University of Pennsylvania
ScholarlyCommons

Publicly Accessible Penn Dissertations

2016

Informational Uniqueness, Corporate Disclosure and Information Environment

Yuen Kit Chau

University of Pennsylvania, yuenkit@wharton.upenn.edu

Follow this and additional works at: <https://repository.upenn.edu/edissertations>

Part of the [Accounting Commons](#), and the [Finance and Financial Management Commons](#)

Recommended Citation

Chau, Yuen Kit, "Informational Uniqueness, Corporate Disclosure and Information Environment" (2016).
Publicly Accessible Penn Dissertations. 1644.
<https://repository.upenn.edu/edissertations/1644>

This paper is posted at ScholarlyCommons. <https://repository.upenn.edu/edissertations/1644>
For more information, please contact repository@pobox.upenn.edu.

Informational Uniqueness, Corporate Disclosure and Information Environment

Abstract

This dissertation examines how the lack of comparable public peers (“informational uniqueness”) is related to a firm’s disclosure policy and information environment.

Having less information spillover from other public firms may present an information deficiency if it is not compensated by other components of the information environment. Using textual similarity in business description among firms to measure the extent of peer presence, I find that informational uniqueness is associated with a higher propensity by firms to provide ongoing bundled guidance. This is consistent with firms attempting to mitigate the information deficiency through strengthening their tacit commitment to continued disclosure by providing more information regularly on predictable schedules. Overall, I find a strong negative relationship between informational uniqueness and the quality of corporate information environment only among firms without regular bundled guidance. My results suggest that, while informational uniqueness can generate significant information deficiency, firms with strong tacit commitment to ongoing disclosure are largely able to compensate for the lack of information spillover from peers.

On the flip side, firms surrounded by many peers may be subject to the influence of peer dynamics on their disclosure behavior. While the presence of comparable public peers likely expands investors’ information endowment on the base firm through information spillover on related exposures, discretionary disclosures by peer firms may also signal information arrival. This can raise investors’ inferred probability that non-disclosure is due to strategic information withholding rather than the absence of new information. In relation to this, I show that the bid-ask spread of the base firm slightly increases when its closest peer initiates discretionary disclosures, but subsequently decreases upon disclosure by the base firm. In addition, I find that the number of comparable public peers is strongly positively associated with the frequency of discretionary disclosure. Overall, these results suggest that the presence of peer dynamics can induce firms with more peers to provide discretionary disclosures more frequently.

Degree Type

Dissertation

Degree Name

Doctor of Philosophy (PhD)

Graduate Group

Accounting

First Advisor

Brian Bushee

Second Advisor

Luzi Hail

Keywords

discretionary disclosure, informational uniqueness, information spillover, peer dynamics, peer identification, tacit disclosure commitment

Subject Categories

Accounting | Finance and Financial Management

INFORMATIONAL UNIQUENESS, CORPORATE DISCLOSURE AND
INFORMATION ENVIRONMENT

Yuen Kit Chau

A DISSERTATION

in

Accounting

For the Graduate Group in Managerial Science and Applied Economics

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy

2016

Supervisor of Dissertation

Co-Supervisor of Dissertation

Brian Bushee
Professor of Accounting

Luzi Hail
Associate Professor of Accounting

Graduate Group Chairperson

Eric Bradlow, K.P. Chao Professor; Professor of Marketing, Statistics, and Education

Dissertation Committee:

Brian Bushee, The Geoffrey T. Boisi Professor of Accounting

Luzi Hail, Associate Professor of Accounting

Paul Fischer, Samuel R. Harrell Professor of Accounting

Wayne Guay, Yageo Professor of Accounting

INFORMATIONAL UNIQUENESS, CORPORATE DISCLOSURE AND
INFORMATION ENVIRONMENT

COPYRIGHT

2016

Yuen Kit Chau

ACKNOWLEDGMENT

I am indebted to my dissertation committee: Brian Bushee (Co-Chair), Luzi Hail (Co-Chair), Paul Fischer, and Wayne Guay for their patient guidance and insightful comments. I would also like to thank Cathy Schrand, Jeremy Michels, and the rest of the faculty for their valuable advice, as well as my fellow Ph.D. classmates for their ongoing support. This dissertation has also benefited greatly from the comments from workshop participants at University of Pennsylvania, University of Hong Kong, Chinese University of Hong Kong, Hong Kong Polytechnic University, Nanyang Technological University, and Singapore Management University.

ABSTRACT

INFORMATIONAL UNIQUENESS, CORPORATE DISCLOSURE AND INFORMATION ENVIRONMENT

Yuen Kit Chau

Brian Bushee

Luzi Hail

This dissertation examines how the lack of comparable public peers (“informational uniqueness”) is related to a firm’s disclosure policy and information environment. Having less information spillover from other public firms may present an information deficiency if it is not compensated by other components of the information environment. Using textual similarity in business description among firms to measure the extent of peer presence, I find that informational uniqueness is associated with a higher propensity by firms to provide ongoing bundled guidance. This is consistent with firms attempting to mitigate the information deficiency through strengthening their tacit commitment to continued disclosure by providing more information regularly on predictable schedules. Overall, I find a strong negative relationship between informational uniqueness and the quality of corporate information environment only among firms without regular bundled guidance. My results suggest that, while informational uniqueness can generate significant information deficiency, firms with strong tacit commitment to ongoing disclosure are largely able to compensate for the lack of information spillover from peers.

On the flip side, firms surrounded by many peers may be subject to the influence of peer dynamics on their disclosure behavior. While the presence of comparable public peers likely expands investors' information endowment on the base firm through information spillover on related exposures, discretionary disclosures by peer firms may also signal information arrival. This can raise investors' inferred probability that non-disclosure is due to strategic information withholding rather than the absence of new information. In relation to this, I show that the bid-ask spread of the base firm slightly increases when its closest peer initiates discretionary disclosures, but subsequently decreases upon disclosure by the base firm. In addition, I find that the number of comparable public peers is strongly positively associated with the frequency of discretionary disclosure. Overall, these results suggest that the presence of peer dynamics can induce firms with more peers to provide discretionary disclosures more frequently.

TABLE OF CONTENTS

ACKNOWLEDGMENT	III
ABSTRACT	IV
LIST OF TABLES	VIII
LIST OF ILLUSTRATIONS	XI
CHAPTER 1: INFORMATIONAL UNIQUENESS, CORPORATE DISCLOSURE AND INFORMATION ENVIRONMENT	
1. INTRODUCTION	1
2. CONCEPT OF INFORMATIONAL UNIQUENESS	7
3. HYPOTHESES DEVELOPMENT	11
4. VARIABLE MEASUREMENT & EMPIRICAL SPECIFICATIONS.....	19
5. RESULTS	27
6. CONCLUSION	36
APPENDIX A: VARIABLE DEFINITION.....	38
APPENDIX B: MEASURING INFORMATIONAL UNIQUENESS	40
APPENDIX C: EXAMPLES OF INFORMATIONALLY UNIQUE FIRMS.....	44
TABLES.....	49
FIGURES	68
CHAPTER 2: PEER DYNAMICS AND DISCRETIONARY DISCLOSURE	
1. INTRODUCTION	72

2. PRIOR LITERATURE & HYPOTHESES DEVELOPMENT	78
3. VARIABLE MEASUREMENT & RESEARCH DESIGN.....	85
4. RESULTS	94
5. CONCLUDING REMARKS	102
APPENDIX A: VARIABLE DEFINITION.....	104
APPENDIX B: MEASURING PEER GROUP SIZE.....	106
APPENDIX C: EXAMPLES OF VOLUNTARY 8-K FILINGS	108
TABLES.....	109
 CHAPTER 3: SEARCHING FOR CLOSE PEERS	
1. INTRODUCTION	120
2. LITERATURE REVIEW	124
3. OVERVIEW OF DIFFERENT METHODOLOGY	126
4. COMPARISON TESTS	130
5. RESULTS OF COMPARISON & ADDITIONAL ANALYSES	134
6. SUMMARY	144
TABLES.....	147
 BIBLIOGRAPHY	 160

LIST OF TABLES

CHAPTER 1: INFORMATIONAL UNIQUENESS, CORPORATE DISCLOSURE AND INFORMATION ENVIRONMENT

Table 1 Summary Statistics

Panel A: Overall Descriptive Statistics	P.49
Panel B: Descriptive Statistics by Group	P.50
Panel C: Correlation Table	P.50

Table 2 Measure Validation

Panel A: Information Transfer from Closest Peer	P.51
Panel B: Uniqueness and Analyst Coverage	P.52

Table 3 Market Reactions during Earnings Announcements

Panel A: Uniqueness and Market Reactions	P.53
--	------

Table 4 Corporate Disclosures

Panel A: Uniqueness and Guidance Policy – Variation across Full Sample	P.54
Panel B: Uniqueness and Guidance Policy – Variation within Guiding Firms	P.55

Table 5 Information Environment

Panel A: Uniqueness and Information Environment	P.56
Panel B: Subsampling on the Presence of Regular Bundled Guidance Policy	P.57

Table 6 Alternative Measures of Informational Uniqueness

Panel A: Using Industry-Level and Industry-Adjusted Uniqueness	P.58
Panel B: Using Alternative Measures Based on Business Segment Data	P.59
Panel C: Using IV on Informational Uniqueness	P.60

Table 7 Controlling for Competition Perspective

Panel A: Controlling for Total No. of U.S. Public & Private Firms in Industry	P.61
Panel B: Controlling for Herfindahl Index Computed from US Census Data	P.63
Table 8 Additional Controls for Robustness Checks	
Panel A: Additional Firm Characteristics as Controls	P.65
Table 9 Time Series Variation	
Panel A: Decreases in Uniqueness due to IPOs of Close Peers	P.67
Panel B: Difference-in-Difference with Matched Firms as Controls	P.67

CHAPTER 2: PEER DYNAMICS AND DISCRETIONARY DISCLOSURE

Table 1 Summary Statistics	P.109
Table 2 Correlation Table	P.110
Table 3 Changes in Bid-Ask Spread When Closest Peer Discloses	P.111
Table 4 Industry Size and Frequency of Discretionary Disclosure	P.112
Table 5 Comparable Peer Size and Frequency of Discretionary Disclosure	P.113
Table 6 Timing of Discretionary Disclosure Using Cox Hazard Model	P.114
Table 7 Peer Similarity and Timing of Discretionary Disclosure	P.115
Table 8 Controlling for Frequency of Events	P.116
Table 9 Using Alternative Measure of Discretionary Disclosure	P.118
Table 10 Using Alternative Measure of Peer Size	P.119

CHAPTER 3: SEARCHING FOR CLOSE PEERS

Table 1 Comparison of % Top Peer in Analyst Co-Coverage Sample	
Panel A: Textual Method and Correlation-Based Method	P.147

Panel B: Correlation-Based Method Combined with Industry	P.148
Panel C: Industry Classification	P.148
Panel D: Industry Classification Combined with Firm Size	P.149
Table 2 Future Comovement in Key Market and Financial Variables	
Panel A: Comovement in Monthly Returns in the Next 5 Years	P.150
Panel B: Comovement in Daily Returns in the Next 1 Year	P.150
Panel C: Comovement in Earnings in the Next 5 Years	P.151
Panel D: Comovement in Operating Cash Flows in the Next 5 Years	P.151
Table 3 Explaining Valuation Multiple and Accounting Ratios	
Panel A: Textual Method and Correlation-Based Methods	P.152
Panel B: Randomly Selected Peer from Industry Classification	P.153
Panel C: Size-Matched Peer from Industry Classification	P.153
Table 4 Variation in Similarity with the Closest Peer	
Panel A: Comovement in Monthly Returns in the Next 5 Years	P.154
Panel B: Comovement in Daily Returns in the Next 1 Year	P.154
Panel C: Comovement in Earnings in the Next 5 Years	P.155
Panel D: Comovement in Operating Cash Flows in the Next 5 Years	P.155
Table 5 Similarity with the Closest Peers and Comovement in Future Periods	
Panel A: Comovement in Monthly Returns in the Next 5 Years	P.156
Panel B: Comovement in Daily Returns in the Next 1 Year	P.157
Panel C: Comovement in Earnings in the Next 5 Years	P.158
Panel D: Comovement in Operating Cash Flows in the Next 5 Years	P.159

LIST OF ILLUSTRATIONS

Figure 1 Uniqueness and Information Environment

Figure 1A: Uniqueness and Spread	P.68
Figure 1B: Uniqueness and Spread Conditional on Guidance	P.69
Figure 1C: Uniqueness and Illiquidity	P.70
Figure 1D: Uniqueness and Illiquidity Conditional on Guidance	P.71

CHAPTER 1

Informational Uniqueness, Corporate Disclosure, and Information Environment

1. Introduction

Information spillover from other publicly listed firms sharing similar business exposures is potentially an important component of a firm's information environment. Prior research has documented the existence of within-industry information transfers from earnings announcements (e.g. Foster 1981) and management forecasts (e.g. Baginski 1987), as well as the widespread use of comparable company analysis by practitioners (e.g. Demirakos, et. al. 2004). However, the availability of comparable peers in the public universe can vary greatly by firm. Certain firms have few or no close public peers, resulting in the absence of one potentially important information channel for investors. Unless other information sources, such as firm-initiated disclosures, can compensate for this missing channel, these “informationally unique” firms will suffer the costs of an opaque information environment, e.g., wider spreads (Glosten and Milgrom 1985), etc. In addition, informational uniqueness may also be related to “opportunistic” concerns due to investors having less knowledge on the information arrival process and less external information channels for cross-checking. These raise the question of whether and how firms' disclosure choices respond to “informational uniqueness” and to what extent firms can mitigate this potential information deficiency.

Due to the lack of information spillover from other public firms, investors may have lower prior precision on “informationally unique” firms in the absence of any firm-

initiated disclosures. In addition, with investors having less knowledge over whether and when managers are endowed with new information as well as less peer information for benchmarking and verification purposes, unique firms potentially have more opportunities to withhold or manipulate information. While investors can generally learn about the probability of information arrival from other peer firms, investors of informationally unique firms may have difficulties assessing when managers receive no new information and thus the absence of disclosure can be interpreted more negatively (e.g. Jung and Kwon 1988). Opacity and the potential for “opportunism” combined can present an exacerbated information asymmetry problem for informationally unique firms. If these firms do not try to compensate for the lack of peer information, outsiders would have higher prior uncertainty on their values, potentially leading to higher spreads and discounts on value (e.g. Glosten and Milgrom 1985).

Consistent with prior literature that disclosure commitment can help alleviate the discount (e.g. Diamond and Verrecchia 1991), I expect that informational uniqueness should be associated with higher incentives by firms to strengthen their commitment to disclosure by adopting a more transparent disclosure policy. However, how managers can credibly commit themselves to continued disclosure remains to be explored. In particular, disclosure may be viewed as more discretionary and opportunistic in nature for informationally unique firms on which investors have little knowledge over when managers receive new information. Without any signals of ongoing disclosure commitment, investors cannot tell if these firms will continue to disclose when new information is received. Thus establishing an *ex-ante* tacit commitment to provide

ongoing disclosure may be an effective way for informationally unique firms to mitigate the deficiency in their information environment.

