2017)'s extended patent database (available until 2010). All variable definitions are outlined in Appendix C. Each specification includes year and industry fixed effects. t-statistics (in brackets) are based on standard errors that are clustered at the firm and year level. ***, **, * represents statistical significance at the 1%, 5% and 10% levels.

VARIABLES	ROA (t+1)
ln (R&D QTY)	-0.008***
	[-5.414]
R&D Tone	0.000
	[0.015]
ROA	0.608***
	[25.602]
Δ ROA	-0.120***
	[-6.390]
ln (1+ Patents/RDC)	-0.010
	[-1.079]
ln (1+ R&D Exp/Asset)	-0.034
	[-0.897]
R&D Growth Dummy	0.003
	[0.761]
Δ APC	0.156
	[0.823]
ln (1+AD/Asset)	0.035
	[1.639]
ln(1+Capex/Asset)	-0.005
	[-0.169]
ln (Asset)	0.006***
	[8.397]
FLS Tone	-0.042**
	[-2.524]
10K Length	-0.012***
	[-7.845]
Big N	0.004
	[1.666]
Fixed Effects:	
Year	Yes
Industry	Yes
Observations	20,257
Adjusted R-squared	0.421

REFERENCES

Aboody, D. and B. Lev, 2000. Information Asymmetry, R&D, and Insider Gains. The Journal of Finance 55(6): 2747-2766.

Alexander, Gordon J., and Mark A. Peterson. "Short selling on the New York Stock Exchange and the effects of the uptick rule." Journal of Financial Intermediation 8, no. 1-2 (1999): 90-116.

Allee, Kristian D., and Matthew D. DeAngelis. "The structure of voluntary disclosure narratives: Evidence from tone dispersion." Journal of Accounting Research 53, no. 2 (2015): 241-274.

Angel, James J. "Short selling on the NYSE." Unpublished manuscript, Georgetown University (1997).

Arslan-Ayaydin, Özgür, Kris Boudt, and James Thewissen. "Managers set the tone: Equity incentives and the tone of earnings press releases." Journal of Banking & Finance (2015).

Audi, Robert, Tim Loughran, and Bill McDonald. "Trust, but Verify: MD&A Language and the Role of Trust in Corporate Culture." Journal of Business Ethics 139, no. 3 (2016): 551-561.

Beschwitz, von Bastian, Oleg Chuprinin, and Massimo Massa. "Why Do Short Sellers Like Qualitative News?." Journal of Financial and Quantitative Analysis 52, no. 2 (2017): 645-675.

Beyer, Anne, Daniel A. Cohen, Thomas Z. Lys, and Beverly R. Walther. "The financial reporting environment: Review of the recent literature." Journal of Accounting and Economics 50, no. 2 (2010): 296-343.

Blau, Benjamin M., Jared R. DeLisle, and S. McKay Price. "Do sophisticated investors interpret earnings conference call tone differently than investors at large? Evidence from short sales." Journal of Corporate Finance 31 (2015): 203-219.

Bodnaruk, Andriy, Tim Loughran, and Bill McDonald. "Using 10-k text to gauge financial constraints." Journal of Financial and Quantitative Analysis 50, no. 4 (2015): 623-646.

Boehmer, Ekkehart, and Juan Julie Wu. "Short selling and the price discovery process." Review of Financial Studies 26, no. 2 (2013): 287-322.

Boehmer, Ekkehart, Charles M. Jones, and Xiaoyan Zhang. "Which shorts are informed?." The Journal of Finance 63, no. 2 (2008): 491-527.

Botosan, C. A., & Stanford, M. 2005. Managers' motives to withhold segment disclosures and the effect of SFAS No. 131 on analysts' information environment. The Accounting Review, 80(3), 751-772.

Bozanic, Zahn, Darren T. Roulstone, and Andrew Van Buskirk. "Management earnings forecasts and other forward-looking statements." Journal of Accounting and Economics (2017).

Bris, Arturo, William N. Goetzmann, and Ning Zhu. "Efficiency and the bear: Short sales and markets around the world." The Journal of Finance 62, no. 3 (2007): 1029-1079.

Brown, Lawrence D., Andrew C. Call, Michael B. Clement, and Nathan Y. Sharp. "Managing the Narrative: Investor Relations Officers and Corporate Disclosure." (2017).

Brown, S., and J. Tucker. 2010. Large-sample evidence on firms' year-over-year MD&A modifications. Journal of Accounting Research 49 (2): 309–346.

Buckley, C. (1993, March). The importance of proper weighting methods. In Proceedings of the workshop on Human Language Technology (pp. 349-352). Association for Computational Linguistics.

Cazier, Richard A., Kenneth J. Merkley, and John Spencer Treu. "When are Firms Sued for Making Optimistic Disclosures?" (2016).

Chan, L. K. C., Lakonishok, J. and T. Sougiannis, 2001. The Stock Market Valuation of Research and Development Expenditures. The Journal of Finance 56(6): 2431-2456.

Chen, H., De, P., Hu, Y. J., & Hwang, B. H. 2014. Wisdom of crowds: The value of stock opinions transmitted through social media. Review of Financial Studies, 27(5), 1367-1403.

Chen, Tao, Jarrad Harford, and Chen Lin. "Do analysts matter for governance? Evidence from natural experiments." Journal of Financial Economics 115, no. 2 (2015): 383-410.

Christophe, Stephen E., Michael G. Ferri, and James J. Angel. "Short-selling prior to earnings announcements." The Journal of Finance 59, no. 4 (2004): 1845-1876.

Christophe, Stephen E., Michael G. Ferri, and Jim Hsieh. "Informed trading before analyst downgrades: Evidence from short sellers." Journal of Financial Economics 95, no. 1 (2010): 85-106.

Ciftci, Mustafa, and William M. Cready. "Scale effects of R&D as reflected in earnings and returns." Journal of Accounting and Economics 52, no. 1 (2011): 62-80.

Clinch, Greg, Wei Li, and Yunyan Zhang. "Short selling and firms' disclosure of bad news: Evidence from Regulation SHO." (2016).

Cohen, Lauren, Dong Lou, and Christopher Malloy. "Playing favorites: How firms prevent the revelation of bad news. " No. w19429. National Bureau of Economic Research, 2013.

Cohen, Lauren, Karl B. Diether, and Christopher J. Malloy. "Supply and demand shifts in the shorting market." The Journal of Finance 62, no. 5 (2007): 2061-2096.

Curtis, Asher, and Neil L. Fargher. "Does short selling amplify price declines or align stocks with their fundamental values?." Management Science 60, no. 9 (2014): 2324-2340.

D'Avolio, Gene. "The market for borrowing stock." Journal of Financial Economics 66, no. 2 (2002): 271-306.

Davis, Angela K., and Isho Tama-Sweet. "Managers' use of language across alternative disclosure outlets: Earnings press releases versus MD&A." Contemporary Accounting Research 29, no. 3 (2012): 804-837.

Davis, Angela K., Jeremy M. Piger, and Lisa M. Sedor. "Beyond the numbers: Measuring the information content of earnings press release language." Contemporary Accounting Research 29, no. 3 (2012): 845-868.

Dechow, Patricia M., Richard G. Sloan, and Amy P. Sweeney. "Detecting earnings management." The Accounting review (1995): 193-225.

Desai, Hemang, Srinivasan Krishnamurthy, and Kumar Venkataraman. "Do short sellers target firms with poor earnings quality? Evidence from earnings restatements." Review of Accounting Studies 11, no. 1 (2006): 71-90.

Diamond, Douglas W., and Robert E. Verrecchia. "Constraints on short-selling and asset price adjustment to private information." Journal of Financial Economics 18, no. 2 (1987): 277-311.

Dickinson, Victoria. "Cash flow patterns as a proxy for firm life cycle." The Accounting Review 86, no. 6 (2011): 1969-1994.

Diether, Karl B., Kuan-Hui Lee, and Ingrid M. Werner. "Short-sale strategies and return predictability." Review of Financial Studies 22, no. 2 (2009): 575-607.

Duffie, Darrell, Nicolae Garleanu, and Lasse Heje Pedersen. "Securities lending, shorting, and pricing." Journal of Financial Economics 66, no. 2 (2002): 307-339.

Durnev, Art, and Claudine Mangen. "The real effects of disclosure tone: Evidence from restatements." (2011).

Eberhart, A., Maxwell, W. and Siddique, A., 2008. A Reexamination of the Tradeoff between the Future Benefit and Riskiness of R&D Increases. Journal of Accounting Research, 46(1), pp.27-52.

Eberhart, A.C., Maxwell, W.F. and Siddique, A.R., 2004. An examination of long-term abnormal stock returns and operating performance following R&D increases. The Journal of Finance, 59(2), pp.623-650.

Engelberg, Joseph E., Adam V. Reed, and Matthew C. Ringgenberg. "How are shorts informed?: Short sellers, news, and information processing." Journal of Financial Economics 105, no. 2 (2012): 260-278.

Entwistle, G. 1999. Exploring the R&D disclosure environment. Accounting Horizons 13 (4): 323–341.

Fama, E. F., & French, K. R. 2000. Forecasting Profitability and Earnings.The Journal of Business, 73(2), 161-175.

Fang, Vivian W., Allen H. Huang, and Jonathan M. Karpoff. "Short selling and earnings management: A controlled experiment." The Journal of Finance (2016).

Feldman, R., Govindaraj, S., Livnat, J., & Segal, B. 2010. Management's tone change, post earnings announcement drift and accruals. Review of Accounting Studies, 15(4), 915-953.

Feldman, Ronen, Suresh Govindaraj, Joshua Livnat, and Benjamin Segal. "Management's tone change, post earnings announcement drift and accruals." Review of Accounting Studies 15, no. 4 (2010): 915-953.

Frankel, Richard, William J. Mayew, and Yan Sun. "Do pennies matter? Investor relations consequences of small negative earnings surprises." Review of Accounting Studies 15, no. 1 (2010): 220-242.

Graham, John R., Campbell R. Harvey, and Shiva Rajgopal. "The economic implications of corporate financial reporting." Journal of Accounting and Economics 40, no. 1 (2005): 3-73.

Grullon, Gustavo, Sébastien Michenaud, and James P. Weston. "The real effects of short-selling constraints." The Review of Financial Studies 28, no. 6 (2015): 1737-1767.

Gu, F., & Li, J. Q. 2003. Disclosure of innovation activities by high-technology firms. Asia-Pacific Journal of Accounting & Economics, 10(2), 143-172.

Gu, F., 2005. Innovation, future earnings, and market efficiency. Journal of Accounting Auditing and Finance 20, 385–418.

Hall, B. H., Jaffe, A. B., & Trajtenberg, M. 2001. The NBER patent citation data file: Lessons, insights and methodological tools (No. w8498). National Bureau of Economic Research.

Harris, Mary Stanford. 1998. "The association between competition and managers' business segment reporting decisions." Journal of Accounting Research: 111-128.

Healy, Paul M., and Krishna G. Palepu. "Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical disclosure literature." Journal of Accounting and Economics 31, no. 1 (2001): 405-440.

Henry, Elaine, and Andrew J. Leone. "Measuring qualitative information in capital markets research: Comparison of alternative methodologies to measure disclosure tone." The Accounting Review 91, no. 1 (2015): 153-178.

Henry, Elaine. "Are investors influenced by how earnings press releases are written?." The Journal of Business Communication (1973) 45, no. 4 (2008): 363-407.

Hirshleifer, D., Hsu, P. H., & Li, D. 2013. Innovative efficiency and stock returns. Journal of Financial Economics, 107(3), 632-654.

Hobson, Jessen L., William J. Mayew, Mark Peecher, and Mohan Venkatachalam. "Improving experienced auditors’ detection of deception in CEO narratives." Journal of Accounting Research (2017).

Hong, Harrison, Jeffrey D. Kubik, and Tal Fishman. "Do arbitrageurs amplify economic shocks?." Journal of Financial Economics 103, no. 3 (2012): 454-470.

Hong, Harrison, Jose Scheinkman, and Wei Xiong. "Asset float and speculative bubbles." The Journal of Finance 61, no. 3 (2006): 1073-1117.

Hope, Ole-Kristian, Danqi Hu, and Wuyang Zhao. "Third-party consequences of short-selling threats: The case of auditor behavior." Journal of Accounting and Economics 63, no. 2 (2017): 479-498.

Huang, Xuan, Siew Hong Teoh, and Yinglei Zhang. "Tone management." The Accounting Review 89, no. 3 (2014): 1083-1113.

James, Sharon D., and J. Myles Shaver. "Motivations for voluntary public R&D disclosures." Academy of Management Discoveries 2, no. 3 (2016): 290-312.

Jegadeesh, Narasimhan, and Di Wu. "Word power: A new approach for content analysis." Journal of Financial Economics 110, no. 3 (2013): 712-729.

Jones, D. A. 2007. Voluntary disclosure in R&D-intensive industries. Contemporary Accounting Research 24 (2): 489–522.

Jones, Jennifer J. "Earnings management during import relief investigations." Journal of Accounting Research (1991): 193-228.

Jurafsky, Daniel and James H. Martin, 2009, Speech and Language Processing, Prentice Hall, New Jersey.

Kahneman, Daniel, and Gary Klein. "Conditions for intuitive expertise: a failure to disagree." American Psychologist 64.6 (2009): 515.

Karpoff, Jonathan M., and Xiaoxia Lou. "Short sellers and financial misconduct." The Journal of Finance 65, no. 5 (2010): 1879-1913.

Karuna, C. 2007. Industry product market competition and managerial incentives. Journal of Accounting and Economics, 43(2), 275-297.

Kearney, C., & Liu, S. 2014. Textual sentiment in finance: A survey of methods and models. International Review of Financial Analysis, 33, 171-185.

Kecskés, Ambrus, Sattar A. Mansi, and Andrew Jianzhong Zhang. "Are short sellers informed? Evidence from the bond market." The Accounting Review 88, no. 2 (2012): 611-639.

Kelly, Bryan, and Alexander Ljungqvist. "Testing asymmetric-information asset pricing models." The Review of Financial Studies 25, no. 5 (2012): 1366-1413.

Kogan, Leonid, Dimitris Papanikolaou, Amit Seru, and Noah Stoffman. "Technological innovation, resource allocation, and growth." The Quarterly Journal of Economics 132, no. 2 (2017): 665-712.

Kothari, S. P., Laguerre, T. E., and A. J. Leone, 2002. Capitalization versus Expensing: Evidence on the Uncertainty of Future Earnings from Capital Expenditures versus R&D Outlays. Review of Accounting Studies 7(4): 355-382.

Kothari, S. P., Xu Li, and James E. Short. "The effect of disclosures by management, analysts, and business press on cost of capital, return volatility, and analyst forecasts: A study using content analysis." The Accounting Review 84, no. 5 (2009): 1639-1670.

Kruger, J., & Dunning, D. 1999. Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. Journal of Personality and Social Psychology, 77(6), 1121.

Lamont, Owen A. "Go down fighting: Short sellers vs. firms." Review of Asset Pricing Studies 2, no. 1 (2012): 1-30.

Law, K. K. F. and Mills, L. F. 2015. Taxes and Financial Constraints: Evidence from Linguistic Cues. Journal of Accounting Research.

Leder, M. 2003. Financial Fine Print: Uncovering a Company's True Value. John Wiley & Sons.

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

Lerner, J., 1994. The importance of patent scope: an empirical analysis. RAND Journal of Economics 25, 319–333.

Lev, B. and T. Sougiannis, 1996. The capitalization, amortization, and value-relevance of R&D. Journal of Accounting & Economics 21: 107-138.

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

Li, Feng, Russell Lundholm, and Michael Minnis. 2013. A Measure of Competition Based on 10-K Filings. Journal of Accounting Research 51.2: 399-436.

Li, Feng. "The information content of forward-looking statements in corporate filings—A naïve Bayesian machine learning approach." Journal of Accounting Research 48, no. 5 (2010a): 1049-1102.

Li, Feng. "Textual analysis of corporate disclosures: A survey of the literature." Journal of Accounting Literature 29 (2010b): 143.

Li, Yinghua, and Liandong Zhang. "Short selling pressure, stock price behavior, and management forecast precision: Evidence from a natural experiment." Journal of Accounting Research 53, no. 1 (2015): 79-117.

Lo, Kin. "Earnings management and earnings quality." Journal of Accounting and Economics 45, no. 2 (2008): 350-357.

Loughran, Tim, and Bill McDonald. "Measuring readability in financial disclosures." The Journal of Finance 69, no. 4 (2014): 1643-1671.

Loughran, Tim, and Bill McDonald. "Textual analysis in accounting and finance: A survey." Journal of Accounting Research 54, no. 4 (2016): 1187-1230.

Loughran, Tim, and Bill McDonald. "When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks." The Journal of Finance 66, no. 1 (2011): 35-65.