In the paper, I hypothesize that bundling additional disclosures with earnings announcements may serve as a tacit commitment device that helps to strengthen a manager's tacit commitment to ongoing disclosure. In particular, providing additional forward-looking information during earnings announcements may also help preempt the information that investors would otherwise learn from peer firms during the subsequent quarter. For example, bundling guidance with earnings announcements may help build up investors' expectations that guidance will be provided regularly during subsequent earnings announcements. While firms may stop their unbundled guidance "quietly" without investors immediately noticing, the cessation of bundled guidance can be immediately revealed to investors due to its predictable schedule. Thus I expect that informational uniqueness is associated with a higher propensity by firms to provide additional information during their earnings announcements as a tacit commitment device to signal their ongoing releases.

To capture "informational uniqueness" in the paper, I attempt to measure the extent of the availability of comparable public peers in the market. My primary measure uses textual analysis and compares firm similarities by computing cosine similarity scores of business descriptions in 10-K filings across all U.S. public firms using a vector space model. Informational uniqueness is then represented by the logarithm of the inverse of the number of peer firms identified through textual analysis. The measure is validated by showing that, as predicted, unique firms receive less information transfer during the earnings announcements of their closest peer firm and attract less analyst coverage (Litov

et al. 2012). For robustness checks, I compute an alternative measure using similarities in business segment weights based on industry classification. The results using both measures are largely consistent in various tests, with the primary textual measure providing more statistical power in almost all cases.

In the first part of my analyses, I explore how a firm's disclosure policy is related to its "informational uniqueness" property. I begin by examining how informational uniqueness is associated with market reactions to earnings announcements. I find strong evidence that firms with higher degree of informational uniqueness generally have higher abnormal return volatility and trading volume during earnings announcements, consistent with news in earnings announcements by informationally unique firms being less preempted by information spillover from peers as well as generally deemed credible. This suggests that "opportunism" does not seem to be a huge concern for the earnings announcements of unique firms on average. It also highlights the potentially more important role of firm-initiated disclosures during earnings announcements in constituting the information environment for firms with high degree of informational uniqueness.

Next, I examine whether informationally unique firms are more likely to regularly provide additional disclosures during their earnings announcements in mitigating their information deficiency. As predicted, I find that informational uniqueness is associated with increased propensity by firms to adopt a regular bundled guidance policy both across the sample of all public firms and within the group of guiding firms, controlling for other aspects of the information environment. Conditional on guidance issuance, I find that informational uniqueness is positively associated with having guidance of longer

horizon and greater specificity. I also find that informationally unique firms appear to be more likely to substitute the use of unbundled guidance with bundled guidance. Overall, my results are consistent with firms on average attempting to mitigate their deficiency from informational uniqueness through bundling additional information releases with earnings announcements in strengthening their tacit commitment to a more transparent disclosure policy.

In the second part of the analyses, I evaluate how disclosure choices interact with “informational uniqueness” in shaping a firm’s overall information environment. The extent to which firm-initiated disclosures can mitigate the deficiency in their information environment from the lack of public peers may depend on whether information spillover from peers is mostly a subset of the information from firm-initiated disclosures and whether these disclosures are generally deemed credible. In the analyses, I use the bid-ask spread and Amihud’s measure of illiquidity (2002) to proxy for the quality of the overall information environment and examine how they relate to the interactions between informational uniqueness and disclosure policy.

My results find that informationally unique firms receive incremental benefits from providing ongoing additional disclosures during earnings announcements relative to other firms. While bundled guidance is positively associated with the quality of a firm’s overall information environment, the relationship is not necessarily true for unbundled guidance. These corroborate with my earlier findings that informationally unique firms are more likely to regularly provide additional disclosures during their earnings announcements in mitigating their information deficiency. In further subsampling analyses, I document a significant negative relationship between informational uniqueness and the quality of the

overall information environment among firms without regular bundled guidance. However, I do not find a significant negative relationship among firms with a regular bundled guidance policy. These results suggest that, while “informational uniqueness” is a highly economically significant factor in constituting a firm’s information environment, firms that demonstrate strong tacit commitment to disclosure are largely able to mitigate this deficiency in their information environment.

My study contributes to the literature by conceptualizing the idea of “informational uniqueness” from an informational prospective at the firm level to represent the extent of information spillover from other public firms. It introduces empirical measures to capture this concept and shows that the availability of peer information is an economically important component that constitutes a firm’s information environment, in addition to the three main forces (voluntary disclosure, mandatory disclosure and information intermediaries) discussed in Beyer et al. (2010). The study also specifically considers the interactions between “informational uniqueness” and corporate disclosure choice in shaping the corporate information environment.

My study also explores how the informational deficiency from lacking public peers are being addressed by firms and investigates if regularly providing additional information during earnings announcements may serve as an effective tacit commitment to ongoing disclosures. I show that informationally unique firms indeed receive incremental benefits from providing ongoing additional information during earnings announcements in the form of bundled guidance. Firms with higher degree of informational uniqueness also appear to have a higher propensity to provide ongoing bundled guidance on average. Overall, my results suggest that firms that demonstrate

strong tacit commitment to ongoing disclosure appear to largely be able to mitigate the problems arising from informational uniqueness.

In addition, the study contributes to the empirical disclosure literature by advancing the idea that bundling additional information releases with earnings announcements may help establish a stronger tacit commitment to disclosure, especially in the context of informationally unique firms. Using management guidance as an example of such additional disclosure, I show that bundled guidance is associated with an overall improvement in corporate information environment. However, this association does not seem to hold for unbundled guidance. Such benefit in signaling its ongoing releases may be one reason that bundled guidance has evolved to be the most prevalent type of guidance over time (Rogers and Van Buskirk 2013).

The remainder of the paper is organized as follows. Section 2 elaborates on the concept of “informational uniqueness” and relates it to existing literature. Section 3 develops my hypotheses. Section 4 describes my variables measurement and empirical specifications. Section 5 presents my results. Section 6 concludes.

2. Concept of Informational Uniqueness

In this study, “informational uniqueness”¹ is used to describe the extent of the availability of information spillover from other public firms, with more unique firms receiving less information spillover. Investors may obtain information on the underlying firm from other public firms with similar business exposures through a variety of channels, such as mandatory filings, voluntary disclosure, analyst coverage, or even

¹ In the paper, I use the terms “informational uniqueness” and “uniqueness” interchangeably to refer to the lack of information spillover from other public firms. It presents difficulties for outsiders to identify comparable public peers with similar business exposures for a firm.

private information acquisition impounded into security prices, etc. An extensive prior literature has documented the presence of within-industry information transfers from earnings announcements (e.g. Foster 1981) and management forecasts (e.g. Baginski 1987). In addition, the use of comparable company analysis is also prevalent by practitioners such as sell-side analysts (e.g., Demirakos, et al. 2004).

Information from other public firms is thus potentially an important component of a firm's information environment². In Beyer et al. (2010), the results from the decomposition of quarterly return variance seem to suggest that over 70% of the variance is unexplained by firm-initiated disclosures and analyst reports. It remains unclear how much of this large unexplained variance is attributable to noise in returns or other sources of information, e.g. information spillover from peers, etc. To help shed light on this issue, this paper examines the cross-sectional variations in the extent of information spillover among firms and seeks to understand its implications on corporate disclosure policy and information environment.

In the study, I define “uniqueness” strictly from an informational perspective at the firm-level. Unique firms thus specifically refer to those firms that have little information spillover from other public firms. In other words, unique firms can be characterized as those that have no or few comparable peers sharing similar business exposures available in the set of public firms. As firms become increasingly dissimilar in terms of business exposure, there is likely decreasing signal and increasing noise to use the information of one firm to infer on the other firms. It is also important to note that information spillover

² In a related literature, Badertscher, et al. (2013) suggests that public firm presence generates positive externalities by reducing industry uncertainty and facilitates more efficient private firm investments. It provides some other evidence on the important role that information spillover can potentially play.

mainly involves peer firms that are public. Close peers that are private are unlikely adding much to the information environment of other firms.

The term “uniqueness” is also often used in various business contexts as a desirable characteristic, such as in marketing and corporate strategy³. Uniqueness, however, may also bring informational problems. One such problem is identified by Litov et al. (2012) showing that uniqueness in corporate strategy choice discourages analyst coverage, possibly due to higher cost in information gathering and analysis. Litov et al. (2012) further proposes the term “uniqueness paradox” to describe the tradeoff between value creation and information problem (in discouraging analyst coverage) manager faces in selecting corporate strategy. However, it did not consider the role of corporate disclosure in potentially mitigating the information problem.

While the conceptual notion of “uniqueness” from existing studies in other fields mostly relates to “uniqueness” from a competition perspective at the product level, my study focuses exclusively on the informational aspects of “uniqueness” at the firm level. There are likely two sources of “informational uniqueness”: (1) a firm is “unique” because it has no competitors conducting similar businesses; (2) a firm is “unique” only from an informational perspective because all other firms engaging in similar businesses are either private or part of a large conglomerate.

The first source of “uniqueness” is defined with respect to all other firms. This may manifest in a variety of business dimensions, such as unique products/services, unique

³ For example, uniqueness is often espoused as a desirable product characteristic in marketing, e.g. the concept of unique selling proposition at the product level. There is also some evidence in finance that having unique products are positively associated with stock market valuations (Hoberg and Philips 2014). From the management literature (e.g. Barney 1986), uniqueness in strategy choice at the firm-level is also regarded as an important path for value creation. In most of these notions, uniqueness is viewed positively from a competition perspective, mainly focusing on the product level.

operating models, unique business or corporate strategy, exclusive relationships with suppliers and customers. On one hand, uniqueness can be attributable to the presence of sustained competitive advantage so that no other firms can replicate easily, e.g. due to access to proprietary technology or scarce resources. On the other hand, some firms may become unique because they are engaging in a business that no other firms want to be in.

The second source of “uniqueness” only concerns comparable peers from the universe of public firms as investors usually cannot obtain much information from peers that are private. Thus, “informational uniqueness” can be simply due to all other firms engaging in similar businesses being either private or part of a large conglomerate⁴. Even though the businesses of these “informationally unique” firms may have close competitors in the private space or from giant conglomerates, it is difficult for outsiders to learn much about their businesses using information from other public firms.

While it is difficult to clearly distinguish between these two sources due to the absence of detailed information on private firms, some initial evidence suggests that most “informationally unique” firms are unique only with respect to public firms from an informational perspective, but not with respect to all firms. Manually reading through the competition section of 10-K filings from a few dozen samples of “informationally unique” firms, I find that almost every firm mentions the presence of competitors from either private companies or giant conglomerates. In addition, according to the data from Statistics of U.S. Businesses, every U.S. public firm has some private peers within the

⁴ For example, Groupon was “informationally unique” as a firm as its peer firm LivingSocial was private and other competitors were either small local private companies or embedded in the businesses of giant public companies such as Google and Amazon. As another example, Acme United, a leading supplier of cutting, measuring and first aid products, was “informationally unique” because its major competitor for each product market were either private, foreign or part of a giant conglomerate such as Johnson & Johnson. Some other examples of “informationally unique” firms around year 2012 are listed in Appendix C.

same primary industry based on its 6-digit NAICS classification even though it may be the only firm that is public among firms in the same primary NAICS code.

3. Hypotheses Development

Regardless of the sources of uniqueness, unique firms may face certain informational challenges. One potential challenge is opacity. Due to the lack of information spillover from other public peers, investors may have less information endowment from external sources on unique firms. For a firm with many close public peers, outsiders can potentially continuously glean some information on the firm from the disclosures or even private information impounded into stock prices from its public peers. In addition, investors will likely find it difficult to identify peer groups for unique firms for valuation and benchmarking exercises. Furthermore, uniqueness may increase the cost of information production by information intermediaries such as sell-side analysts and thereby discouraging analyst coverage (Litov et al. 2012).

Another potential challenge for uniqueness is the concern for “opportunism”. As firm-initiated disclosure is likely the predominant information source to investors for unique firms, unique firms can be considered as a “monopoly” provider of information on their own businesses. Disclosures by unique firms are likely less immediately verifiable using peer information relative to firms with abundant information spillover from peers⁵. The abundance of peer disclosures may provide some constraints on opportunistic behaviors as drastic unexplained differences can often cause suspicions by investors. In addition, investors of unique firms likely possess little knowledge over whether and when

⁵ For example, in the case of forward-looking information disclosures, investors of unique firms may only be able to verify the information when audited earnings reports are released or cash flows are realized in the future.

managers are endowed with new information, potentially giving managers more rooms to strategically time their disclosures or even temporarily withhold certain new information. Furthermore, managers of unique firms may receive more benefits from information manipulation if investors put a larger relative weight on firm-initiated disclosures in updating their beliefs when other information sources are not readily available.

Despite these potential informational challenges, mandatory disclosure requirements (e.g. audited financial reports), associated litigation risks, and market forces can place some constraints on opportunistic behaviors by unique firms. While uniqueness may provide enhanced capability for firms to withhold or manipulate information opportunistically, it is unclear if most unique firms choose to engage in opportunistic behaviors given various disciplinary forces. In addition, firm-initiated disclosures and other sources of information may be a “superset” of information spillover from peer firms. Thus the relationship between uniqueness and the quality of overall information environment is likely dependent on how firms respond to uniqueness in their disclosures to compensate for the lack of information spillover from peers and to what extent information spillover is a “subset” of information from firm-initiated disclosures.

3.1. Informational Uniqueness and Market Responses

With fewer comparable public peers available, “informationally unique” firms likely receive less pre-announcement information flows due to the lack of information transfers from other public firms. The information contained in the disclosure announcements of more unique firms is thus less likely to be pre-empted by other sources of information. With less information spillover from peers, the earnings announcements by more unique firms should provide more surprises to investors. This should trigger more pronounced

market reactions in the market if information contained in earnings announcements are generally deemed credible for unique firms.

Due to potentially higher prior uncertainty on firm value, investors of unique firms may put a heavier weight on firm-initiated disclosures in updating their beliefs on firm value. This has a direct effect in increasing investors' reaction to announcement news for more unique firms⁶. On the other hand, concerns for "opportunism" may reduce the credibility of disclosures and therefore their market reactions, especially for the discretionary disclosure of informationally unique firms on which investors have little knowledge on the information endowment of managers and limited access to external information channels for verification.

As the accounting system provides relatively consistent and verifiable performance measures across firms due to external auditing, financial information during earnings announcements can potentially serve as an important information source for investors of unique firms due to possibly fewer concerns for rampant manipulation relative to non-financial information contained in other occasions of discretionary disclosure. Thus I expect that market reactions during earnings announcements should be stronger for more unique firms, similar to prior findings in Eberhart (2001) that the number of firms in an industry is related to lower stock return volatility around earnings announcements. This leads to my first hypothesis:

H1: Informational uniqueness is associated with more pronounced market reactions during earnings announcements.

⁶ However, an indirect effect of a heavier weight on disclosure is an increased benefit from information manipulation. According to Fischer and Verrecchia (2000), the direct effect should dominate the indirect effect and thus higher prior uncertainty should be associated with more pronounced market responses.