Massa, Massimo, Bohui Zhang, and Hong Zhang. "The Invisible Hand of Short Selling: Does Short Selling Discipline Earnings Management?" Review of Financial Studies 28, no. 6 (2015): 1701-1736.

Matolcsy, Z.P. and Wyatt, A., 2008. The association between technological conditions and the market value of equity. The Accounting Review, 83(2), pp.479-518.

Matsumoto, Dawn, Maarten Pronk, and Erik Roelofsen. "What makes conference calls useful? The information content of managers' presentations and analysts' discussion sessions." The Accounting Review 86, no. 4 (2011): 1383-1414.

Merkley, K. J. 2014. Narrative Disclosure and Earnings Performance: Evidence from R&D Disclosures. The Accounting Review, 89(2), 725-757.

Miller, Edward M. "Risk, uncertainty, and divergence of opinion." The Journal of Finance 32, no. 4 (1977): 1151-1168.

Mitchell, Mark, Todd Pulvino, and Erik Stafford. "Price pressure around mergers." The Journal of Finance 59, no. 1 (2004): 31-63.

Muslu, Volkan, Suresh Radhakrishnan, K. R. Subramanyam, and Dongkuk Lim. "Forward-looking MD&A disclosures and the information environment."Management Science 61, no. 5 (2014): 931-948.

Nickell, S. 1981. Biases in dynamic models with fixed effects. Econometrica: Journal of the Econometric Society, 1417-1426.

Nini, G., Smith, D. C., & Sufi, A. 2012. Creditor control rights, corporate governance, and firm value. Review of Financial Studies, 25(6), 1713-1761.

Opinion Research Corporation, 2008, "Short selling study: The views of corporate issuers", prepared on behalf of NYSE Euronext.

Pandit, S., Wasley, C. E., & Zach, T. 2011. The effect of research and development (R&D) inputs and outputs on the relation between the uncertainty of future operating performance and R&D expenditures. Journal of Accounting, Auditing & Finance, 26, 121-144.

Pownall, Grace, and Paul J. Simko. "The information intermediary role of short sellers." The Accounting Review 80, no. 3 (2005): 941-966.

Previts, G. J., Bricker, R. J., Robinson, T. R., & Young, S. J. 1994. A content analysis of sell-side financial analyst company reports. Accounting Horizons,8(2), 55.

Price, S. McKay, James S. Doran, David R. Peterson, and Barbara A. Bliss. "Earnings conference calls and stock returns: The incremental informativeness of textual tone." Journal of Banking & Finance 36, no. 4 (2012): 992-1011.

Public Company Accounting Oversight Board (PCAOB). Auditing Standard Related to the Auditor’s Assessment of and Response to Risk and Related Amendments to PCAOB Standards. Washington, DC: PCAOB, 2010.

Raith, M., 2003. Competition, risk and managerial incentives. American Economic Review 93 (4), 1425–1436.

Rogers, Jonathan L., Andrew Van Buskirk, and Sarah LC Zechman. "Disclosure tone and shareholder litigation." The Accounting Review 86, no. 6 (2011): 2155-2183.

Saffi, Pedro AC, and Kari Sigurdsson. "Price efficiency and short selling." Review of Financial Studies (2010): hhq124.

Savor, Pavel G., and Mario Gamboa-Cavazos. "Holding on to your shorts: When do short sellers retreat?." (2011).

Schrand, Catherine M., and Sarah LC Zechman. "Executive overconfidence and the slippery slope to financial misreporting." Journal of Accounting and Economics 53, no. 1 (2012): 311-329.

Scott, T. W. 1994. Incentives and disincentives for financial disclosure: Voluntary disclosure of defined benefit pension plan information by Canadian firms. Accounting Review, 26-43.

Securities, U. S., and Exchange Commission. "Economic analysis of the short sale price restrictions under the regulation SHO pilot." Office of Economic Analysis, Washington, DC (2007).

Sen, Rik, and Robert Tumarkin. "Stocking up: Executive optimism, option exercise, and share retention." Journal of Financial Economics 118, no. 2 (2015): 399-430.

Zhou, Dexin. "The blame game." (2014).

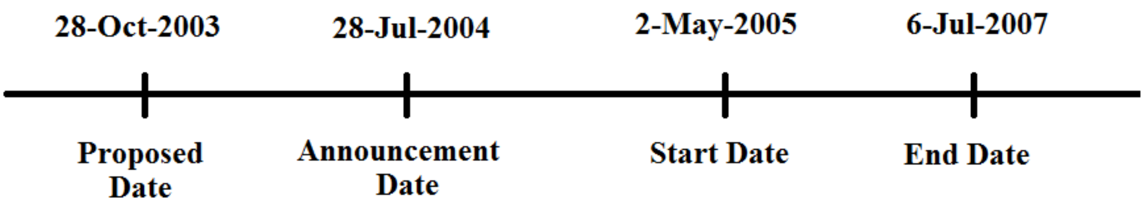

Figure 1.1: Timeline of Regulation-SHO

	Dropped	Observations
Conference call data merged with Compustat (Non-financial firms)		7,571
Exclude observations during the announcement period	1,471	6,100
Check if a firm exists during pre-event as well as post-event period	1,167	4,933
Control variables missing	286	4,647

No of unique treatment firms	386
No of unique control firms	941
No of unique firms	1,327

Figure 1.2: Sample Selection

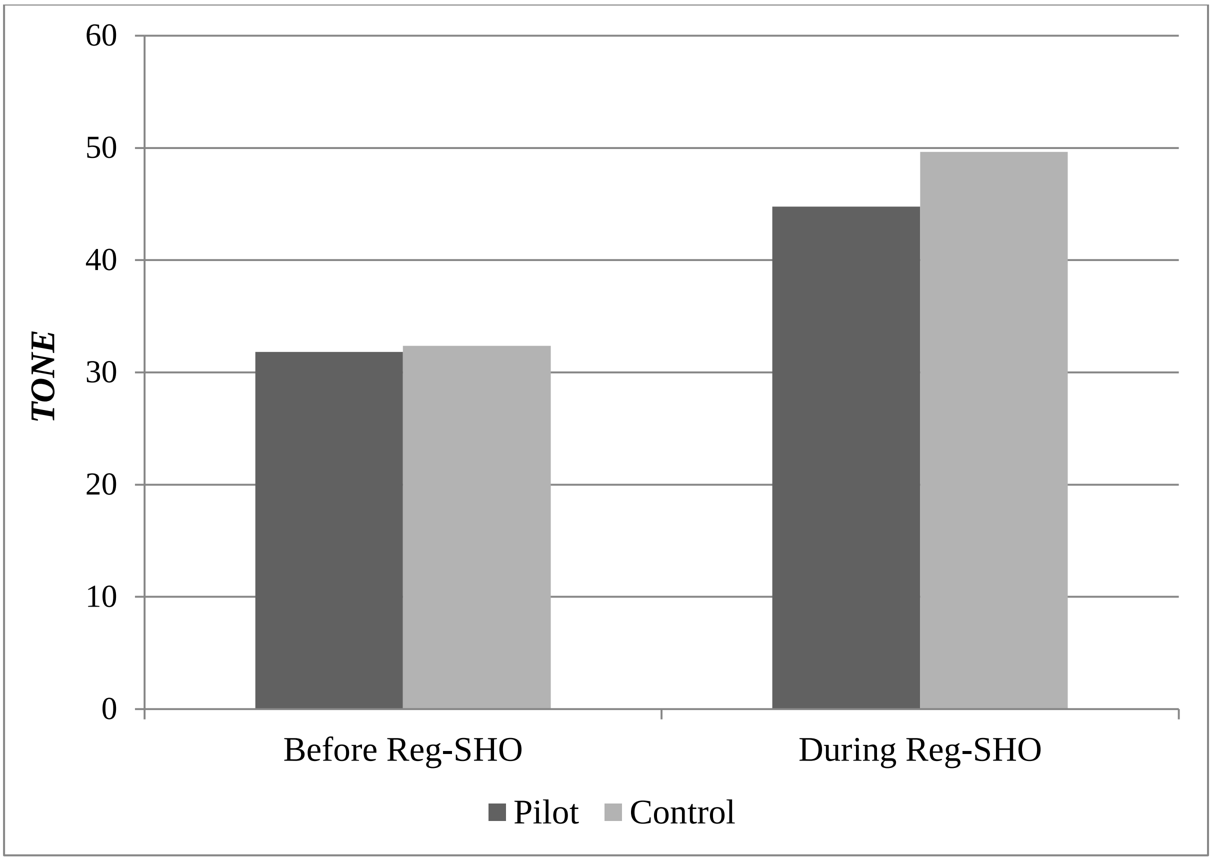

Figure 1.3: Impact of Regulation-SHO on *TONE*

	Dropped	Sample Size
SEC 10-K files 1994 to 2006		87,132
Drop if number of words in 10-K < 2,000	5,667	81,465
Merge with Compustat	43,863	37,602
(Exclude firms with negative value of book equity)		
(Exclude firms with only one year of data in Compustat)		
Drop financial firms (Two digits SIC: 6000-6999)	10,867	26,735
Include only first filing in a given year	230	26,505
At least 180 days between a given firm's 10-K filings	258	26,247
Missing Control Variables	5,799	20,448
Firms with no R&D disclosure	4,869	15,579
Firm -Year Observations		15,579
Number of unique firms		3,069