3.2. Informational Uniqueness and Disclosure Policy

Opacity from the lack of peer information spillover and the concern for “opportunism” can present an exacerbated information problem for unique firms, especially if unique firms of worse type can easily mimic unique firms of better type. If firms do not try to compensate for the lack of peer information and signal their values, outsiders would have lower information precision on firm value or may infer a worse type, potentially leading to higher spreads and value discounts (e.g. Akerlof 1970, Glosten and Milgrom 1985, Diamond and Verrecchia 1991, Lambert et al. 2012). While investors can generally learn about the probability of information arrival from other peer firms, investors of informationally unique firms may have difficulties assessing when managers receive no new information and thus the absence of disclosure can be interpreted more negatively (e.g. Dye 1985, Jung and Kwon 1988). In addition, difficulties in hedging positions for unique firms may discourage trading by certain types of investors (e.g. long/short equity hedge funds) and further reduce its liquidity.

Consistent with prior literature that disclosure commitment can help alleviate the discount (e.g. Diamond and Verrecchia 1991, Baiman and Verrecchia 1996, Leuz and Verrecchia 2000), I expect that unique firms should generally have higher incentives to mitigate the information deficiency through strengthening their commitment to continuous disclosure. If informationally unique firms can demonstrate a strong commitment to ongoing disclosure, the absence of disclosure at any given time is then more likely to be interpreted as receiving no new signals instead of possessing adverse information. On the flip side, firms with many close public peers can potentially free-ride on information spillover and thus have reduced incentives to implement potentially costly

commitment mechanisms to disclose continuously. Heinle and Verrecchia (2015) also generate similar theoretical predictions that firms with higher prior uncertainty on cash flows and lower cash flow correlation with peers are more likely to commit to disclose.

Thus establishing an ex-ante tacit commitment to provide ongoing disclosure can possibly be an effective way for unique firms to mitigate their information deficiency. However, how managers can credibly commit themselves to continued disclosure remains to be explored. In particular, disclosure may be viewed as more discretionary and opportunistic in nature for informationally unique firms on which investors have little knowledge on when managers receive new information. Without any signals of ongoing disclosure commitment, investors cannot tell if these firms will continue to disclose when new information is received. To mitigate the deficiency in their information environment, informationally unique firms may need to credibly signal their commitment to provide ongoing disclosures through an auxiliary commitment device.

I hypothesize that bundling additional disclosures with earnings announcements may serve the role of this auxiliary commitment device by helping to strengthen a manager's tacit commitment to ongoing disclosure. For example, bundling guidance with earnings announcements may help to build up an expectation among investors that guidance will be issued regularly during subsequent quarterly earnings announcements no matter whether the signal is good or bad⁷. This is supported by prior research showing that

⁷ Having a track record of guidance during prior earnings announcements likely suggest that managers usually possess enough information to provide guidance by the time of the earnings announcements, following the assumption that firm's possession of information is history dependent (Einhorn and Ziv 2008). Stopping bundled guidance is likely interpreted more negatively and deemed more costly relative to stopping unbundled guidance.

bundled guidance is highly sticky over time (e.g. Billings, et al. 2015)⁸. In addition, guidance cessation can be immediately revealed to investors when guidance is bundled with earnings announcements⁹. On the other hand, information releases outside a predictable window, e.g. unbundled guidance, can be viewed as more discretionary and opportunistic in nature (with regards to both the disclosure decision and the timing of disclosure) upon managers receiving a signal, the existence of which investors may not even be aware of for informationally unique firms. In the case of unbundled guidance, unless firms voluntarily announce the cessation of guidance, it is difficult for investors to tell whether and when exactly firms “quietly” stop providing unbundled guidance.

Providing additional disclosures during earnings announcements (e.g. having a regular bundled guidance policy) is thus tantamount to an increase in tacit commitment to continued disclosure, which may help reduce outsiders’ assessed variance on firm value. In addition, providing additional forward-looking information during earnings announcements may also help to preempt the information that investors would otherwise learn from peer firms during the subsequent quarter. However, these informational benefits likely do not come without any costs. As conference calls held concurrent with earnings announcement have become increasingly common in recent years (e.g. Matsumoto and Roelofsen 2011), these additional information provided during earnings announcements may invite more scrutiny during conference calls. In the case of bundled

⁸ In Billings, et al. (2015), they find that firms which bundled their guidance with earnings announcements in the prior quarter and in the same quarter of the previous year provide bundled guidance in the current quarter 87.2% of the time in their sample.

⁹ Chen et al. (2011) finds that the three-day return window around guidance stoppage announcements is negative, suggesting that there is immediate cost associated with guidance cessation if investors can identify them. The cost of discontinuing bundled guidance can help to make the tacit commitment more credible.

guidance, managers may also lose the flexibility to wait for more certain signals before issuing their guidance in the future and risk damaging their reputation by providing inaccurate or imprecise guidance (e.g. Lee, et al. 2012).

Facing the deficiency in their information environments, unique firms should have more incentives to bear the costs and strengthen their tacit commitment to a more transparent disclosure policy by bundling additional disclosures with earnings announcements. This leads to second hypotheses below:

H2: Informational uniqueness is associated with a higher propensity by firms to strengthen their tacit commitment to continued disclosure by providing additional disclosures during earnings announcements.

3.3. Informational Uniqueness and Overall Information Environment

The relationship between informational uniqueness and the overall information environment is likely linked to how significant the information deficiency is from the lack of information spillover from peers and how firms respond to this deficiency through their disclosure policy based on their cost-benefit tradeoff. In particular, it may depend on how much investors can learn from public peer firms is a “subset” of the information from firm-initiated disclosures or other information sources. If information spillover from peers is indeed an important component that constitutes a firm’s information environment, I expect that uniqueness should be associated with a weaker overall information environment when the other components of the information environment (e.g. firm-initiated disclosures) cannot compensate for the loss of information spillover from peers. In particular, informational uniqueness may generate significant

informational barriers to investors for firms that have not demonstrated a strong tacit commitment to ongoing disclosure. This leads to my third hypothesis:

H3: Informational uniqueness is associated with a weaker overall information environment among firms without strong tacit commitment to continued disclosure.

As investors likely have little knowledge over whether and when managers are endowed with new information for informationally unique firms, having a regular disclosure policy that provides additional information during earnings announcements can help strengthen the *ex-ante* tacit commitment to ongoing disclosure for these firms and preempt the need for information transfer from peers. For informationally unique firms, unbundled disclosure may not be as effective as bundled disclosure in signaling a tacit commitment to ongoing disclosure since it can be viewed as more discretionary and opportunistic in nature. From our prior literature, higher *ex-ante* commitment to disclosure is shown to be positively associated with lower discount or cost of capital (e.g. Leuz and Verrecchia 2000, Hail 2011). Thus, I expect that bundled guidance should provide incremental benefits on the corporate information environment for more unique firms in addressing their information deficiency. These notions lead to my fourth hypothesis:

H4: Strengthening tacit commitment to continued disclosure provides incremental benefits to the corporate information environment of informationally unique firms by mitigating their information deficiency.

To what extent strengthening a firm's tacit commitment to providing ongoing additional disclosures may help to mitigate the information deficiency due to uniqueness is ultimately an empirical question. For firms with high level of tacit commitment to

disclosure, it is possible that information spillover from peers is just a “subset” of the other sources of the information, such as firm-initiated disclosures. Thus, “informational uniqueness” may not be necessarily associated with a weaker overall information environment when other components of the information environment are sufficiently strong and fully compensate for the information deficiency. In the case of guidance, prior literature provides some weak evidence that the presence of management guidance helps preempt information transfer from earnings announcements of firms in the same industry (Pownall and Waymire 1989). However, whether firms adopting a regular bundled disclosure policy can completely eliminate the information deficiency from uniqueness remains to be empirically tested. If information spillover from peers is largely subsumed by information from bundled disclosure, I would expect that informational uniqueness is not significantly associated with a weaker overall informational environment among firms with strong tacit commitment to continued disclosure.

4. Variable Measurement and Empirical Specifications

4.1. Measuring Informational Uniqueness

4.1.1 Measure Based on Business Segment Data

To measure “informational uniqueness” at the firm-level, I start off with using the business segment data¹⁰. Based on business segment information, this measure is intended to capture whether there are other public firms that operate in similar business segments as the underlying firm. It assumes that business exposures are similar within

¹⁰ In my measure, I compare the business segment distribution of each firm to all other public firms using COMPUSTAT segment data and attempt to measure the extent of the presence of comparable peers. Litov et al. (2012) also uses similar business segment data source to compute their measure of corporate strategy uniqueness by computing uniqueness as deviation in business segment weights from the centroid firm in the industry based on primary industry classification.

each business segment as defined by 6-digit NAICS codes. Thus, a single-segment firm is highly unique if it is the only publicly listed single-segment firm in its industry with no other multi-segment firms having a high segment weighting on that particular industry. Similarly, a multi-segment firm is unique if there are no or few other public firms having similar business segment weight distribution. Operationally, the measure is implemented by computing similarity scores across all public firms using a vector space model where each business segment is regarded as a dimension in the model¹¹. The natural logarithm of the inverse of the number of peer firms with high similarity score above a defined threshold is used to represent “informational uniqueness”. In the paper, this measure (*segunique*) is mostly used as a secondary measure for robustness checks.

4.1.2 Measure Based on Textual Business Description

Instead of assuming homogeneity within 6-digit NAICS industry classification across firms, I attempt to expand the scope of information set in comparing similarities across firms by utilizing the textual business description section in mandatory 10-K filings. This measure assumes truthful disclosure in the business description section of 10-K filings as governed by regulation S-K. Compared to the measure based on business segment data, it has the advantage of being able to distinguish among firms within an industry classification using a richer information set based on textual information. Operationally, this textual measure uses the terms in company business description (instead of business segment) as the dimensions in the vector space model to compute cosine similarity scores

¹¹ The vector space model here represents each firm as a vector in an n-dimensional Euclidean space, where n is the number of unique 6-digit NAICS business segment and the value of each vector element is the proportion of sales contribution by each segment of the firm. The similarity of any two firms can be viewed as the angle between the two vectors representing the two firms, with a smaller angle representing more similar firms. More details on measure construction are included in Appendix B.

across all public firms¹². The natural logarithm of the inverse of the number of peer firms identified through having high similarity scores is used as my primary measure (*txtunique*) to represent “informational uniqueness”¹³. I also define an indicator variable (*txtunique indicator*) to represent the top 25% unique firms, i.e. firms that have similarity score with their closest peer in the bottom quartile.

4.1.3 Measure Validation

4.1.3.1 Internal Consistency

While using business segment and textual business description to measure similarities across firms may both suffer from certain measurement errors, measures of uniqueness computed using the two methods should show significant positive correlation. As predicted, correlation analysis confirms that the uniqueness measures calculated using both methodology have a statistically significant positive correlation coefficient of 0.29.

4.1.3.2 Information Transfer

The extent of the information transfer during earnings announcements is likely related to how similar the firms are to one another. Thus, I expect that more unique firms should have less information transfer from the earnings releases of their closest peer. As

¹² The vector space model here represents the 10-K business description of each firm as a vector in an n-dimensional Euclidean space, where n is the number of unique phrases in the textual business description and the value of each vector element is represented by the commonly used *tf-idf* (term frequency-inverse document frequency) weighting scheme in the information retrieval literature (e.g. Salton and Buckley 1988). Using *tf-idf* scheme weighs down the effects of too frequently occurring terms and weighs up the effects of less frequently occurring terms. The similarity of any two firms can be viewed as the angle between the two vectors representing the two firms, with a smaller angle representing more similar firms. Similar textual analysis methodology using a vector space model is also used in Brown and Tucker (2010) to compare within-firm MD&A modifications and Hoberg and Philip (2015) to reclassify industries based on product similarities. Further details and differences are discussed in Appendix B. The performance of my textual method in identifying peer firms is further tested in Chapter 3 of the dissertation.

¹³ The similarity threshold is arbitrarily chosen to define 25% firms as having no peers in each year to aid interpretation. The results are robust to using other similarity thresholds around the chosen cutoff, as varying the similarity threshold mostly just changes the scale of the measure.

expected, I show that unique firms identified using textual analysis receive significantly less information transfer from the earnings announcements of their closest peer firm in Panel A of Table 2.

4.1.3.3 Analyst Coverage

Litov et al. (2012) shows that “corporate strategy uniqueness”, defined using a firm’s choice of business segment, is associated with lower analyst coverage. Similar to corporate strategy uniqueness in Litov et al. (2012), I expect that my measure of “informational uniqueness” should also be significantly negatively related to analyst coverage. Consistent with this conjecture, I show that my uniqueness measures are indeed significantly negatively associated with analyst coverage in Panel B of Table 2.

4.2 Sample Selection

In my full sample, I include all U.S. public firms from 1996 to 2013 where major financial reporting items, stock returns, business segment and valid textual business description data are available from CRSP, COMPUSTAT and SEC Edgar filings. The earliest year goes back to 1996 when electronic Edgar data started to become available. The sample used in most test specifications that require management guidance data, however, only dates back to 2003 when more consistent IBES guidance data started to become available. I exclude firms in the regulated industries such as utilities and banking as there are often different disclosure requirements for these firms. After imposing further data requirements on major control variables, the post-2002 sample used in most specifications comprises of 29,839 observations.

4.3 Research Design

In this study, the identification of the relationship among uniqueness, disclosure and information environment mainly relies on the cross-sectional variations across firms. Informational uniqueness is an inherent and relatively sticky firm-level property in relationship to all other firms. It depends not only on the decision of the firm itself but also on other firms. As a firm's entry into its major businesses is likely determined long before the firm becomes public, such business decisions are likely a first-order consideration for most firms before they start considering disclosure policy and information environment. Thus, there should be little concern over reverse causality bias in this setting. As shown in my data, "informational uniqueness" is quite stable over time and the autocorrelation in the textual uniqueness measure is close to 0.75.

Major identification threats include measurement errors with respect to the uniqueness measures and correlated omitted variable bias. To mitigate concerns for measurement errors, I use both uniqueness measures constructed using textual business description and business segment data in my tests for robustness checks. I also include various controls for firm characteristics, such as analyst coverage, institutional ownership, firm size, stock returns, stock beta, market-to-book ratio, stock return volatility, years of listing, profitability, leverage, etc., to mitigate concerns for correlated omitted variables. In additional analyses, I use the proportion of private peers as an instrument variable to informational uniqueness to capture the more exogenous sources of variation and include more control variables for robustness checks.

4.3.1 Market Responses during Earnings Announcements

To examine how market reactions during annual earnings announcements are related to informational uniqueness, I regress measures of market responses on measures of

uniqueness, measures of guidance policy (e.g. frequency of bundled guidance, frequency of unbundled guidance, frequency of earnings preannouncements), and various controls on other firm characteristics. I consider two dimensions of market responses: abnormal return volatility and abnormal trading volume. Measures of abnormal volatility and trading volume during annual earnings announcements are calculated relative to non-event windows following Landsman et al. (2012). The sample includes observations after 2002 due to requirement on IBES guidance data. To be consistent with *H1*, the sign for the coefficient on uniqueness should be positive in (1).

$$\text{Market Responses} = \alpha + \beta_1 \text{Uniqueness} + \beta_2 \text{Guidance} + \beta_3 \text{Controls} + \beta_4 \text{Year FE} + \varepsilon \quad (1)$$

4.3.2 Tacit Commitment to Disclosure

As a typical example of additional disclosures provided by firms during earnings announcements, the presence of a regular bundled guidance policy is used to represent the strengthening of a firm's tacit commitment to ongoing disclosure in testing *H2*. Prior literature suggests that management guidance contributes significantly to the information reflected in stock prices (e.g. Ball and Shivakumar 2008) and may help reduce information asymmetry (e.g. Coller and Yohn 1997). Other prior empirical studies also suggest that firms generally have an incentive to mitigate the deficiency in their information environment through voluntary disclosures in the form of management guidance (e.g. Gong et al. 2013, Balakrishnan et al. 2014). Thus establishing an ongoing bundled guidance policy to investors can be regarded as an important form of

strengthening a firm's tacit commitment to continued disclosure through building a reputation for being "forthcoming" as opposed to being "strategic"¹⁴.

To examine how a firm's guidance policy varies with informational uniqueness, I regress the guidance policy of firms on measures of uniqueness and other controls for firm characteristics. To represent tacit commitment to a more transparent disclosure policy, I use the annual frequency of bundled guidance or an indicator for regular bundled guiders (with three or more bundled guidance annually)¹⁵. I also examine guidance properties such as horizon and specificity as part of the tacit commitment. These dimensions of guidance tend to be relatively stable and investors can easily form an expectation about its ongoing releases. I measure guidance horizon and guidance specificity by taking the annual averages of the forecast horizons of EPS guidance and the number of forecasted items per guidance respectively. The sample is post-2002 due to requirement on IBES guidance data¹⁵.

To be consistent with *H2*, I expect that unique firms are more likely to provide bundled guidance and guidance of longer horizon and specificity. Thus, the sign for the coefficient on uniqueness should be positive in (2) when frequency of bundled guidance, indicator for regular bundled guiders, guidance horizon and guidance specificity are used as dependent variables. I also expect that the coefficient on uniqueness in (2) should be more positive for high-growth firms that care about their valuation levels in capital

¹⁴ Terminology follows Beyer and Dye (2012) in which "forthcoming" managers disclose all forecasts they receive while "strategic" managers disclose forecasts they receive only when it is in their self-interest to do so.

¹⁵ Guidance is categorized as bundled if it is issued in less than two days after earnings announcements. As Chuk et. al (2013) point out that a previously commonly used guidance database CIG was not able to identify all guidance, I further restrict my sample observations to those that are covered by analysts as defined by IBES to minimize errors in incorrectly identifying guiding firms as non-guiding firms.

markets as low-growth firms may receive fewer benefits from more precisely revealing their type.

$$Disclosure\ Policy = \alpha + \beta_1 Uniqueness + \beta_2 Controls + \beta_3 Year\ FE + \varepsilon \quad (2)$$

4.3.3 Overall Information Environment

To investigate how informational uniqueness and disclosure policy interact in constituting the overall information environment, I regress proxies for the quality of the overall information environment on measures of uniqueness, disclosure policy, the interaction terms between uniqueness and disclosure policy, as well as other controls for firm characteristics. I use both the ratio of bid-ask spread over price and Amihud's measure of illiquidity (2002) to represent the quality of the overall information environment. To represent disclosure policy, I use the frequency of bundled guidance and unbundled guidance or an indicator variable for regular bundled guiders. To be consistent with *H3*, the coefficient on uniqueness should be positive. To be consistent with *H4*, the coefficient on the indicator for regular bundled guider or frequency of bundled guidance as well as the coefficients on the interaction terms between uniqueness and the indicator for regular bundled guider or frequency of bundled guidance should also be negative when either bid-ask spread or illiquidity is used as the dependent variable in (3).

$$Information\ Environment = \alpha + \beta_1 Uniqueness + \beta_2 Disclosure + \beta_3 Uniqueness * Disclosure + \beta_4 Controls + \beta_5 Year\ FE + \varepsilon \quad (3)$$

In further tests, I use subsampling analyses by the presence of regular bundled guidance and re-run (3) without the interaction terms to examine whether under certain conditions the adverse effects of uniqueness on a firm's overall information environment can be completely eliminated. Jointly estimating the model using both subsamples and

testing for the difference in coefficients is equivalent to examining (3) but fully interacting all independent variables with the presence of regular bundled guidance. To facilitate interpretation, I use the indicator variable to represent the top quartile firms in terms of uniqueness in the test. To be consistent with *H3 and H4*, the coefficient on uniqueness indicator should be significantly positive in a subsample of firms without regular bundled guidance but insignificant in a subsample of firms with regular bundled guidance.

5. Results

5.1 Summary Statistics

Table 1 provides the basic summary statistics for the main variables in the paper in Panel A. For variables related to guidance, the sample starts from 2003. The table shows that roughly 43% of firms in the sample provide regular bundled guidance after imposing data requirements on controls. Panel B provides the descriptive statistics broken down into whether the firm is regarded as the top quartile unique firms or not. It shows that unique firms have a worse overall information environment on average, measured by the bid-ask spread and Amihud's measure of illiquidity (2002). In addition, unique firms tend to have less analyst coverage and institutional ownership. To be conservative, I control for both analyst coverage and institutional ownership in all my tests to focus on the direct effect from the lack of information spillover. The table also shows that unique firms tend to be smaller, less levered and have slightly lower beta and higher market-to-book ratio. These variables are all included as controls in subsequent tests. Panel C provides the pairwise correlation among some of the main variables in the tests. All the signs of correlation are consistent with my hypotheses in univariate setting.

5.2 Market Reactions during Earnings Announcements

Table 3 displays the results on the relationship between informational uniqueness and market reactions during annual earnings announcements of firms. I find that informational uniqueness is highly positively associated with abnormal return volatility and abnormal trading volume during annual earnings announcements. Controlling for guidance behaviors and other firm characteristics, abnormal return volatility changes from 0.50¹⁶ for firms in the lowest quartile of uniqueness to 0.94¹⁶ for firms in the highest quartile of uniqueness. Similarly, abnormal trading volume changes from 0.34¹⁶ for firms in the lowest quartile of uniqueness to 0.54¹⁶ for firms in the highest quartile of uniqueness. These highly economically significant results are consistent with investors having lower prior precision on unique firms and therefore more surprises during earnings announcements due to the lack of information spillover from peer firms. My results strongly support *H1* and confirm prior similar finding by Eberhart (2001). The larger magnitude of market reactions for unique firms also suggests that the earnings of unique firms are generally deemed credible by investors¹⁷. Thus, “opportunism” does not seem to be a huge concern for the earnings announcements of unique firms¹⁸.

5.3 Disclosure Policy

¹⁶ The numbers are calculated using the sample median for other firm characteristics.

¹⁷ In untabulated tests, I find that the same results hold for quarterly earnings announcements. In addition, I find that the earnings response coefficients using analyst consensus as expectation also increase with uniqueness on average.

¹⁸ As financial information needs to be subsequently verified by external auditors, opportunism is likely not a big concern for information contained in financial reporting. It is not clear if opportunism is not a concern for other types of more qualitative discretionary disclosures of unique firms. In untabulated test, I do not find a more pronounced market reaction for unique firms for other types of more qualitative discretionary disclosures (Cooper et al. 2015) that are subsequently filed through 8-Ks. This may either suggest that the content for these disclosures are equally surprising for both unique and non-unique firms or that the investors may find these types of disclosures less credible for unique firms.

In Table 4, I examine how the disclosure policy of firms varies with uniqueness. In Panel A, I focus on the decision to provide additional disclosure during earnings announcements as tacit commitment to disclosure. I find that informational uniqueness is significantly positively associated with a firm's propensity to provide bundled guidance. Controlling for other firm characteristics, the propensity to provide regular bundled guidance changes from 37%¹⁶ for firms in the lowest quartile of uniqueness to 50%¹⁶ for firms in the highest quartile of uniqueness. This is consistent with firms generally attempting to mitigate the information deficiency from lacking information spillovers through regular bundled guidance¹⁹. In columns (3) and (4), I find that the positive association between uniqueness and bundled guidance are driven mostly by firms with high growth potential²⁰. This is consistent with high-growth firms having more capital market incentives to address the information problem by credibly communicating their values to avoid pooling with firms of worse types.

In Panel B of Table 4, I examine the relationship between uniqueness and properties of guidance among guiding firms. Conditional on firms providing management guidance, the results show that uniqueness is positively associated with the frequency of bundled guidance, but negatively associated with the frequency of unbundled guidance. This is consistent with unique firms being more likely to substitute unbundled guidance with bundled guidance to strengthen their tacit commitment to continued guidance, as investors likely have more uncertainty over the information endowment of the managers

¹⁹ My results in the paper are robust to alternative specifications such as using Poisson regression for frequency variables and using probit model for indicator variables in the dependent variables.

²⁰ In further cross-sectional tests, I find that the results are also stronger in subsamples of firms that are more profitable, have higher analyst coverage/institutional ownership, and have multiple business segments. Statistical significance of the results is present within nearly all subsamples except for the subsample of neglected firms (defined by no analyst coverage).

of unique firms and unbundled guidance can be regarded as more discretionary and opportunistic in nature. In columns (3) and (4), I further find that uniqueness is positively associated with guidance of longer horizon and higher specificity. Overall, my results are consistent with firms attempting to mitigate the information deficiency by strengthening their tacit commitment to continued disclosure through providing additional disclosure such as guidance during earnings announcements, providing strong support for $H2^{21}$.

5.4 Overall Information Environment

In Table 5, I assess how informational uniqueness interacts with disclosure policy in constituting the overall information environment. The results from Table 4 suggest that different firms may have different incentives and cost-benefit tradeoffs in addressing the informational deficiency from uniqueness. Heterogeneity in disclosure choices provides the cross-sectional variations to isolate the direct influence of informational uniqueness on the information environment. In Panel A, I show that my measure of informational uniqueness is indeed significantly negatively associated with the quality of the overall information environment, providing strong support to $H3$. With the inclusion of controls for the presence of other informational channels such as guidance policy, analyst coverage, institutional ownership and other firm characteristics, the bid-ask spread changes from 0.69%¹⁶ for firms in the lowest quartile of uniqueness to 0.92%¹⁶ for firms in the highest quartile of uniqueness while the Amihud's measure of illiquidity (2002)

²¹ In addition, Ball et al. (2012) proposes the confirmation hypothesis that audited financial reporting and voluntary disclosure of managerial private information are complement in communications with investors due to enhanced credibility from ex-post verifiability. Consistent with the confirmation hypothesis, I find that more unique firms also seem to provide reported earnings of high quality on average, measured by accruals quality (Dechow and Dichev 2002). This also helps to refute an alternative explanation that more unique firms provide more bundled guidance because guidance is generally easier to beat through opportunistic manipulation.

changes from 1.53¹⁶ for firms in the lowest quartile of uniqueness to 3.65¹⁶ for firms in the highest quartile of uniqueness (see Figure 1A and 1B).

The results in columns (1) and (3) of Table 5 Panel A also confirm that bundled guidance, which resembles a stronger tacit commitment to disclosure, is positively associated with the quality of the overall information environment²². However, the results suggest that unbundled guidance can be negatively associated with the quality of the overall information environment. This might be because unbundled guidance is more discretionary in nature and may often be triggered by some adverse exogenous events (e.g. Rogers, et al. 2009). These results corroborate with earlier findings in Table 4 and help explain why unique firms are more likely to bundle additional disclosure with earnings announcements as tacit commitment mechanism instead of other types of more discretionary disclosure mechanisms, e.g. unbundled guidance, to address their information deficiency.

In columns (2) and (4) of Table 5 Panel A, I examine whether strengthening a firm's tacit commitment to disclosure provides incremental benefits in mitigating the deficiency associated with informational uniqueness by including an interaction term between uniqueness and the presence of regular bundled guidance. Controlling for bundled guidance and other firm characteristics, I find that the interaction term is significantly negatively associated with the quality of the overall information environment. This

²² In addition, I find that informationally unique firms incur statistically negative returns upon the cessation of bundled guidance, suggesting that it is costly to breach the tacit commitment. In untabulated test, I find that the cumulative returns around the three-day window of earnings announcements during which firms stop providing bundled guidance to be significantly negative for informationally unique firms, amounting to about -2% on average. Such incremental costs may also help to strengthen their tacit commitment to continued disclosure through bundling guidance with earnings announcements.

supports *H4*, suggesting that unique firms receive incremental benefits from stronger tacit commitment to ongoing disclosure with respect to their information environment. Coefficients on the interaction terms in columns (2) and (4) have almost the same magnitude as the coefficients on informational uniqueness, providing preliminary evidence that firms with a regular bundled guidance policy may be able to largely eliminate the information deficiency resulting from uniqueness.

To provide further evidence on whether firms with stronger tacit commitment to ongoing disclosure can largely eliminate the information deficiency from the lack of information spillover from peers, I use subsampling analyses by splitting the sample into firms with regular bundled guidance and those without regular bundled guidance in Panel B of Table 5. I find that informational uniqueness is negatively associated with the quality of the overall information environment only among firms without bundled guidance in columns (1) and (3). However, informational uniqueness does not show a significant negative association with the quality of the overall information environment among firms with regular bundled guidance in columns (2) and (4). Figure 1B and 1D provide some additional graphical illustrations on these relationships as well as their economic magnitude across uniqueness of different quartiles. Overall, my results are consistent with both *H3* and *H4* and suggest that firms with strong tacit commitment to ongoing disclosure through regular bundled guidance policy can largely overcome the information problem from lacking information spillover from peers.

5.5 Additional Analyses

5.5.1 Alternative Measures of Informational Uniqueness

In Table 6, I rerun the main test specification in the paper using various alternative measures of informational uniqueness, including an industry-level and industry-adjusted measure based on textual data, an alternative measure based on business segment data as well as an instrument variable constructed using the proportion of private peers within the industry a firm operates.

In Panel A, I further split my textual uniqueness measure into an industry-level component and industry-adjusted component based on Fama-French 48 industry classification for my main tests. Both the industry-level and industry-adjusted uniqueness measures show consistent signs, and are significant in all specifications. However, the industry-level uniqueness measure is much more significant in most specifications²³.

In Panel B, I implement my main tests using an alternative measure of uniqueness based on business segment data. The results indicate that my uniqueness measures using both textual and business segment data produce consistent and significant signs in the main tests, with the primary measure being much more significant in all specifications. This is consistent with findings in earlier measure validation tests and suggests that the alternative measure is likely coarser than the textual measure in capturing informational uniqueness.

In Panel C, I further use the proportion of private firms in the firm's primary NAICS 6-digit industry (excluding the underlying firm in calculation) as an instrument variable

²³ In theory, if there is a perfect industry classification scheme that is fine enough so that all firms within an industry can be regarded as homogeneous, all variations in uniqueness should be at the industry level. While I expect that a significant source of variation in uniqueness is at the industry level, I have also included industry fixed effects using Fama-French 48 industry classification to control for unobserved heterogeneity at the industry level for robustness checks in untabulated tests. The statistical significance of uniqueness variables generally drops in all specifications with the inclusion of industry fixed effect but uniqueness still shows up statistically significant.

for informational uniqueness for my main tests. Conditional on the underlying firm being public, higher proportion of private firms in the firm's industry likely correspond to less information spillover from other public firms²⁴. The findings using the instrumental variable approach is also consistent with my prior findings.