Figure 2.1: The Selection of 10-K Sample

TABLE 1.1: Calculation of Abnormal Tone

Panel A reports coefficient estimates from the annual cross-sectional regressions of *TONE* on its determinants from the model (1). Panel B shows summary statistics for *ABTONE*. All variable definitions are outlined in the Appendix. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A – Determinants of *TONE*

	TONE		
EARNINGS	-0.006	*GEO_SEG*	-0.005
	[-0.125]		[-0.053]
RETURNS	0.168***	*AGE*	-0.031
	[6.869]		[-1.599]
SIZE	0.055***	*LOSS*	-0.206***
	[6.650]		[-10.929]
BTM	-0.168**	*Δ EARNINGS*	0.145
	[-4.017]		[1.711]
STD_RET	-0.025	*AFE*	1.006***
	[-0.400]		[7.685]
STD_EARN	0.021	*AF*	0.083*
	[1.885]		[2.400]
BUSINESS_SEG	-0.056*	*INTERCEPT*	0.759***
	[-2.073]		[7.712]
Observations	11,193		
Adjusted R-squared	0.081		

Panel B – Summary Statistics

Variable	Mean	S.D.	0.25Q	Median	0.75Q
ABTONE	0.00	26.77	-16.26	2.78	19.09

TABLE 1.2: Summary Statistics

Panel A provides summary statistics of earnings conference call sample. Panel B reports variance decomposition for *TONE* and *ABTONE* and Panel C compares firm characteristics of pilot and control firms. All continuous variables are winsorized at top and bottom 1% to mitigate the effect of outliers. All variable definitions are outlined in the Appendix A.

Panel A - Summary Statistics of Earnings Conference Calls

Variable	N	Mean	S.D.	0.25Q	Median	0.75Q
Length:						
Total Words	4,647	3,040	1,368	2,168	2,857	3,684
Total Sentences	4,647	162	74	116	152	195
ln (Total Words)	4,647	7.93	0.43	7.68	7.96	8.21
ln (Total Sentences)	4,647	5.00	0.41	4.75	5.02	5.27
Tone Variables:						
Optimistic Words	4,647	57	34	34	51	72
Pessimistic Words	4,647	26	18	14	22	34
TONE	4,647	35.88	27.37	18.68	39.24	56.14
ABTONE	4,647	0.46	26.39	-15.47	3.73	19.24
Forward/Non-Forward Looking Disclosures:						
FLD ABTONE	4,599	0.00	26.41	-16.84	2.48	18.72
Non-FLD ABTONE	4,599	-0.11	123.55	-27.77	1.32	28.47
The Structure of Tone:						
POSITIVE_ARF	4,597	0.56	0.05	0.52	0.56	0.59
ADJ_POSITIVE_ARF	4,597	-0.08	0.05	-0.11	-0.08	-0.05
Readability:						
Common Words	4,625	48.85	4.33	45.92	48.40	51.41
Financial Terminology	4,625	10.79	0.86	10.23	10.78	11.34
Vocabulary	4,625	1.52	0.27	1.35	1.54	1.7

Panel B – Variance Decomposition

	Between Variation	Within Variation
TONE	46.68%	53.32%
ABTONE	46.04%	53.96%

Panel C: Summary Statistics of Pilot and Control Firms

This table presents univariate differences in firm characteristics between pilot and control firms during the *PRE* period. The significance of differences in mean (median) between two samples is based on two-tailed t-tests (Wilcoxon rank-sum test). ***, **, and * correspond to 1%, 5%, and 10% significance levels.

	PILOT FIRMS			CONTROL FIRMS				
	MEAN	**MEDIAN**	**SD**	**MEAN**	**MEDIAN**	**SD**	**ΔMEAN**	**ΔMEDIAN**
EARNINGS	0.035	0.051	0.137	0.022	0.044	0.158	0.013	0.006
SIZE	7.290	7.031	1.388	7.051	6.831	1.354	0.238***	0.200**
BTM	0.408	0.381	0.271	0.416	0.362	0.298	-0.008	0.018
STD_RET	0.420	0.351	0.231	0.453	0.383	0.268	-0.033**	-0.032
STD_EARN	0.142	0.042	0.476	0.170	0.050	0.489	-0.027	-0.008**
BUSINESS_SEG	0.243	0.000	0.433	0.218	0.000	0.415	0.025	0.000
GEO_SEG	0.004	0.000	0.053	0.012	0.000	0.090	-0.008	0.000
AGE	2.758	2.708	0.697	2.654	2.565	0.727	0.104**	0.143**
LOSS	0.208	0.000	0.407	0.252	0.000	0.434	-0.043	0.000
Δ EARNINGS	0.021	0.006	0.127	0.025	0.008	0.141	-0.004	-0.002
AFE	0.000	0.001	0.020	0.000	0.000	0.026	0.000	0.000
AF	-0.008	-0.001	0.041	-0.020	-0.001	0.120	0.013*	0.000

TABLE 1.3: Impact of Short-Selling Pressure on Disclosure Tone

This table presents the impact of short-selling pressure on *TONE* (Panel A) and *ABTONE* (Panel B) from the difference-in-differences analysis. *PILOT* is a dummy variable which is equal to 1 for pilot firms in Regulation-SHO and zero otherwise. *POST* is dummy variable which is equal to 1 for observations during the Regulation-SHO program and zero otherwise. *PRE* is dummy variable which is equal to 1 for observations during the one-year period before the Regulation-SHO was announced and zero otherwise. All variable definitions are outlined in the Appendix A. t-statistics (in brackets) are based on heteroscedasticity-robust standard errors that are clustered at the firm level. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Impact of Short-Selling Pressure on *TONE*

Specification (1) includes industry and year fixed effects, specification (2) includes industry*year fixed effects, and specifications (3) and (4) include firm and year fixed effects. Control variables in the specification (4) include all the determinants of *TONE* mentioned in the model (1).

	(1)	**(2)**	**(3)**	**(4)**
		TONE		
POST	1.467	1.624	-2.846	0.207
	[0.238]	[0.194]	[-0.547]	[0.040]
*PILOT * POST*	-3.554**	-4.303**	-3.676**	-3.910**
	[-2.112]	[-2.501]	[-2.181]	[-2.502]
PILOT	1.575	1.838		
	[1.133]	[1.277]		
Controls	No	No	No	Yes
Firm FE	No	No	Yes	Yes
Industry FE	Yes	No	No	No
Industry * Year FE	No	Yes	No	No
Year FE	Yes	No	Yes	Yes
Observations	4,647	4,647	4,647	4,647
Adjusted R-squared	0.087	0.095	0.372	0.427

Panel B: Impact of Short-Selling Pressure on *ABTONE*

Specification (1) includes industry and year fixed effects, specifications (2) and (3) include industry*year fixed effects, and specifications (4) and (5) include firm and year fixed effects.

	(1)	(2)	(3)	(4)	(5)
			ABTONE		
POST	0.826	-0.605	3.983	-1.729	4.357
	[0.108]	[-0.055]	[0.338]	[-0.267]	[0.585]
*PILOT * POST*	-3.591**	-4.154**	-4.321**	-3.934**	-3.827*
	[-2.267]	[-2.548]	[-2.126]	[-2.485]	[-1.890]
PILOT	0.770	0.967	1.133		
	[0.569]	[0.688]	[0.637]		
PRE			4.629		4.886
			[1.153]		[1.298]
*PILOT * PRE*			-0.453		0.063
			[-0.215]		[0.030]
Firm FE	No	No	No	Yes	Yes
Industry FE	Yes	No	No	No	No
Industry * Year FE	No	Yes	Yes	No	No
Year FE	Yes	No	No	Yes	Yes
Observations	4,647	4,647	4,647	4,647	4,647
Adjusted R-squared	0.064	0.065	0.065	0.349	0.349

TABLE 1.4: Impact of Short-Selling Pressure on Overoptimistic and Overconfident Managers

This table presents a sub-sample analysis of the impact of short-selling pressure on *ABTONE* for managers who were either overoptimistic or were overconfident when Regulation-SHO was announced (i.e. during the *PRE* period). Column (1) reports results for firms with overoptimistic managers (*Overoptimism = 1*) and column (5) reports results for firms without overoptimistic managers (*Overoptimism = 0*). Columns (2)-(4) report results for firms with overconfident managers (*Overconfidence = 1*) and columns (6)-(8) report results for firms without overconfident managers (*Overconfidence = 0*). I identify overoptimistic managers using the sign of *ABTONE* and overconfident managers using *RETAINER*, *OC_FIRM4*, and *OC_FIRM5* proxies for overconfidence. *PILOT* is a dummy variable which is equal to 1 for pilot firms in Regulation-SHO and zero otherwise. *POST* is dummy variable which is equal to 1 during the Regulation-SHO program and zero otherwise. All specifications include firm and year fixed effects. All variable definitions are outlined in the Appendix A. t-statistics (in brackets) are based on heteroscedasticity-robust standard errors that are clustered at the firm level. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	ABTONE							
	Overoptimism = 1	***Overconfidence = 1***			***Overoptimism = 0***	***Overconfidence = 0***		
	ABTONE≥0	***RETAINER***	***OC_FIRM4***	***OC_FIRM5***	***ABTONE<0***	***RETAINER***	***OC_FIRM4***	***OC_FIRM5***
POST	-9.419*	-5.757	1.460	6.837	-4.791	3.045	-27.474	-24.454*
	[-1.753]	[-0.789]	[0.279]	[1.142]	[-0.238]	[0.364]	[-0.921]	[-1.674]
*PILOT * POST*	-3.922**	-6.228**	-4.104**	-4.086**	-3.177	-0.411	0.166	-3.498
	[-2.043]	[-2.096]	[-2.380]	[-2.098]	[-1.214]	[-0.092]	[0.031]	[-1.074]
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,447	1,147	3,642	3,040	1,787	609	546	1,065
Adjusted R-squared	0.248	0.390	0.347	0.330	0.278	0.279	0.359	0.377

TABLE 1.5: Impact of Short-selling Pressure on Type of Disclosure Tone

This table presents the impact of short-selling pressure on *FLD ABTONE* (specification 1) and *Non-FLD ABTONE* (specification 2) from the difference-in-differences analysis. *FLD ABTONE* is the abnormal tone from forward-looking disclosures and *Non-FLD ABTONE* is the abnormal tone from non-forward-looking disclosures. Forward looking disclosures are identified using Muslu, Radhakrishnan, Subramanyam, and Lim (2014). *PILOT* is a dummy variable which is equal to 1 for the pilot firms in Regulation-SHO and zero otherwise. *POST* is dummy variable which is equal to 1 for observations during the Regulation-SHO program and zero otherwise. All specifications include firm and year fixed effects. All variable definitions are outlined in the Appendix A. t-statistics (in brackets) are based on heteroscedasticity-robust standard errors that are clustered at the firm level. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1) ***FLD ABTONE***	**(2)** ***Non-FLD ABTONE***
POST	2.234	-11.290
	[0.361]	[-0.547]
*PILOT * POST*	-1.571	-8.483**
	[-0.981]	[-1.985]
Firm FE	Yes	Yes
Year FE	Yes	Yes
Observations	4,546	4,546
Adjusted R-squared	0.138	0.297

TABLE 1.6: Sub-Sample Analysis

This table presents a sub-sample analysis of the impact of short-selling pressure on *ABTONE*. Classification into *High* and *Low* is based on the proxies of short-selling constraints in Panel A and analyst coverage in Panel B, based on their median values during the *PRE* period. *PILOT* is a dummy variable which is equal to 1 for pilot firms in Regulation-SHO and zero otherwise. *POST* is dummy variable which is equal to 1 for observations during the Regulation-SHO program and zero otherwise. All variable definitions are outlined in the Appendix A. t-statistics (in brackets) are based on heteroscedasticity-robust standard errors that are clustered at the firm level. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Sub-Sample Analysis based on Short-Selling Constraints

Classification into *High* and *Low* is based on the level of institutional ownership in columns (1) and (2), and *SIZE* in columns (3) and (4).

	(1)	(2)	(3)	(4)
	ABTONE			
	Institutional Ownership		***SIZE***	
	High	***Low***	***High***	***Low***
POST	-7.119	12.709***	1.356	-13.109
	[-0.583]	[2.702]	[0.224]	[-0.968]
*PILOT * POST*	-3.080	-5.402*	-1.946	-6.297**
	[-1.127]	[-1.880]	[-0.896]	[-2.524]
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	1,494	1,424	2,217	2,017
Adjusted R-squared	0.319	0.360	0.340	0.357

Panel B: Sub-Sample Analysis based on Analyst Coverage

	(1)	(2)
	ABTONE	
	ANALYST	
	High	***Low***
POST	-1.096	-7.696
	[-0.179]	[-0.602]
*PILOT * POST*	-2.517	-5.566**
	[-1.113]	[-2.354]
Firm FE	Yes	Yes
Year FE	Yes	Yes
Observations	2,208	2,026
Adjusted R-squared	0.354	0.338

TABLE 1.7: Impact of Short-selling Pressure on Positive Tone Dispersion

This table presents the impact of short-selling pressure on *POSITIVE_ARF* (Column 1) and *ADJ_POSITIVE_ARF* (Panel B) from the difference-in-differences analysis. *POSITIVE_ARF* and *ADJ_POSITIVE_ARF* capture tone dispersion of optimistic words and are calculated using Allee and DeAngelis (2015). *PILOT* is a dummy variable which is equal to 1 for the pilot firms in Regulation-SHO and zero otherwise. *POST* is dummy variable which is equal to 1 for observations during the Regulation-SHO program and zero otherwise. Control variables in specifications (1) and (2) include all the determinants of *TONE* mentioned in the model (1). All specifications include firm and year fixed effects. All variable definitions are outlined in the Appendix A. t-statistics (in brackets) are based on heteroscedasticity-robust standard errors that are clustered at the firm level. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1) ***POSITIVE_ARF***	**(2)** ***ADJ _POSITIVE_ARF***
POST	-0.009	-0.008
	[-0.816]	[-0.751]
*PILOT * POST*	-0.006*	-0.005*
	[-1.697]	[-1.664]
ln (Total Words)	0.022***	0.020***
	[2.820]	[2.880]
ln (Positive Words)	-0.016**	-0.001
	[-2.486]	[-0.115]
ln (Negative Words)	-0.005**	-0.005**
	[-2.458]	[-2.421]
Controls	Yes	Yes
Firm FE	Yes	Yes
Year FE	Yes	Yes
Observations	4,492	4,492
Adjusted R-squared	0.248	0.262

TABLE 1.8: Investors' Reaction to Positive Tone Dispersion

This table presents investors' reaction to *POSITIVE_ARF* around earnings announcements. *D_POSITIVE_ARF* and *D_SUE* are annual decile ranking for *POSITIVE_ARF* and *SUE* respectively. *PILOT* is a dummy variable which is equal to 1 for the pilot firms in Regulation-SHO and zero otherwise. *POST* is dummy variable which is equal to 1 for observations during the Regulation-SHO program and zero otherwise. Control variables include discretionary accruals (Modified Jones Model), *SIZE*, *BTM*, *RETURNS*, *STD_RET*, and *STD_EARN*. All variable definitions are outlined in the Appendix A. t-statistics (in brackets) are based on heteroscedasticity-robust standard errors that are clustered at the firm level. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	**(2)**
	CAR [-1, +1]	
D_POSITIVE_ARF	0.001*	0.000
	[1.742]	[0.450]
*PILOT * POST * D_POSITIVE_ARF*		-0.003*
		[-1.798]
D_SUE	0.006***	0.006***
	[14.098]	[14.132]
EARNINGS	0.003	0.010
	[0.229]	[0.831]
Controls	Yes	Yes
Industry FE	Yes	Yes
Year FE	Yes	Yes
Observations	4,589	4,498
Adjusted R-squared	0.061	0.063