5.5.2 Controlling for Competition Perspective

In Table 7, I attempt to further distinguish between the information perspective and the competition perspective on uniqueness through using additional data covering both public and private firms. In Panel A, I include an additional control for the natural logarithm of total number of U.S. public and private firms in the firm's primary NAICS 6-digit collected from the *Statistics of U.S. Businesses*. In Panel B, I include an additional control for the industry concentration ratio (Herfindahl-Hirschman index for the 50 largest companies) of each NAICS 6-digit industry from US 2007 *Economic Census* calculated using data from all US firms within the manufacturing sector, similar to Ali et al. (2009). Under both panels, I find that my results continue to hold after putting in the extra controls for competition.

5.5.3 Additional Controls

In Table 8, I include additional control variables in my main tests for robustness checks, such as an indicator for loss firms, asset turnover, litigation risks proxied by high-risk industry dummies (Francis et al. 1994), proprietary costs proxied by the frequency of words on protected sources of competitive advantage in 10-Ks (e.g. patents, licenses, copyrights), industry concentration proxied by Herfindahl index, business complexity

²⁴ The instrument satisfies the relevance condition as evidenced by the high F-statistic of 74.58 in the first-stage Wald test. It also likely satisfies the exclusion restriction as the proportion of private peers in an industry is mostly beyond the control and influence of the underlying firm.

proxied by the number of business segments and the number of geographical segments, accounting comparability (Franco et al. 2011), etc. While data requirements on additional controls further shrink the samples, the results are not significantly changed with the addition of these extra controls²⁵.

5.5.4 Time-Series Variation – Difference-in-Difference Analysis

While I expect that variations across informational uniqueness are mostly cross-sectional, I also attempt to examine some residual time-series variations as an alternative research design by identifying a sample of firms experiencing a seemingly exogenous decrease in their information uniqueness due to comparable private peers completing their initial public offerings. While this may mitigate concerns for correlated omitted variable that is stable through time, such time-series variations may not provide a powerful setting to test the relationship. In addition to being limited to a tiny sample of 44 firms, adding a new peer from IPO may not provide a significant decrease in information uniqueness. It may also not represent a very “clean” event as there may be some other confounding effect from peer IPOs, e.g. increased competition.

Panel A of Table 9 shows the first difference in earnings announcement volatility, regular bundled guidance policy, spread and illiquidity for the subsample of 44 firms that experience a decrease in information uniqueness after the IPOs of their close peers. Panel B of Table 10 shows the difference-in-difference of the variables for the 44 “treated”

²⁵ In untabulated tests, I have included more additional specific controls for each specification, including prior probability of beating guidance, proportion of guiding firms within the industry, frequency of capital raising activities, research and development expenses (Brown and Kimbrough 2011), return synchronicity (Morck et al. 2000), return skewness (as an additional variable for predicting litigation risk in Kim and Skinner 2012), etc. Similar results are obtained. I have also used a firm-quarter specification and followed the same specification as Table 3 of Rogers and Van Buskirk 2013 and added my measures of informational uniqueness to the specification. My results on the propensity to issue bundled guidance continue to hold in the firm-quarter specification.

firms matched with 44 size-matched “control” firms that exhibit similar prior uniqueness, analyst coverage and profitability. Samples only include at most 3 years before and after the event depending on data availability. The signs for the coefficients are consistent with my other prior tables but only the changes in spread and illiquidity in columns (3) and (4) are statistically significant in this time-series setting, possibly due to the lack of power in such a tiny sample. The results also indicate that there is not much change in guidance policy after the event, consistent with the sticky nature of regular bundled guidance.

6. Conclusion

In the paper, I conceptualize the idea of “informational uniqueness” as the lack of information spillovers from comparable public peers and examine its relationship with corporate disclosure and information environment. To capture uniqueness, I construct two measures using business segment data and textual business description to represent the availability of comparable public peers by computing similarity scores across firms in a vector space model.

The lack of information spillover from peers and concerns for potential “opportunism” can present an exacerbated information problem for unique firms. Consistent with firms having an incentive to mitigate such an information deficiency, I find that informational uniqueness is associated with a higher propensity by firms to provide additional disclosures during earnings announcements to strengthen its tacit commitment to ongoing disclosure.

Overall, I find a significant negative association between uniqueness and the quality of the overall information environment among firms without regular bundled guidance, but not among firms with regular bundled guidance. My results suggest that, while

“informational uniqueness” is a highly economically significant factor in constituting a firm’s information environment, firms with strong tacit commitment to disclosure are able to largely compensate for the lack of information spillover from peers.

Appendix A: Variable Definition

<i>Variable</i>	<i>Definition</i>
Measures of Informational Uniqueness	
<i>segunique</i>	Natural logarithm of the inverse of the number of comparable public peers using business segment information
<i>txtunique</i>	Natural logarithm of the inverse of the number of comparable public peers using textual business description in 10-K filings
<i>txtunique indicator</i>	Indicator for firms in the top quartile of “informational uniqueness” using textual business description
Dependent Variables	
<i>abnormal volume</i>	Abnormal trading volume during annual earnings announcements [-1 day, +1 day] relative to non-announcement periods [-60 days, -10 days] & [+10 days, +60 days] following Landsman, et. al. [2012]
<i>abnormal volatility</i>	Abnormal return volatility during annual earnings announcements [-1 day, +1 day] relative to non-announcement periods [-60 days, -10 days] & [+10 days, +60 days] following Landsman, et. al. [2012]
<i>regular bundled guider</i>	Indicator for regular bundled guidance issuer with three or more bundled guidance in a fiscal year
<i>frequency- bundled</i>	Frequency of bundled guidance in a fiscal year
<i>frequency- unbundled</i>	Frequency of unbundled guidance in a fiscal year
<i>guidance specificity</i>	Average number of forecasted items per annual guidance in a year
<i>guidance horizon</i>	Average EPS forecast horizon per guidance in a year
<i>spread</i>	Average bid-ask spread over price in a fiscal year
<i>illiquid</i>	Amihud’s illiquidity measure (2002)
Standard Firm Controls	
<i>mb ratio</i>	Market-to-book ratio
<i>stock return</i>	Cumulative stock returns over the fiscal year
<i>stock volatility</i>	Standard deviation of monthly returns over the fiscal year
<i>% inst ownership</i>	% institutional ownership
<i>stock beta</i>	Stock beta
<i>roa ratio</i>	Income before extraordinary items / total assets at the beginning of the year
<i>firm size</i>	Natural logarithm of total assets
<i>years of listing</i>	Number of years since initial public offering
<i>coverage</i>	Natural logarithm of (number of analysts + 1)
<i>leverage</i>	Total debt / total equity
Other Variables	

<i>frequency- preann.</i>	Frequency of preannouncement of earnings in a fiscal year
<i>industry-level uniqueness</i>	Industry average uniqueness (<i>txtunique</i>) at Fama-French 48 industry level
<i>industry-adj. uniqueness</i>	Deviation from average uniqueness (<i>txtunique</i>) at Fama-French 48 industry level
<i>totindpeersize</i>	Natural logarithm of the total number of U.S. public & private firms in the firm's primary NAICS industry
<i>US Census herfindahl</i>	Herfindahl-Hirschman index for the 50 largest companies in each manufacturing NAICS 6-digit industry classification from the US 2007 Economic Census data
<i>loss firms</i>	Indicator for loss firms (<i>roa</i> < 0)
<i>asset turnover</i>	Revenue / total asset
<i>FPS litigation industries</i>	Indicator for industries with high litigation risks (Francis et al. 1994)
<i>word counts about protected advantage</i>	Counts of words in 10-K to describe protected sources of competitive advantage, e.g. patents, licenses, copyrights
<i>industry concentration</i>	Industry herfindahl index calculated using public firms
<i># industry segments</i>	Number of business segments
<i># geo. segments</i>	Number of geographical segments
<i>accounting comparability</i>	Accounting comparability measure (Franco et al. 2011)

Appendix B: Measuring Informational Uniqueness

1. Alternative Measure Based on Business Segment Data

To measure “informational uniqueness” in the universe of publicly listed firms based on similarities in business exposures, one feasible source of data is using the business segment disclosure data by firms under the assumption that business exposures are similar within each business segment as defined by 6-digit NAICS codes.

This measure is constructed using similar Compustat Industrial and Segments data as used in Litov et al. (2012) but differs in the following sense. Litov et al. (2012) is more concerned about corporate strategy uniqueness and computes uniqueness as deviation in business segment weights from the centroid firm in the industry based on primary industry classification. If the centroid firm is a multi-segment firm, then all single-segment firms will be regarded as more unique in their corporate strategy. For my measure of uniqueness, I compare firms against all other public firms instead of against the centroid firm in their primary industry classification. If there are more than 2 single-segment firms having the same business segment, these firms will have a similarity score of 1, which is regarded as highly non-unique. Thus, this measure assumes that firms/segments within a 6-digit NAICS code are homogeneous and makes no distinction across these firms/segments in any other dimensions, such as business strategy, products and services, geographical exposure, etc.

Operationally it is implemented by computing cosine similarity scores across all publicly listed firms on an annual basis using a vector space model, which represents each firm as a vector in an n-dimensional Euclidean space, where n is the number of unique 6-digit NAICS business segment and the value of each vector element is the proportion of sales contribution by each segment of the firm. Cosine similarity scores are calculated using the following formula:

$$\text{similarity} = \cos(\theta) = \frac{A \cdot B}{\|A\| \|B\|}$$

Suppose in a simplified world with only 3 industries X, Y and Z, company A has 100% revenue from industry X. Company B has 80% revenue from industry X and 20% revenue from industry Y while company C has 10 % revenue from industry X, 20% revenue from industry Y and 70% from industry Z. The vector space model thus contains 3 dimensions: X, Y, Z.

Applying the formula, the cosine similarity score between A and B is calculated by $0.8/[1*\sqrt{0.8^2+0.2^2}]$, which is 0.97. The cosine similarity score between B and C is calculated by $(0.8*0.1+0.2*0.2+0)/[\sqrt{0.8^2+0.2^2}*\sqrt{0.1^2+0.2^2+0.7^2}]$, which is 0.12. While this simple scenario has only 3 dimensions in the vector space model, there can be as many as 1,200 dimensions in the vector space model when 6-digit NAICS industry codes are used to represent the dimensions in the vector space.

Uniqueness based on business segment information can then be represented in a variety of methods. In Table 2 Panel B and Table 7, my business segment based measure of “informational uniqueness” (*segunique*) is represented by the logarithm of the inverse of the number of close peers above certain similarity score threshold. In counting the number of close peers for a given firm using business segment information, a similarity score of 0.867, equivalent to $\cos 30^\circ$, is chosen as the similarity score threshold. While the threshold is arbitrarily chosen, robustness checks are done using a variety of thresholds and the results are largely similar.

2. Primary Measure Based on Textual Business Description

Similar to the measure based on business segment information, my measure based on textual business description replaces business segment weights by the relative term frequency in the business description section of 10-K filings in the vector space model on an annual basis. Thus, a firm-year level of uniqueness measure can also be constructed. For this primary measure, I only include the first 2,000 words in the business description for textual analysis for more apples-to-apples comparison as firms tend to put the most essential parts of business description upfront but may often diverge in content on the latter part of the business description. For example, some firms may present much more lengthy details on a variety of items such as regulatory environment, management team profiles, etc., in the latter part of the business description section.

Instead of considering all words as being equally important as in Hoberg and Philip (2015), I apply the commonly used tf-idf (term frequency-inverse document frequency) weighting factor, which weighs down the effects of too frequently occurring terms and weighs up the effects of less frequently occurring terms, where terms are defined by word/phrases in my vector space model²⁶. This is because most common terms are unlikely providing much useful information in differentiating among firms. Mathematically, each business description in the corpus d_i is represented by a vector in the word space considering the distinctive power of each word (*idf*), where tf_{ij} represents the term frequency of word j for firm i , idf_j represents the inverse document frequency of word j and M is the number of unique words in the entire corpus.

²⁶ One major difference is that I use all pre-processed meaningful words/phrases within the first 2,000 words of business description and Hoberg and Philip (2015) considers the binary occurrence of nouns and pronouns in the whole section of business description as it attempts to focus more on capturing product similarity. Similar to Hoberg and Philip (2015) using a 25% threshold (excluding 75% most common nouns and proper nouns) in calculating product similarities among firms, I also use a stopword threshold of 25% (excluding the 75% most common terms) but exclude extremely rare words/phrases that only occur once across all business descriptions annually in calculating the similarity score in my main measure. As tf-idf weighting scheme is being used, my measure is not sensitive to the choice of stopword threshold. Brown and Tucker (2010) also use the tf-idf weight scheme in comparing within-firm MD&A changes. Table 5 in Hoberg and Philip (2015) performs some benchmarking tests on various parameters. It shows that tf-idf weighting scheme using all words works better than tf-idf weighting scheme using only nouns and proper nouns in all cases as well as stopword threshold of 25% using only nouns and proper nouns in most cases. The results of the paper are robust to different combinations of parameters in stopword thresholds and the number of words retained from business description section.

$$d_i = \{tf_{i1} * idf_1, \dots, tf_{ij} * idf_j, tf_{iM} * idf_M\}$$

Uniqueness based on business description can then be similarly represented by a number of ways. My primary measure of informational uniqueness (*txtunique*) is represented by the logarithm of the inverse of the number of close peers above certain similarity score threshold. In each year, I define firms in the bottom 25% in similarity score with their closest peer as unique firms (*txtunique indicator*). In other words, these unique firms have no peers above a certain similarity threshold. In calculating the size of the peer group for *txtunique*, I use the same similarity threshold to count the number of close peers for a given firm. While the threshold is arbitrarily chosen based on defining 25% firms as having no close peers in each year, robustness checks are done using a variety of thresholds and the results are largely similar. In addition, although not tabulated in the paper, I also use some other representations to capture the concept of uniqueness, such as the negative of the cosine similarity with the closest peer (*txtmaxcosine*), the negative of the average cosine similarity with the top 4 closest peers (*txttop4cosine*), etc.

Various measures of uniqueness computed using either the negative of the cosine similarity with the closest peer, the negative of the average cosine similarity of the top 4 closest peers or the logarithm of the inverse of the number of close peers above certain similarity thresholds are highly correlated with one another. For example, the cosine similarity score with the closest peer (*txtmaxcosine*) has a correlation coefficient of 0.91 with the average cosine similarity score of the top 4 closest peers (*txttop4cosine*). The negative of the average cosine similarity score with the top 4 closest peers (*txttop4cosine*) has a correlation coefficient of 0.80 with the logarithm of the inverse of the number of close peers above certain similarity score threshold (*txtunique*). For the ease of presentation, I mainly use *txtunique* as my primary measure of uniqueness in my analyses.

Instead of assuming that all operations within a 6-digit NAICS business segment are homogeneous, this measure attempts to expand the scope of information set in considering the similarities across firms by utilizing the textual business description information in 10-K filings. It adds additional information in calculating the uniqueness measure compared to business segment data and has the advantage of being able to distinguish among firms within an industry segment classification. As business description section is part of mandatory disclosure governed by regulation S-K and contains mostly factual information, it is unlikely that certain firms would strategically choose words that are common or uncommon to describe their businesses, causing systematic biases in the measure²⁷.

²⁷ There is fairly detailed framework under regulation S-K on the content of business description in the 10-K filings, with over 3,800 words of extensive guidance, e.g. on products and services, raw materials, seasonality, intangibles, working capital, customers, order backlogs, competition, research and development, regulation, number of employees, etc. However, it is possible that textual analysis based on business description section may suffer from noise in measurement due to differential writing styles by different firms.