TABLE 1.9: Robustness Test using Additional Controls

This table presents the impact of short-selling pressure on *ABTONE* from the difference-in-differences analysis. *PILOT* is a dummy variable which is equal to 1 for pilot firms in Regulation-SHO and zero otherwise. *POST* is dummy variable which is equal to 1 for observations during the Regulation-SHO program and zero otherwise. All variable definitions are outlined in the Appendix A. t-statistics (in brackets) are based on heteroscedasticity-robust standard errors that are clustered at the firm level. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	**(2)**	**(3)**
		ABTONE	
POST	7.582	11.945*	2.721
	[1.514]	[1.718]	[0.574]
*PILOT * POST*	-4.861***	-5.534***	-4.166**
	[-2.909]	[-3.166]	[-2.478]
PILOT	1.195	1.413	
	[0.843]	[0.954]	
DA	-8.652**	-9.268**	0.000
	[-2.445]	[-2.347]	[0.000]
EQUITY_ISSUANCE	4.337	5.404	-4.464
	[0.973]	[1.197]	[-0.970]
DEBT_ISSUANCE	1.038	1.409	0.913
	[0.521]	[0.691]	[0.408]
Firm FE	No	No	Yes
Industry FE	Yes	No	No
Industry * Year FE	No	Yes	No
Year FE	Yes	No	Yes
Observations	4,267	4,267	4,267
Adjusted R-squared	0.066	0.061	0.358

TABLE 1.10: Impact of Short-Selling Pressure on Length and Readability

This table presents the impact of short-selling pressure on length and readability of conference calls from the difference-in-differences analysis. *PILOT* is a dummy variable which is equal to 1 for pilot firms in Regulation-SHO and zero otherwise. *POST* is dummy variable which is equal to 1 for observations during the Regulation-SHO program and zero otherwise. Control variables in all specifications include all variables from the model (1). All specifications include firm and year fixed effects. All variable definitions are outlined in the Appendix A. t-statistics (in brackets) are based on heteroscedasticity-robust standard errors that are clustered at the firm level. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)
	ln (Total Words)	***ln (Total Sentences)***	***Common Words***	***Financial Terminology***	***Vocabulary***
POST	-0.027	-0.017	0.113	0.429***	-0.011
	[-0.473]	[-0.314]	[0.127]	[2.693]	[-0.212]
*PILOT * POST*	-0.011	-0.006	-0.049	-0.034	0.005
	[-0.563]	[-0.295]	[-0.249]	[-0.760]	[0.409]
Controls	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Observations	4,647	4,647	4,625	4,625	4,625
Adjusted R-squared	0.649	0.657	0.585	0.458	0.580

TABLE 2.1: Summary Statistics

This table provides summary statistics for the overall sample of 15,579 firm-year observations from 1993-2006. All variable definitions are outlined in Appendix C.

Variable	N	Mean	S.D.	0.25 Q	Median	0.75 Q
Firm Characteristics:						
Sales ($ millions)	15,579	2,005.47	5,529.75	84.77	310.31	1,176.70
Assets ($ millions)	15,579	2,046.07	5,923.53	80.60	288.67	1,093.25
Market Value ($ millions)	15,579	3,002.41	9,596.06	75.36	323.62	1,381.47
ROA	15,579	0.03	0.13	0.01	0.05	0.09
Adj ROA	15,579	0.07	0.12	0.02	0.07	0.13
CFO	15,579	0.08	0.11	0.03	0.09	0.14
R&D Exp ($ millions)	15,579	51.30	180.80	0.00	1.85	17.78
Advertising Exp ($ millions)	15,579	25.62	118.48	0.00	0.00	0.89
Capital Exp ($ millions)	15,579	107.48	346.66	2.60	11.44	52.50
R&D Exp/Total Asset	15,579	0.04	0.07	0.00	0.01	0.06
10-K Characteristics:						
Total Words (10K Length)	15,579	37,583	29,017	19,122	29,289	45,794
Total Sentences (10K)	15,579	1,395	933	803	1,176	1,695
Forward Looking Tone (FLS Tone)	15,579	-0.06	0.05	-0.09	-0.05	-0.03
R&D Disclosures:						
R&D related Words	15,579	192	154	82	147	257
ln (R&D QTY)	15,579	1.74	0.97	0.93	1.71	2.43
R&D Sentences	15,579	14	18	3	8	20
ln (R&D Sentences)	15,579	2.19	1.05	1.39	2.20	3.04
Forward Looking R&D Sentences	15,579	2	4	0	1	3
Non-Forward Looking R&D Sentences	15,579	12	14	2	7	17
R&D Optimism	15,579	20.00	33.84	0.30	8.00	23.77
R&D Pessimism	15,579	24.80	50.93	0.00	2.77	27.86
R&D Uncertainty	15,579	4.90	12.79	0.00	0.26	3.92
R&D Tone	15,579	-4.52	56.60	-12.72	0.00	10.99
Competition Measure:						
HHI	15,211	0.16	0.14	0.06	0.12	0.19
PC_MARGIN	9,608	1.06	0.21	1.01	1.07	1.14
MKT_SIZE	13,822	9.83	2.16	8.83	10.16	11.31

TABLE 2.2: R&D Disclosure Quantity and Future Profitability

This table reports the coefficient estimates of a regression of subsequent-period ROA on current R&D disclosure quantity. All variable definitions are outlined in Appendix C. Specifications (1)-(4) include year and industry fixed effects and specification (5) include industry*year fixed effects. t-statistics (in brackets) are based on standard errors that are clustered at the firm and year level. ***, **, * represents statistical significance at the 1%, 5% and 10% levels.

VARIABLES			**ROA (t+1)**		
	(1)	**(2)**	**(3)**	**(4)**	**(5)**
ln (R&D QTY)	-0.008***	-0.008***	-0.008***	-0.009***	-0.008***
	[-5.295]	[-5.271]	[-5.185]	[-5.531]	[-5.499]
R&D Tone	-0.000	0.000	0.000	0.000	0.000
	[-1.014]	[0.292]	[0.339]	[0.266]	[0.138]
ln (R&D QTY) * R&D Tone		-0.000	-0.000	-0.000	-0.000
		[-0.647]	[-0.693]	[-0.517]	[-0.521]
ROA	0.643***	0.643***	0.637***	0.613***	0.617***
	[29.447]	[29.425]	[30.752]	[31.125]	[32.489]
Δ ROA	-0.116***	-0.116***	-0.113***	-0.103***	-0.111***
	[-6.345]	[-6.345]	[-6.312]	[-5.825]	[-5.972]
ln (1+ Patents/RDC)			0.030***	0.023***	0.020***
			[4.817]	[3.566]	[3.068]
ln (1+ R&D Exp/Asset)			-0.032	-0.011	-0.012
			[-0.813]	[-0.278]	[-0.296]
R&D Growth Dummy			0.003	0.003	0.003
			[0.725]	[0.645]	[0.883]
Δ APC			0.328	0.125	0.132
			[1.592]	[0.802]	[0.947]
ln (1+AD/Asset)				0.035*	0.034*
				[1.865]	[1.803]
ln(1+Capex/Asset)				-0.011	-0.010
				[-0.325]	[-0.326]
ln (Asset)				0.006***	0.005***
				[6.958]	[6.994]
FLS Tone				-0.047**	-0.043**
				[-2.709]	[-2.524]
10K Length				-0.012***	-0.012***
				[-7.331]	[-7.400]
Big N				0.000	0.001
				[0.051]	[0.562]
Fixed Effects:					
Year	Yes	Yes	Yes	Yes	-
Industry	Yes	Yes	Yes	Yes	-
Industry * Year	No	No	No	No	Yes
Observations	15,579	15,579	15,579	15,579	15,579
Adjusted R-squared	0.420	0.420	0.421	0.426	0.430

TABLE 2.3: Association of R&D Disclosure with Future ROA

This table reports the coefficient estimates of a regression of subsequent-period ROA (up to five years ahead) on current R&D disclosure quantity. All variable definitions are outlined in Appendix C. All specifications include year and industry fixed effects. t-statistics (in brackets) are based on standard errors that are clustered at the firm and year level. ***, **, * represents statistical significance at the 1%, 5% and 10% levels.