However, my informational uniqueness measure using 10-K business description has the limitation that it captures informational uniqueness with respect to all U.S. public firms that file 10-Ks only. Some identified unique firms under my measure may not be really “informational unique” from a global perspective if these firms have close comparable public peers from foreign countries²⁸. This may suggest that my measure may have larger measurement errors for firms with large foreign market presence. In untabulated tests, I split my sample of firms based on the proportion of foreign pre-tax income and find that the results are robust in both subsamples but stronger in the subsample of firms with domestic focus. This is consistent with the nature of the limitation of my measure described above.

²⁸ For example, Autodesk appears to be “informationally unique” from my measure, but its close competitor Dassault Systemes is a French company which does not file 10-Ks and therefore is excluded from my universe of public firms for comparison. In most case (except for some multinational companies), this may not be a huge concern as foreign firms mostly focus on local foreign markets and thus their markets are potentially different from the U.S. public firms in my sample.

Appendix C: Examples of Informationally Unique Firms

Acme United

- A leading worldwide supplier of innovative cutting, measuring and first aid products to the school, home, office, hardware, sporting goods and industrial markets
- 10-K Competition Section
 - o “The Company competes with many companies in each market and geographic area. The Company believes that the principal points of competition in these markets are product innovation, quality, price, merchandising, design and engineering capabilities, product development, timeliness and completeness of delivery, conformity to customer specifications and post-sale support.”
 - o “The major competitor in the cutting category is Fiskars Corporation. The major competitor in the measuring category is Maped. The major competitor in the safety category is Johnson and Johnson.”

Brinks

- Premium provider of secure logistics and security solutions
- 10-K Competition Section
 - o “Brinks competes with large multinational, regional and smaller companies throughout the world. Our largest multinational competitors are G4S plc (U.K.); Loomis AB (Sweden); Prosegur, Compania de Seguridad, S.A. (Spain); and Garda World Security Corporation (Canada).”
- Analyst (Macquarie Initiating Report 2015)
 - o Less of a “transportation” or “security” company than it is a proxy for cash as a method of payment
 - o 40% EPS exposure to Venezuela
 - o Controls the largest share (about 20%) of global cash and valuables transportation market
 - o Use Commercial and Professional Services (e.g. Casella Waste Systems, Xerium Technologies, Performant Financial Corp, Cenveo, Heritage-Crystal Clean, Energy Recovery) as comps

CECO Environmental

- A leading global industrial technology company focused on providing innovative solutions for the environmental, energy, and fluid handling/filtration industries
- 10-K Competition Section
 - o “We believe that there are no singly dominant companies in the product recovery, air pollution control, fluid handling and filtration product and service industries. These markets are fragmented with numerous small and regional participants.”
- Analyst (Cowen Initiating Report 2012)

- “Although there is stiff competition in each specific segment and product, no other public company offers the full breadth of CECE products and services. Companies were selected for our comparable valuation analysis because they either have some overlap, compete in closely related markets, or serve the same customer set with complementary products.”

Dun & Bradstreet

- The world's leading source of commercial information and insight on businesses
- 10-K Competition Section
 - “We are subject to highly competitive conditions in all aspects of our business. However, we believe no competitor offers our complete line of solutions or can match our global data and analytic capabilities including proprietary capabilities quality resulting from our DUNSRight Quality Process.”
- Analyst (Piper Jaffray Initiating Report 2010)
 - “D&B’s database lies at the core of all its product offerings and serves as a significant competitive barrier to entry. Underscoring its database is the DUNSRight process, which quality controls all inputs and outputs and which the company views as a major source of its competitive advantage. “
 - Use information services companies as peer group

Groupon

- A global leader in local commerce, making it easy for people around the world to search and discover great businesses and merchandise
- 10-K Competition Section
 - “Since our inception, a substantial number of competing group buying sites have emerged around the world attempting to replicate our business model, from very small startups to some of the largest companies in the world.”
- Analyst (UBS Initiating Report 2012)
 - Competition in Daily Deals comes from Amazon Local, LivingSocial, Google Offers, and a host of smaller companies / Competition in Goods comes from Amazon, eBay, Overstock.com, woot!, etc.
 - “Groupon’s 1P business is tilted toward inventory disposal on a large scale. We believe Overstock.com to be a fair comparable for this segment.” / “As for 3P, it is more difficult to find a true comparable given the business tilts more toward lead generation (versus third-party fulfillment or marketplace for goods models).”

Hershey

- Largest producer of quality chocolate in North America/ global leader in chocolate and sugar confectionery
- 10-K Competition Section

- “We sell our brands in highly competitive markets with many other multinational, national, regional and local firms. Some of our competitors are much larger firms that have greater resources and more substantial international operations.”
- Analyst (Jefferies Initiating Report 2011)
 - Use large-cap U.S. packaged food as peers, such as General Mills, McCormick, Kellogg, etc.

Liquidity Services

- Leading online auction marketplaces for surplus and salvage assets
- 10-K Competition Section
 - “We currently compete with other e-commerce providers; auction websites; government agencies that have created websites to sell surplus and salvage assets; and traditional liquidators and fixed-site auctioneers.”
- Analyst (William Blair Initiating Report 2007)
 - “LSI is a distinctive business model that combines facets of other publicly traded companies.”
 - “The basic procurement and sales model is like a distributor, but with a unique source of goods from the reverse supply chain, and a sales model that is different in two ways—it is an auction model and it is online-only. LSI is an auctioneer, but without the capital and buyer-acquisition constraints of bricks and mortar. Like e-tailers, it uses the Internet to inexpensively reach a huge number of buyers, yet its buying base is commercial rather than consumers.”
 - List related categories as comparable companies in online e-tailers (e.g. Amazon), online infrastructure (e.g. Digital River), Auction Companies (e.g. eBay), traditional distributors (e.g. CDW)

Peerless Systems

- Licenses imaging and networking technologies and components to the digital document markets (became private in 2015)
- 10-K Competition Section
 - “The market for outsourced imaging systems for digital document products is highly competitive and characterized by continuous pressure to enhance performance, add functionality, reduce costs and accelerate the release of new products. We compete on the basis of the set of core technologies we sublicense from third parties, technology expertise, product functionality and price. Our technology primarily competes with solutions developed internally by OEMs.”

Pool Corp

- Largest wholesale distributor of swimming pool and related backyard products
- 10-K Competition Section

- “We are the largest wholesale distributor of swimming pool and related backyard products and the only truly national wholesale distributor focused on the swimming pool industry in the United States. We face intense competition from many regional and local distributors in our markets...We also face competition, both directly and indirectly, from mass-market retailers and large pool supply retailers...”
- Analyst (JPMorgan Initiating Report 2005)
 - “Only pure play in attractive pool distribution market”
 - Use high-quality, niche distribution companies with solid secular growth business models as peers

OpenTable

- Provide solutions that form an online network connecting reservation-taking restaurants and people who dine
- 10-K Competition Section
 - The traditional competitor for our restaurant software and hardware solutions is the pen-and-paper reservation book. We also compete with reservation management software companies, point of sale providers and waitlist management system providers.
 - Our primary competitor for diners making reservations is the phone.
- Analyst (JPMorgan Initiating Report 2012)
 - “A leading online restaurant reservations site in North America”
 - “Faces significant competition from Livebookings in Germany, and Livebookings launched Freebookings in the US in 4Q11”
 - Use online bookings/reservations companies such as Priceline and local online players such as Angie’s List and Yelp as peer group

Snap-On

- A leading global innovator, manufacturer and marketer of tools, equipment, diagnostics, repair information and systems solutions for professional users performing critical tasks
- 10-K Competition Section
 - Snap-on competes on the basis of its product quality and performance, product line breadth and depth, service, brand awareness and imagery, technological innovation and availability of financing (through SOC or its international finance subsidiaries). While Snap-on does not believe that any single company competes with it across all of its product lines and distribution channels, various companies compete in one or more product categories and/or distribution channels.
- Analyst (Wunderlich Initiating Report 2014)
 - “Snap-on develops innovative products to make work easier for professionals in the automotive repair sector and in multiple industrial segments.”

- “Snap-on operates an industrial and tools business with high barriers to entry and geographic diversification that results in sustainable fundamentals. The valuation for the core business should be more in line with a high-quality industrial company.”

Steinway Musical Instruments

- Largest domestic manufacturing of musical instruments (became private in late 2013)
- 10-K Competition Section
 - “Steinway & Sons Pianos – The level of competition our pianos face depends on the market definition. Steinway & Sons pianos hold a unique position at the top of the grand piano market, both in terms of quality and price. While there are many makers of pianos, only a few compete directly with our Steinway brand. Other manufacturers of primarily higher priced pianos include Bösendorfer and Fazioli. Because Steinway pianos are built to last for generations, a relatively large market exists for used Steinways. It is difficult to estimate the significance of used piano sales because most are conducted in the private aftermarket. However, we believe that used Steinway pianos provide the most significant competition in the high end piano market.”
 - “Boston and Essex Pianos – Our mid-priced pianos compete with brands such as Bechstein, Schimmel, Kawai, and Yamaha.”

Tesla

- First company to commercially produce a federally-compliant electric vehicle
- 10-K Competition Section
 - “The worldwide automotive market, particularly for alternative fuel vehicles, is highly competitive today and we expect it will become even more so in the future...Many established and new automobile manufacturers have entered or have announced plans to enter the alternative fuel vehicle market.”
- Analyst (JPM Initiating Report 2012)
 - “We are attracted to Tesla’s highly differentiated business model: the firm uniquely owns its distribution and is saddled with none of the pension, OPEB, and other legacy costs that frequently burden large entrenched automakers.”
 - “We are attracted to Tesla’s highly differentiated products.”

Table 1 Summary Statistics

Panel A: Overall Descriptive Statistics

The table below shows the summary descriptive statistics for the main variables in the paper for the full sample from 1996 to 2013 after imposing data requirements. Variables on guidance are from 2003 onwards after imposing data requirements.

VARIABLES	N	mean	sd	p25	p50	p75
<i>txtunique</i>	49,738	-1.295	1.031	-2.079	-1.099	-0.693
<i>txtunique indicator</i>	49,738	0.246	0.431	0	0	0
<i>segunique</i>	49,738	-2.973	1.757	-4.500	-2.944	-1.609
<i>abnormal volatility</i>	49,738	0.349	1.572	-0.628	0.407	1.407
<i>abnormal volume</i>	49,738	0.314	0.756	-0.0940	0.350	0.781
<i>regular bundled guider</i>	25,123	0.432	0.496	0	0	1
<i>frequency- bundled</i>	25,123	1.881	1.864	0	1	4
<i>frequency- unbundled</i>	25,123	0.682	1.221	0	0	1
<i>frequency- preann.</i>	25,123	0.223	0.555	0	0	0
<i>guidance specificity</i>	13,262	1.722	0.899	1	1.481	2
<i>guidance horizon</i>	9,980	135.7	91.89	64	138.3	189
<i>spread</i>	49,738	1.781	2.464	0.196	0.867	2.403
<i>illiquid</i>	49,738	2.328	19.67	0.00238	0.0212	0.273
<i>analyst coverage</i>	49,738	1.448	0.951	0.693	1.386	2.197
<i>inst. ownership</i>	49,738	0.502	0.299	0.239	0.517	0.755
<i>firm size</i>	49,738	5.889	2.006	4.413	5.775	7.232
<i>stock return</i>	49,738	17.17	73.76	-24.16	5.628	38.29
<i>mb ratio</i>	49,738	3.258	4.161	1.226	2.018	3.547
<i>roa ratio</i>	49,738	0.000127	1.577	-0.0221	0.0332	0.0762
<i>stock beta</i>	49,738	1.253	0.979	0.593	1.091	1.727
<i>stock volatility</i>	49,738	0.148	0.0944	0.0841	0.124	0.184
<i>leverage</i>	49,738	0.885	2.061	0.0142	0.292	0.843

Table 1 Summary Statistics (Cont'd)**Panel B: Descriptive Statistics by Group**

The table below shows the summary descriptive statistics by groups defined by uniqueness indicator (*txtunique indicator*) for the variables on the overall information environment and basic firm characteristics for the full sample from 1996 to 2013 after imposing data requirements.

VARIABLES	<i>txtunique indicator</i> =0				<i>txtunique indicator</i> =1			
	N	mean	sd	p50	N	mean	sd	p50
<i>spread</i>	37,483	1.654	2.321	0.777	12,255	2.169	2.822	1.185
<i>illiquid</i>	37,483	1.740	16.28	0.0162	12,255	4.128	27.46	0.0584
<i>analyst coverage</i>	37,483	1.534	0.940	1.609	12,255	1.185	0.939	1.099
<i>inst ownership</i>	37,483	0.519	0.297	0.538	12,255	0.453	0.301	0.447
<i>firm size</i>	37,483	6.031	1.984	5.936	12,255	5.456	2.009	5.245
<i>stock return</i>	37,483	17.45	74.68	5.469	12,255	16.31	70.88	6.117
<i>mb ratio</i>	37,483	3.322	4.254	2.041	12,255	3.063	3.859	1.948
<i>roa ratio</i>	37,483	0.00380	1.805	0.0313	12,255	-0.0111	0.344	0.0389
<i>stock beta</i>	37,483	1.270	0.991	1.103	12,255	1.203	0.939	1.057
<i>stock volatility</i>	37,483	0.148	0.0945	0.125	12,255	0.148	0.0942	0.123
<i>leverage</i>	37,483	0.938	2.151	0.303	12,255	0.723	1.747	0.259

Panel C: Correlation Table

The table below shows the correlation table between uniqueness (*txtunique*) and market responses and major guidance variables.

	<i>txtunique</i>	<i>abnormal volatility</i>	<i>abnormal volume</i>	<i>bundled guider</i>	<i>guidance specificity</i>	<i>guidance horizon</i>
<i>txtunique</i>	1					
<i>abnormal volatility</i>	0.0571	1				
<i>abnormal volume</i>	0.0550	0.5192	1			
<i>bundled guider</i>	0.0182	0.2273	0.1860	1		
<i>guidance specificity</i>	0.1201	0.1091	0.1027	0.2531	1	
<i>guidance horizon</i>	0.0617	0.0963	0.0654	0.0256	0.2298	1

Table 2 Measure Validation

Panel A: Information Transfer from Closest Peer

(i) Correlation in Returns with Closest Peer during Peer Earnings Announcements

The table below examines the correlation in returns between the underlying firm and its closet peer identified using textual business description during the peer firm's earnings announcement. *car1d2* refers to the underlying firm's abnormal return during the [-1D, +1D] window around the earning announcement of its closest peer while *car2* refers to the peer firm's abnormal return during the [-1D, +1D] window around its earnings announcement.

	(1) <i>txtunique indicator=0</i>	(2) <i>txtunique indicator=1</i>
Correlation between <i>car2</i> and <i>car1d2</i>	0.10510	0.04378
p-value	<.0001	0.0001
# observations	53,099	18,940

(ii) Regression of Firm Return on Peer Return during Peer Earnings Announcements

The table below examines how information transfer from the closest peer is related to "informational uniqueness" of the underlying firm. Subsample 1 restricts the sample to significant information events only (with absolute abnormal returns of closet peer firms over 5% during their earnings announcements). Subsample 2 further restricts the sample to closet peer firms whose earnings announcements dates are more than 7 days apart from the underlying firm. Firm clustered t-statistics in parentheses *** p<0.01, ** p<0.05, * p<0.1

VARIABLES	(1) Full Sample <i>car1d2</i>	(2) Subsample 1 <i>car1d2</i>	(3) Subsample 2 <i>car1d2</i>
<i>car2</i>	0.080*** (15.70)	0.077*** (14.70)	0.074*** (12.54)
<i>car2 * txtunique indicator</i>	-0.046*** (-5.30)	-0.046*** (-5.15)	-0.045*** (-4.91)
<i>txtunique indicator</i>	-0.001 (-0.97)	-0.002 (-1.44)	-0.001 (-0.43)
Observations	67,120	29,416	20,650
R-squared	0.009	0.014	0.014
Industry FE	YES	YES	YES
Year FE	YES	YES	YES

Table 2 Measure Validation (Cont'd)**Panel B: Uniqueness and Analyst Coverage**

The table below examines how analyst coverage varies with measures of “information uniqueness” using both textual business description (*txtunique* and *txtunique indicator*) and business segment data (*segunique*). The first three columns use the full sample from 1996 to 2013 while Columns (4) and (5) use the post-2002 sample with guidance data. Firm and year clustered t-statistics in parentheses *** p<0.01, ** p<0.05, * p<0.1

VARIABLES (predicted sign)	(1) <i>Analyst Coverage</i>	(2) <i>Analyst Coverage</i>	(3) <i>Analyst Coverage</i>	(4) <i>Analyst Coverage</i>	(5) <i>Analyst Coverage</i>
<i>txtunique</i> (-)	-0.058*** (-6.85)			-0.087*** (-11.10)	
<i>txtunique indicator</i> (-)		-0.096*** (-7.60)			-0.118*** (-9.25)
<i>segunique</i> (-)			-0.022*** (-3.31)		
<i>inst ownership</i>	0.962*** (30.43)	0.960*** (30.08)	0.968*** (30.07)	0.770*** (15.96)	0.774*** (15.62)
<i>firm size</i>	0.284*** (33.91)	0.286*** (34.30)	0.290*** (35.05)	0.259*** (21.46)	0.264*** (21.41)
<i>stock return</i>	-0.001*** (-8.88)	-0.001*** (-8.69)	-0.001*** (-8.85)	-0.001*** (-5.65)	-0.001*** (-5.55)
<i>mb ratio</i>	0.051*** (11.82)	0.051*** (11.73)	0.050*** (11.86)	0.055*** (18.85)	0.055*** (19.53)
<i>roa ratio</i>	-0.007** (-2.56)	-0.006** (-2.46)	-0.006** (-2.47)	-0.009*** (-2.64)	-0.008** (-2.49)
<i>years of listing</i>	-0.006*** (-8.99)	-0.006*** (-9.16)	-0.006*** (-9.82)	-0.006*** (-9.66)	-0.007*** (-10.60)
<i>stock beta</i>	0.089*** (6.37)	0.089*** (6.40)	0.085*** (6.30)	0.067*** (3.38)	0.067*** (3.46)
<i>stock volatility</i>	0.647*** (3.62)	0.656*** (3.69)	0.640*** (3.66)	0.584* (1.95)	0.609** (2.05)
<i>leverage</i>	-0.088*** (-17.21)	-0.088*** (-17.28)	-0.086*** (-17.28)	-0.086*** (-13.81)	-0.087*** (-14.15)
<i>frequency - bundled</i>				0.075*** (9.94)	0.070*** (8.97)
Observations	49,738	49,738	49,738	29,839	29,839
R-squared	0.599	0.597	0.597	0.622	0.617
Year FE	YES	YES	YES	YES	YES

Table 3 Market Reactions during Earnings Announcements**Panel A: Uniqueness and Market Reactions**

The table below examines how market reactions in terms of abnormal return volatility and abnormal trading volume vary with measures of “information uniqueness” (*txtunique* and *txtunique indicator*) during annual earnings announcements, controlling for guidance policy as well as other standard firm characteristics. Firm and year clustered t-statistics in parentheses *** p<0.01, ** p<0.05, * p<0.1

VARIABLES (predicted sign)	(1) <i>Abnormal Volatility</i>	(2) <i>Abnormal Volatility</i>	(3) <i>Abnormal Volume</i>	(4) <i>Abnormal Volume</i>
<i>txtunique</i> (+)	0.163*** (11.06)		0.075*** (13.27)	
<i>txtunique indicator</i> (+)		0.200*** (8.09)		0.090*** (10.09)
<i>frequency - bundled</i>	0.094*** (10.18)	0.104*** (10.98)	0.039*** (15.16)	0.044*** (15.81)
<i>frequency - unbundled</i>	-0.047*** (-8.60)	-0.051*** (-9.21)	-0.011*** (-4.40)	-0.013*** (-5.04)
<i>frequency- preann.</i>	-0.053** (-2.47)	-0.047** (-2.20)	-0.003 (-0.46)	-0.000 (-0.03)
<i>inst ownership</i>	0.571*** (11.53)	0.583*** (11.76)	0.253*** (11.32)	0.258*** (11.34)
<i>analyst coverage</i>	0.226*** (14.38)	0.200*** (12.89)	0.124*** (16.95)	0.112*** (15.07)
<i>firm size</i>	0.018** (2.39)	0.013* (1.69)	-0.013** (-2.05)	-0.015** (-2.37)
<i>stock return</i>	0.001*** (3.68)	0.001*** (3.55)	0.001*** (3.91)	0.001*** (3.82)
<i>mb ratio</i>	-0.005 (-1.34)	-0.004 (-1.22)	-0.003 (-1.39)	-0.003 (-1.33)
<i>roa ratio</i>	0.003 (0.62)	0.002 (0.46)	0.002 (0.67)	0.002 (0.56)
<i>years of listing</i>	-0.001 (-0.65)	0.001 (0.66)	-0.001 (-1.48)	0.000 (0.01)
<i>stock beta</i>	0.046** (2.52)	0.047** (2.46)	0.035*** (3.81)	0.035*** (3.71)
<i>stock volatility</i>	0.453** (2.48)	0.417** (2.14)	0.322** (2.27)	0.304** (2.05)
<i>leverage</i>	-0.021*** (-4.10)	-0.022*** (-4.17)	-0.006*** (-2.86)	-0.007*** (-3.16)
Observations	29,839	29,839	29,839	29,839
R-squared	0.102	0.095	0.110	0.102
Year FE	YES	YES	YES	YES

Table 4 Corporate Disclosure

Panel A: Uniqueness and Disclosure Policy – Variation across Full Sample

The table below examines how the disclosure policy of firms varies in cross-section with “information uniqueness” (*txtunique*). The first two columns utilize the full post-2002 sample for firms with analyst coverage with column (1) using the frequency of bundled guidance and column (2) using an indicator for regular bundled guiders as dependent variables. The last two columns split the sample into high growth firms and low growth firms by the median market-to-book (*MB*) ratio. The difference column shows the chi-square statistics of the difference in coefficients between columns (3) and (4). Firm and year clustered t-statistics in parentheses *** p<0.01, ** p<0.05, * p<0.1

	<u>Full Sample</u>	<u>Full Sample</u>	<u>High MB Sample</u>	<u>Low MB Sample</u>	
VARIABLES (predicted sign)	(1) <i>Frequency of Bundled Guidance</i>	(2) <i>Regular Bundled Guider</i>	(3) <i>Regular Bundled Guider</i>	(4) <i>Regular Bundled Guider</i>	<i>Diff. F-Stat.</i>
<i>txtunique</i> (+)	0.213*** (9.41)	0.050*** (8.35)	0.060*** (8.49)	0.026*** (3.56)	15.88 ***
<i>inst ownership</i>	1.654*** (10.99)	0.393*** (10.47)	0.350*** (9.49)	0.409*** (9.36)	2.36
<i>analyst coverage</i>	0.599*** (10.06)	0.155*** (11.24)	0.144*** (9.52)	0.124*** (7.30)	1.50
<i>years of listing</i>	-0.001 (-0.54)	-0.000 (-0.47)	-0.001* (-1.84)	-0.000 (-0.38)	1.41
<i>firm size</i>	-0.035 (-1.42)	-0.010* (-1.65)	0.007 (1.07)	-0.016** (-2.08)	7.86 ***
<i>stock return</i>	0.001*** (4.89)	0.000*** (4.53)	0.000*** (3.06)	0.000 (0.45)	5.45 **
<i>roa ratio</i>	0.010 (0.76)	0.003 (0.84)	0.002 (1.06)	0.108*** (4.60)	28.77 ***
<i>stock beta</i>	0.062** (2.48)	0.012* (1.85)	0.006 (0.84)	0.022** (2.07)	3.57 *
<i>mb ratio</i>	-0.003 (-0.36)	-0.001 (-0.40)	-0.005** (-2.48)	0.105*** (5.42)	61.31 ***
<i>stock volatility</i>	-1.330*** (-3.34)	-0.284*** (-2.71)	-0.403*** (-3.26)	0.084 (0.74)	20.23 ***
<i>leverage</i>	-0.005 (-0.28)	-0.002 (-0.43)	0.005 (1.06)	-0.016*** (-4.90)	26.55 ***
Observations	25,123	25,123	12,562	12,561	
R-squared	0.228	0.192	0.214	0.181	
Year FE	YES	YES	YES	YES	

Table 4 Corporate Disclosure (Cont'd)**Panel B: Uniqueness and Disclosure Policy – Variation within Guiding Firms**

The table below examines how the disclosure policy of firms varies with “information uniqueness” (*txtunique*) conditional on providing management guidance in the year. The first two columns use the frequency of bundled guidance and unbundled guidance respectively as dependent variables and the last two columns use measures of earnings guidance horizon and guidance specificity as dependent variables. Firm and year clustered t-statistics in parentheses
 *** p<0.01, ** p<0.05, * p<0.1

VARIABLES (predicted sign)	(1) <i>Frequency of Bundled Guidance</i>	(2) <i>Frequency of Unbundled Guidance</i>	(3) <i>Guidance Horizon</i>	(4) <i>Guidance Specificity</i>
<i>txtunique</i>	(+) 0.065*** (3.32)	(+/-) -0.081*** (-4.06)	(+) 6.657*** (3.61)	(+) 0.144*** (5.74)
<i>inst ownership</i>	1.173*** (8.72)	0.067 (0.89)	0.298 (0.05)	0.142** (2.35)
<i>analyst coverage</i>	0.356*** (7.64)	0.345*** (8.31)	-10.715*** (-4.24)	0.100*** (3.82)
<i>years of listing</i>	0.000 (0.12)	0.006*** (3.49)	-0.112 (-0.82)	-0.005*** (-4.53)
<i>firm size</i>	-0.039 (-1.62)	0.074*** (3.43)	10.041*** (6.87)	-0.038*** (-3.57)
<i>stock return</i>	0.002*** (6.81)	0.001*** (4.17)	-16.660 (-1.40)	0.272*** (3.35)
<i>roa ratio</i>	0.979*** (6.16)	0.279** (2.34)	0.054** (1.99)	0.000 (0.14)
<i>stock beta</i>	-0.001 (-0.04)	-0.021 (-1.36)	-13.676*** (-5.89)	-0.076*** (-2.90)
<i>mb ratio</i>	0.004 (0.40)	0.006 (1.15)	2.462*** (5.11)	0.009** (2.41)
<i>stock volatility</i>	-1.729*** (-3.97)	-0.042 (-0.12)	-75.469** (-2.47)	-0.027 (-0.09)
<i>leverage</i>	0.006 (0.39)	-0.011 (-1.00)	-0.600 (-0.77)	-0.004 (-0.73)
Observations	15,124	15,124	9,980	13,262
R-squared	0.157	0.098	0.076	0.073
Year FE	YES	YES	YES	YES

Table 5 Information Environment**Pane A: Uniqueness and Information Environment**

The table below examines how the quality of the overall information environment varies with “information uniqueness” (*txtunique*) and guidance choices. The first two columns use the bid-ask spread (*spread*) while the latter two columns use Amihud’s measure of illiquidity (*illiquid*) to proxy for an inverse relationship with the quality of the overall information environment. Firm and year clustered t-statistics in parentheses *** p<0.01, ** p<0.05, * p<0.1

VARIABLES (predicted sign)	(1) <i>Spread</i>	(2) <i>Spread</i>	(3) <i>Illiquid</i>	(4) <i>Illiquid</i>
<i>txtunique</i> (+)	0.086*** (5.91)	0.126*** (6.25)	0.718*** (3.60)	1.061*** (3.21)
<i>txtunique</i> * <i>bundled guider</i> (-)		-0.125*** (-5.47)		-1.079*** (-2.99)
<i>bundled guider</i> (-)		-0.246*** (-4.78)		-1.991*** (-3.04)
<i>frequency – bundled</i> (-)	-0.042*** (-3.47)		-0.308*** (-2.82)	
<i>frequency – unbundled</i> (+/-)	0.045*** (3.39)		0.379** (2.44)	
<i>analyst coverage</i>	-0.196*** (-6.84)	-0.192*** (-6.75)	-0.435 (-1.53)	-0.389 (-1.51)
<i>inst. ownership</i>	-1.216*** (-7.95)	-1.247*** (-7.89)	-6.058*** (-3.35)	-6.283*** (-3.34)
<i>firm size</i>	-0.187*** (-8.55)	-0.181*** (-8.46)	-1.044*** (-3.79)	-0.997*** (-3.81)
<i>stock return</i>	-0.000 (-0.50)	-0.000 (-0.51)	-0.009** (-2.14)	-0.009** (-2.15)
<i>mb ratio</i>	-0.048*** (-5.81)	-0.047*** (-5.74)	-0.303*** (-4.17)	-0.295*** (-4.21)
<i>roa ratio</i>	-0.012** (-2.03)	-0.011** (-1.98)	-0.047 (-1.14)	-0.045 (-1.10)
<i>stock beta</i>	-0.173*** (-5.63)	-0.173*** (-5.68)	-1.348*** (-2.73)	-1.346*** (-2.74)
<i>stock volatility</i>	2.333*** (4.22)	2.348*** (4.26)	13.562* (1.89)	13.666* (1.91)
<i>leverage</i>	0.063*** (5.09)	0.062*** (5.02)	0.310*** (4.27)	0.305*** (4.19)
Observations	29,839	29,839	29,839	29,839
R-squared	0.445	0.445	0.045	0.045
Year FE	YES	YES	YES	YES

Table 5 Information Environment (Cont'd)**Panel B: Subsampling on the Presence of Regular Bundled Guidance Policy**

The table below examines how the quality of the overall information environment of “informationally unique” firms (*txtunique indicator*) varies in subsamples of different guidance policy. Columns (1) and (3) examine a subsample of firms without a regular bundled guidance policy while columns (2) and (4) examine a subsample of firms with a regular bundled guidance policy. The dependent variables are the bid-ask spread (*spread*) in columns (1) and (2) and Amihud’s measure of illiquidity (*illiquid*) in columns (3) and (4) respectively. The difference columns show the chi-square statistics of the difference in coefficients in the prior two columns. Firm and year clustered t-statistics in parentheses *** p<0.01, ** p<0.05, * p<0.1