VARIABLES	**ROA (t+1)**	**ROA (t+2)**	**ROA (t+3)**	**ROA (t+4)**	**ROA (t+5)**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**
ln (R&D QTY) (t)	-0.008***	-0.010***	-0.009***	-0.007**	-0.007*
	[-5.491]	[-5.002]	[-3.387]	[-2.369]	[-2.007]
R&D Tone (t)	-0.000	-0.000	-0.000	-0.000*	-0.000
	[-0.588]	[-0.324]	[-1.557]	[-2.133]	[-1.589]
ROA	0.613***	0.422***	0.315***	0.237***	0.183***
	[31.149]	[12.354]	[9.314]	[7.278]	[5.282]
Δ ROA	-0.103***	-0.094***	-0.074***	-0.085***	-0.038*
	[-5.825]	[-4.452]	[-4.623]	[-3.555]	[-2.138]
ln (1+ Patents/RDC)	0.023***	0.040***	0.040***	0.049***	0.051***
	[3.589]	[5.189]	[3.735]	[3.828]	[4.177]
ln (1+ R&D Exp/Asset)	-0.011	-0.063	-0.089	-0.093	-0.157**
	[-0.279]	[-1.333]	[-1.405]	[-1.258]	[-2.399]
R&D Growth Dummy	0.003	0.003	-0.008	-0.002	0.005
	[0.649]	[0.559]	[-1.281]	[-0.317]	[0.773]
Δ APC	0.125	0.281	0.501**	0.201	0.162
	[0.801]	[1.503]	[2.283]	[0.924]	[0.479]
ln (1+AD/Asset)	0.035*	0.046	0.036	0.044	0.039
	[1.877]	[1.482]	[0.884]	[1.330]	[0.929]
ln(1+Capex/Asset)	-0.011	-0.013	-0.049	-0.038	-0.033
	[-0.328]	[-0.359]	[-1.636]	[-1.115]	[-0.747]
ln (Asset)	0.006***	0.007***	0.008***	0.009***	0.010***
	[6.964]	[7.501]	[8.066]	[7.409]	[5.977]
FLS Tone	-0.047**	-0.044*	-0.048	-0.017	-0.004
	[-2.724]	[-1.970]	[-1.718]	[-0.572]	[-0.089]
10K Length	-0.012***	-0.014***	-0.014***	-0.015***	-0.017***
	[-7.354]	[-5.491]	[-4.665]	[-4.889]	[-4.214]
Big N	0.000	-0.007*	-0.008*	-0.014**	-0.008
	[0.038]	[-2.012]	[-1.884]	[-2.509]	[-1.790]
Fixed Effects:					
Year	Yes	Yes	Yes	Yes	Yes
Industry	Yes	Yes	Yes	Yes	Yes
Observations	15,579	11,097	9,009	7,215	5,782
Adjusted R-squared	0.426	0.262	0.196	0.156	0.141

TABLE 2.4: The Implications of Earnings Persistence for the Association between R&D Disclosure Quantity and ROA

This table reports the coefficient estimates of a regression of change in ROA (column 1) and ROA (columns 2 and 3) in subsequent period on current R&D disclosure quantity. All variable definitions are outlined in Appendix C. Each specification includes year and industry fixed effects. t-statistics (in brackets) are based on standard errors that are clustered at the firm and year level. ***, **, * represents statistical significance at the 1%, 5% and 10% levels.

VARIABLES	Δ ROA (t+1)	ROA (t+1)	
	(1)	(2)	(3)
		Gain	Loss
ln (R&D QTY)	-0.006***	-0.022***	-0.004***
	[-3.117]	[-4.905]	[-4.084]
R&D Tone	-0.000	-0.000	0.000
	[-0.443]	[-0.662]	[0.614]
ROA		0.476***	0.761***
		[15.425]	[34.721]
Δ ROA		-0.113***	-0.057**
		[-6.168]	[-2.526]
ln (1+ Patents/RDC)	0.011	0.027	0.022***
	[1.229]	[1.323]	[5.117]
ln (1+ R&D Exp/Asset)	0.232***	0.022	-0.130**
	[5.637]	[0.265]	[-2.587]
R&D Growth Dummy	-0.011*	0.011	0.005
	[-2.010]	[0.975]	[1.080]
Δ APC	-0.069	0.384	-0.031
	[-0.490]	[0.535]	[-0.195]
ln (1+AD/Asset)	0.013	0.045	0.017
	[0.631]	[0.848]	[0.754]
ln(1+Capex/Asset)	-0.109**	-0.050	-0.037
	[-2.217]	[-0.768]	[-1.316]
ln (Asset)	-0.002	0.010***	0.005***
	[-1.664]	[3.016]	[7.198]
FLS Tone	-0.110***	-0.256***	0.003
	[-8.818]	[-4.561]	[0.190]
10K Length	-0.001	-0.019***	-0.009***
	[-0.383]	[-3.965]	[-6.675]
Big N	-0.002	-0.006	0.002
	[-0.508]	[-0.755]	[0.741]
Fixed Effects:			
Year	Yes	Yes	Yes
Industry	Yes	Yes	Yes
Observations	11,097	3,599	11,980
Adjusted R-squared	0.031	0.277	0.280

TABLE 2.5: Measurement Error Tests for the Association between R&D Disclosure Quantity and ROA

This table reports the coefficient estimates of a regression of subsequent-period ROA on current R&D disclosure quantity after including two measures of negative R&D sentiment [*R&D Pessimism* (column 1) and *R&D Uncertainty* (column 2)]. All variable definitions are outlined in Appendix C. Each specification includes year and industry fixed effects. t-statistics (in brackets) are based on standard errors that are clustered at the industry and year level. ***, **, * represents statistical significance at the 1%, 5% and 10% levels.

VARIABLES	**ROA (t+1)**	
	(1)	**(2)**
ln (R&D QTY)	-0.008***	-0.008***
	[-5.492]	[-5.506]
R&D Pessimism	-0.000	
	[-0.286]	
ln (R&D QTY) * R&D Pessimism	0.000	
	[0.665]	
R&D Uncertainty		0.000
		[1.061]
ln (R&D QTY) * R&D Uncertainty		0.000
		[0.927]
ROA	0.613***	0.613***
	[31.088]	[31.313]
Δ ROA	-0.103***	-0.103***
	[-5.825]	[-5.818]
ln (1+ Patents/RDC)	0.023***	0.023***
	[3.583]	[3.615]
ln (1+ R&D Exp/Asset)	-0.011	-0.011
	[-0.282]	[-0.284]
R&D Growth Dummy	0.003	0.003
	[0.648]	[0.635]
Δ APC	0.124	0.123
	[0.799]	[0.790]
ln (1+AD/Asset)	0.035*	0.035*
	[1.867]	[1.870]
ln(1+Capex/Asset)	-0.011	-0.010
	[-0.326]	[-0.291]
ln (Asset)	0.006***	0.006***
	[6.980]	[7.001]
FLS Tone	-0.047**	-0.047**
	[-2.697]	[-2.745]
10K Length	-0.012***	-0.012***
	[-7.321]	[-7.421]
Big N	0.000	-0.000
	[0.054]	[-0.001]

Fixed Effects:		
Year	Yes	Yes
Industry	Yes	Yes
Observations	15,579	15,579
Adjusted R-squared	0.426	0.426

TABLE 2.6: Examining the Effect of Competition on the Association between R&D Disclosure Quantity and ROA

This table reports the coefficient estimates of a regression of subsequent-period ROA on current R&D disclosure quantity after including the respective competition measure [HHI (column 1); PC_MARGIN (column 2); MKT_SIZE (column 3)] and a corresponding interaction term. All variable definitions are outlined in Appendix C. Each specification includes year fixed effects. t-statistics (in brackets) are based on standard errors that are clustered at the firm and year level. ***, **, * represents statistical significance at the 1%, 5% and 10% levels.

VARIABLES	**ROA (t+1)**		
	(1)	**(2)**	**(3)**
	HHI	**PC_MARGIN**	**MKT_SIZE**
ln (R&D QTY)	-0.006***	-0.007***	-0.006***
	[-3.902]	[-3.348]	[-4.081]
HHI	-0.019		
	[-1.581]		
HHI * ln (R&D QTY)	0.010		
	[1.155]		
PC_MARGIN		0.018	
		[1.622]	
PC_MARGIN * ln (R&D QTY)		0.000	
		[0.074]	
MKT_SIZE			0.000
			[0.684]
MKT_SIZE * ln (R&D QTY)			-0.000
			[-0.574]
R&D Tone	-0.000	-0.000**	-0.000
	[-0.218]	[-2.677]	[-0.583]
ROA	0.623***	0.620***	0.620***
	[31.924]	[26.612]	[30.281]
Δ ROA	-0.104***	-0.094***	-0.108***
	[-6.111]	[-4.663]	[-6.134]
ln (1+ Patents/RDC)	0.029***	0.036***	0.029***
	[4.328]	[4.330]	[4.026]
ln (1+ R&D Exp/Asset)	-0.008	-0.019	-0.005
	[-0.190]	[-0.347]	[-0.120]
R&D Growth Dummy	0.003	0.004	0.002
	[0.676]	[0.665]	[0.565]
Δ APC	0.076	0.165	0.169
	[0.448]	[0.647]	[0.962]
ln (1+AD/Asset)	0.036*	0.035	0.037
	[1.842]	[1.471]	[1.707]
ln(1+Capex/Asset)	-0.012	-0.019	-0.017
	[-0.357]	[-0.387]	[-0.439]
ln (Asset)	0.005***	0.005***	0.005***
	[6.908]	[4.297]	[6.379]
FLS Tone	-0.055**	-0.057**	-0.055**
	[-2.830]	[-2.232]	[-2.813]
10K Length	-0.012***	-0.012***	-0.012***
	[-7.441]	[-4.508]	[-6.854]
Big N	0.001	0.002	0.002
	[0.239]	[0.623]	[0.532]

Fixed Effects:			
Year	Yes	Yes	Yes
Observations	15,211	9,608	13,822
Adjusted R-squared	0.422	0.429	0.416

TABLE 2.7: Decomposing R&D Disclosures into Forward-looking (FLS) and Non-Forward-looking (Non-FLS)

This table reports the coefficient estimates of a regression of subsequent-period ROA on both FLS and Non-FLS R&D disclosure quantity, for both the full sample and only the R&D-intensive firms. All variable definitions are outlined in Appendix C. Each specification includes year and industry fixed effects. t-statistics (in brackets) are based on standard errors that are clustered at the industry and year level. ***, **, * represents statistical significance at the 1%, 5% and 10% levels.

VARIABLES	**ROA (t+1)**
ln (FLS)	-0.010***
	[-5.654]
ln (Non-FLS)	-0.002
	[-1.370]
R&D Tone	-0.000
	[-1.033]
ROA	0.613***
	[31.131]
Δ ROA	-0.103***
	[-5.851]
ln (1+ Patents/RDC)	0.021***
	[3.277]
ln (1+ R&D Exp/Asset)	0.020
	[0.494]
R&D Growth Dummy	0.003
	[0.826]
Δ APC	0.173
	[1.088]
ln (1+AD/Asset)	0.039*
	[2.095]
ln(1+Capex/Asset)	-0.012
	[-0.368]
ln (Asset)	0.006***
	[7.345]
FLS Tone	-0.040**
	[-2.262]
10K Length	-0.007***
	[-5.159]
Big N	0.000
	[0.130]
Fixed Effects:	
Year	Yes
Industry	Yes
Observations	15,579
Adjusted R-squared	0.427

TABLE 2.8: Additional Analysis

Panel A reports the coefficient estimates of a regression of subsequent-period ROA on current R&D disclosure quantity run on the sample of only R&D-intensive firms. Panel B reports the coefficient estimates of a regression of subsequent-period Adj ROA on current R&D disclosure quantity run on the entire sample. All variable definitions are outlined in Appendix C. The specification includes year and industry fixed effects. t-statistics (in brackets) are based on standard errors that are clustered at the firm and year level. ***, **, * represents statistical significance at the 1%, 5% and 10% levels.

Panel A: R&D intensive Firms

VARIABLES	ROA (t+1)
ln (R&D QTY)	-0.013***
	[-6.566]
R&D Tone	-0.000
	[-0.848]
ROA	0.600***
	[28.574]
Δ ROA	-0.109***
	[-5.201]
ln (1+ Patents/RDC)	0.029***
	[3.649]
ln (1+ R&D Exp/Asset)	-0.000
	[-0.008]
R&D Growth Dummy	0.003
	[0.799]
Δ APC	0.103
	[0.634]
ln (1+AD/Asset)	0.069**
	[2.299]
ln(1+Capex/Asset)	-0.002
	[-0.038]
ln (Asset)	0.006***
	[5.799]
FLS Tone	-0.045
	[-1.698]
10K Length	-0.017***
	[-6.753]
Big N	-0.002
	[-0.609]
Fixed Effects:	
Year	Yes
Industry	Yes

Observations	8,733
Adjusted R-squared	0.433

Panel B: Adj ROA

VARIABLES	Adj ROA (t+1)
ln (R&D QTY)	-0.005***
	[-3.329]
R&D Tone	-0.000
	[-0.617]
Adj ROA	0.626***
	[30.488]
Δ Adj ROA	-0.092***
	[-4.665]
ln (1+ Patents/RDC)	0.018**
	[2.532]
ln (1+ R&D Exp/Asset)	0.213***
	[4.715]
R&D Growth Dummy	-0.002
	[-0.414]
Δ APC	0.232
	[1.474]
ln (1+AD/Asset)	0.032
	[1.548]
ln(1+Capex/Asset)	-0.026
	[-0.924]
ln (Asset)	0.004***
	[6.040]
FLS Tone	-0.040**
	[-2.429]
10K Length	-0.010***
	[-6.456]
Big N	0.001
	[0.646]
Fixed Effects:	
Year	Yes
Industry	Yes
Observations	15,579
Adjusted R-squared	0.458

TABLE 2.9: Robustness Analysis using Alternative Proxies

This table reports the coefficient estimates of a regression of future performance on current R&D disclosure quantity using alternative measures. Panel A shows results using alternative measures of performance, cash flows. Panel B shows results using alternative measures of R&D disclosure quantity computed using plain word count (column 1), sentence count (column 2), and unique sentence count (column 3) approaches. All variable definitions are outlined in Appendix C. All specifications include year and industry fixed effects. t-statistics (in brackets) are based on standard errors that are clustered at the firm and year level. ***, **, * represents statistical significance at the 1%, 5% and 10% levels.

Panel A: Alternative Performance Measure

VARIABLES	CF (t+1)
ln (R&D QTY)	-0.002***
	[-3.756]
R&D Tone	0.000
	[1.128]
CF	0.608***
	[43.035]
Δ CF	-0.185***
	[-14.043]
ln (1+ Patents/RDC)	0.021***
	[3.012]
ln (1+ R&D Exp/Asset)	-0.070*
	[-1.844]
R&D Growth Dummy	0.002
	[0.787]
Δ APC	0.330*
	[2.056]
ln (1+AD/Asset)	0.023
	[1.238]
ln(1+Capex/Asset)	0.085***
	[3.995]
ln (Asset)	0.005***
	[7.895]
FLS Tone	-0.029
	[-1.760]
10K Length	-0.008***
	[-4.096]
Big N	0.004*
	[1.827]
Fixed Effects:	
Year	Yes

Industry	Yes
Observations	15,579
Adjusted R-squared	0.422

Panel B: Alternative R&D Disclosure Measures

VARIABLES	ROA (t+1)		
	(1)	**(2)**	**(3)**
	Plain Word Count	**Sentence Count**	**Unique Sentence Count**
ln (R&D QTY)	-0.022***		
	[-8.107]		
ln (R&D Sentences)		-0.007***	-0.007***
		[-4.679]	[-4.544]
R&D Tone	-0.042	-0.000	-0.000
	[-1.281]	[-0.913]	[-0.906]
ROA	0.610***	0.616***	0.616***
	[30.742]	[30.837]	[30.901]
Δ ROA	-0.101***	-0.104***	-0.104***
	[-5.700]	[-5.855]	[-5.861]
ln (1+ Patents/RDC)	0.022***	0.023***	0.023***
	[3.543]	[3.595]	[3.609]
ln (1+ R&D Exp/Asset)	0.017	0.002	0.004
	[0.449]	[0.050]	[0.091]
R&D Growth Dummy	0.003	0.003	0.003
	[0.834]	[0.759]	[0.754]
Δ APC	0.134	0.130	0.131
	[0.881]	[0.810]	[0.817]
ln (1+AD/Asset)	0.037*	0.042**	0.042**
	[1.973]	[2.283]	[2.278]
ln(1+Capex/Asset)	-0.003	-0.013	-0.013
	[-0.108]	[-0.391]	[-0.383]
ln (Asset)	0.006***	0.006***	0.006***
	[7.058]	[7.247]	[7.214]
FLS Tone	0.031	-0.046**	-0.045**
	[0.264]	[-2.580]	[-2.565]
10K Length	-0.015***	-0.008***	-0.008***
	[-6.816]	[-5.435]	[-5.453]
Big N	0.001	-0.000	-0.000
	[0.279]	[-0.013]	[-0.020]

Fixed Effects:			
Year	Yes	Yes	Yes
Industry	Yes	Yes	Yes
Observations	15,579	15,579	15,579
Adjusted R-squared	0.428	0.426	0.426