	Sample w/o Regular Bundled Guidance	Sample w/ Regular Bundled Guidance		Sample w/o Regular Bundled Guidance	Sample w/ Regular Bundled Guidance	
VARIABLES (predicted sign)	(1) <i>Spread</i>	(2) <i>Spread</i>	Diff. F- Stat.	(3) <i>Illiquid</i>	(4) <i>Illiquid</i>	Diff. F-Stat.
<i>txtunique indicator</i> (+)	0.210*** (5.45)	0.009 (0.70)	29.44 ***	2.017*** (2.58)	0.010 (0.14)	7.08 ***
<i>analyst coverage</i>	-0.239*** (-6.61)	-0.104*** (-5.45)	36.72 ***	-0.538 (-1.41)	-0.067** (-2.58)	4.06 **
<i>inst. ownership</i>	-1.234*** (-7.58)	-0.559*** (-7.67)	88.83 ***	-6.484*** (-3.01)	-0.809*** (-3.44)	28.43 ***
<i>firm size</i>	-0.241*** (-8.07)	-0.052*** (-7.68)	208.4 0 ***	-1.409*** (-3.63)	-0.081*** (-2.92)	67.78 ***
<i>stock return</i>	-0.000 (-0.87)	0.001*** (4.55)	9.79 ***	-0.011* (-1.96)	0.001** (2.36)	12.34 ***
<i>mb ratio</i>	-0.066*** (-5.80)	-0.010*** (-5.01)	158.7 3 ***	-0.441*** (-3.86)	-0.022*** (-3.15)	67.02 ***
<i>roa ratio</i>	-0.009** (-1.98)	-0.398*** (-5.29)	53.00 ***	-0.029 (-0.87)	-0.220 (-1.48)	1.54
<i>stock beta</i>	-0.197*** (-5.57)	-0.055*** (-3.41)	72.90 ***	-1.692*** (-2.59)	-0.047** (-2.21)	23.36 ***
<i>stock volatility</i>	2.774*** (4.10)	0.800*** (4.01)	158.7 3 ***	18.271** (2.01)	-0.464 (-0.86)	9.86 ***
<i>leverage</i>	0.088*** (5.54)	0.012*** (3.52)	79.34 ***	0.461*** (4.54)	0.024** (1.98)	25.77 ***
Observations	18,818	11,021		18,818	11,021	
R-squared	0.436	0.370		0.051	0.026	
Year FE	YES	YES		YES	YES	

Table 6 Alternative Measures of Informational Uniqueness**Panel A: Main Tests Using Industry-Level and Industry-Adjusted Uniqueness**

The table below examines the main test specifications from Tables 3 to 5 using abnormal volatility during earnings announcements in column (1), the frequency of bundled guidance in column (2), the bid-ask spread in column (3) and Amihud's measure of illiquidity in column (4), where uniqueness measure (*txtunique*) is split into industry-level uniqueness and industry-adjusted uniqueness based on Fama-French 48 industry classification. Firm and year clustered t-statistics in parentheses *** p<0.01, ** p<0.05, * p<0.1

VARIABLES (predicted sign)	(1) <i>Abnormal Volatility</i>	(2) <i>Frequency of Bundled Guidance</i>	(3) <i>Spread</i>	(4) <i>Illiquid</i>
<i>industry-level uniqueness</i> (+)	0.387*** (9.83)	0.548*** (12.53)	0.062*** (2.71)	0.864* (1.95)
<i>industry-adjusted uniqueness</i> (+)	0.076*** (6.36)	0.037* (1.79)	0.076*** (4.55)	0.491*** (3.49)
<i>inst. ownership</i>	0.510*** (10.03)	1.483*** (10.62)	-1.190*** (-7.86)	-0.304** (-2.52)
<i>analyst coverage</i>	0.220*** (14.38)	0.564*** (10.66)	-0.176*** (-6.65)	0.332** (2.29)
<i>frequency - bundled</i>	0.079*** (8.99)		-0.039*** (-3.08)	-5.911*** (-3.26)
<i>frequency - unbundled</i>	-0.060*** (-6.91)		0.040*** (3.08)	-0.272 (-1.05)
<i>firm size</i>	0.039*** (5.46)	0.019 (1.04)	0.006*** (5.47)	0.053*** (2.86)
<i>stock return</i>	0.001*** (3.40)	0.001*** (6.09)	-0.209*** (-8.31)	-1.215*** (-3.84)
<i>mb ratio</i>	-0.001 (-0.43)	0.006 (1.02)	-0.000 (-0.34)	-0.009** (-2.12)
<i>roa ratio</i>	0.003 (0.67)	0.003 (0.67)	-0.012** (-2.00)	-0.051 (-1.16)
<i>years of listing</i>	-0.001 (-1.24)	-0.003 (-1.59)	-0.169*** (-5.74)	-1.325*** (-2.73)
<i>stock beta</i>	0.033* (1.92)	0.032 (1.45)	-0.050*** (-5.82)	-0.319*** (-4.20)
<i>stock volatility</i>	0.475*** (2.87)	-0.943*** (-3.68)	2.425*** (4.43)	14.420** (2.03)
<i>leverage</i>	-0.023*** (-4.24)	-0.007 (-0.47)	0.068*** (5.46)	0.353*** (4.81)
Observations	29,839	29,839	29,839	29,839
R-squared	0.109	0.322	0.448	0.046
Year FE	YES	YES	YES	YES

Table 6 Alternative Measures of Informational Uniqueness (Cont'd)**Panel B: Main Tests Using Alternative Measure Based on Business Segment Data**

The table below examines the main test specifications from Tables 3 to 5 using abnormal volatility during earnings announcements in column (1), the frequency of bundled guidance in column (2), bid-ask spread in column (3) and Amihud's measure of illiquidity in column (4), where uniqueness measure (*segunique*) is constructed based on comparing weightings on business segments across firms. Firm and year clustered t-statistics in parentheses *** p<0.01, ** p<0.05, * p<0.1

VARIABLES (predicted sign)	(1) <i>Abnormal Volatility</i>	(2) <i>Frequency of Bundled Guidance</i>	(3) <i>Spread</i>	(4) <i>Illiquid</i>
<i>segunique</i> (+)	0.054*** (6.27)	0.024** (2.04)	0.020*** (2.70)	0.159* (1.87)
<i>inst. ownership</i>	0.553*** (11.04)	1.633*** (11.88)	-1.196*** (-7.81)	-5.867*** (-3.34)
<i>analyst coverage</i>	0.205*** (12.76)	0.560*** (9.32)	-0.187*** (-6.96)	-0.355 (-1.36)
<i>frequency - bundled</i>	0.103*** (10.98)		-0.035*** (-2.78)	-0.247** (-2.47)
<i>frequency - unbundled</i>	-0.062*** (-7.33)		0.039*** (2.97)	0.322** (2.25)
<i>firm size</i>	0.006 (0.76)	-0.026 (-1.21)	-0.214*** (-8.33)	-1.287*** (-3.76)
<i>stock return</i>	0.001*** (3.46)	0.001*** (5.27)	-0.000 (-0.42)	-0.009** (-2.14)
<i>mb ratio</i>	-0.002 (-0.59)	0.002 (0.37)	-0.049*** (-5.84)	-0.317*** (-4.10)
<i>roa ratio</i>	0.001 (0.28)	0.001 (0.19)	-0.013** (-2.08)	-0.057 (-1.26)
<i>years of listing</i>	0.000 (0.19)	-0.001 (-0.28)	0.006*** (5.71)	0.057*** (2.89)
<i>stock beta</i>	0.054*** (2.73)	0.059*** (2.89)	-0.166*** (-5.90)	-1.283*** (-2.70)
<i>stock volatility</i>	0.401* (1.91)	-1.097*** (-3.98)	2.402*** (4.40)	14.190** (2.00)
<i>leverage</i>	-0.026*** (-4.79)	-0.007 (-0.50)	0.066*** (5.47)	0.342*** (4.72)
Observations	29,839	29,839	29,839	29,839
R-squared	0.095	0.298	0.446	0.046
Year FE	YES	YES	YES	YES

Table 6 Alternative Measures of Informational Uniqueness (Cont'd)**Panel C: Main Tests Using IV on Informational Uniqueness**

The table below examines the main test specifications from Tables 3 to 5 using abnormal volatility during earnings announcements in column (1), the frequency of bundled guidance in column (2), bid-ask spread in column (3) and Amihud's measure of illiquidity in column (4), where uniqueness measure (*txtunique_instrument*) is the predicted value of informational uniqueness from the first-stage regression using the proportion of private firms in the firm's primary NAICS 6-digit industry as instrument variable. Firm and year clustered t-statistics in parentheses *** p<0.01, ** p<0.05, * p<0.1

VARIABLES (predicted sign)	(1) <i>Abnormal Volatility</i>	(2) <i>Frequency of Bundled Guidance</i>	(3) <i>Spread</i>	(4) <i>Illiquid</i>
<i>txtunique_instrument</i> (+)	0.560*** (6.56)	0.793*** (5.82)	0.298*** (4.16)	3.094*** (2.82)
<i>inst. ownership</i>	0.547*** (9.34)	1.470*** (10.43)	-1.248*** (-8.20)	-6.433*** (-3.44)
<i>analyst coverage</i>	0.300*** (12.01)	0.730*** (12.47)	-0.125*** (-4.03)	0.369 (1.51)
<i>frequency - bundled</i>	0.063*** (5.55)		-0.060*** (-4.18)	-0.543*** (-3.45)
<i>frequency - unbundled</i>	-0.055*** (-5.24)		0.046*** (3.08)	0.418** (2.37)
<i>firm size</i>	0.051*** (5.30)	0.019 (0.90)	-0.191*** (-8.28)	-1.128*** (-3.71)
<i>stock return</i>	0.001*** (3.91)	0.001*** (4.21)	-0.000 (-0.07)	-0.009** (-2.14)
<i>mb ratio</i>	-0.001 (-0.19)	0.004 (0.72)	-0.050*** (-6.20)	-0.327*** (-4.33)
<i>roa ratio</i>	0.005* (1.66)	0.007* (1.79)	-0.010* (-1.73)	-0.034 (-0.81)
<i>years of listing</i>	-0.005*** (-4.17)	-0.008*** (-3.74)	0.003*** (2.64)	0.027 (1.64)
<i>stock beta</i>	0.049*** (3.01)	0.056** (2.20)	-0.180*** (-5.40)	-1.386*** (-2.76)
<i>stock volatility</i>	0.598*** (4.11)	-0.431 (-1.63)	2.400*** (5.25)	14.985** (2.02)
<i>leverage</i>	-0.022*** (-3.42)	-0.004 (-0.27)	0.075*** (5.14)	0.400*** (4.54)
Observations	26,813	26,813	26,813	26,813
1 st Stage K-P Wald F statistic	74.58	74.58	74.58	74.58
R-squared	0.041	0.213	0.427	0.037
Year FE	YES	YES	YES	YES

Table 7 Controlling for Competition Perspective**Panel A: Main Tests Controlling for Total Number of U.S. Public and Private Firms in the Industry**

The table below examines the main test specifications from Tables 3 to 5 using abnormal volatility during earnings announcements in column (1), the frequency of bundled guidance in column (2), the bid-ask spread in column (3) and Amihud's measure of illiquidity in column (4), where the logarithm of the total number of U.S. public and private firms operating in the same primary 6-digit NAICS code (*totindpeersize*) is added alongside my main measure of informational uniqueness (*txtunique*) as an additional control. Firm and year clustered t-statistics in parentheses *** p<0.01, ** p<0.05, * p<0.1

VARIABLES (predicted sign)	(1) <i>Abnormal Volatility</i>	(2) <i>Frequency of Bundled Guidance</i>	(3) <i>Spread</i>	(4) <i>Illiquid</i>
<i>txtunique</i> (+)	0.158*** (10.21)	0.206*** (8.94)	0.068*** (5.48)	0.653*** (3.33)
<i>totindpeersize</i>	-0.000 (-0.01)	0.019* (1.66)	0.017** (2.51)	0.195** (2.41)
<i>inst. ownership</i>	0.584*** (10.90)	1.595*** (10.18)	-1.232*** (-7.82)	-6.273*** (-3.31)
<i>analyst coverage</i>	0.201*** (12.45)	0.616*** (12.89)	-0.179*** (-6.34)	-0.199 (-0.71)
<i>frequency - bundled</i>	0.098*** (9.47)		-0.041*** (-3.28)	-0.337*** (-3.03)
<i>frequency - unbundled</i>	-0.063*** (-6.95)		0.041*** (3.14)	0.372** (2.39)
<i>firm size</i>	0.026*** (3.53)	-0.022 (-1.12)	-0.206*** (-8.32)	-1.288*** (-3.66)
<i>stock return</i>	0.001*** (3.50)	0.001*** (4.80)	0.000 (0.19)	-0.009** (-2.04)
<i>mb ratio</i>	-0.003 (-0.76)	0.001 (0.18)	-0.051*** (-6.03)	-0.336*** (-4.12)
<i>roa ratio</i>	0.002 (0.52)	0.002 (0.58)	-0.012** (-2.02)	-0.053 (-1.17)
<i>years of listing</i>	-0.001 (-0.81)	-0.001 (-0.41)	0.006*** (5.13)	0.057*** (2.74)
<i>stock beta</i>	0.049*** (2.80)	0.059*** (2.68)	-0.178*** (-5.51)	-1.371*** (-2.66)
<i>stock volatility</i>	0.333* (1.94)	-0.858*** (-3.03)	2.256*** (4.82)	13.551* (1.82)
<i>leverage</i>	-0.023*** (-4.05)	-0.005 (-0.35)	0.074*** (4.91)	0.391*** (4.21)

Observations	26,791	26,791	26,791	26,791
R-squared	0.101	0.310	0.450	0.047
Year FE	YES	YES	YES	YES

Table 7 Controlling for Competition Perspective (Cont'd)**Panel B: Main Tests Controlling for Herfindahl Index Computed from US Census Data**

The table below examines the main test specifications from Tables 3 to 5 using abnormal volatility during earnings announcements in column (1), the frequency of bundled guidance in column (2), the bid-ask spread in column (3) and Amihud's measure of illiquidity in column (4), where industry concentration ratio represented by the Herfindahl-Hirschman index for the 50 largest companies from the US 2007 *Economic Census* data is added alongside my main measure of informational uniqueness (*txtunique*) as an additional control. The sample is restricted to US manufacturing firms only due to data availability on industry concentration ratio. Firm and year clustered t-statistics in parentheses *** p<0.01, ** p<0.05, * p<0.1

VARIABLES (predicted sign)	(1) <i>Abnormal Volatility</i>	(2) <i>Frequency of Bundled Guidance</i>	(3) <i>Spread</i>	(4) <i>Illiquid</i>
<i>txtunique</i> (+)	0.131*** (5.39)	0.144*** (4.72)	0.111*** (4.70)	0.788** (2.47)
<i>US Census herfindahl</i>	-0.000 (-0.21)	0.000*** (2.91)	-0.000 (-0.83)	0.000 (0.45)
<i>inst. ownership</i>	0.549*** (7.22)	1.338*** (7.93)	-1.386*** (-6.77)	-6.324*** (-3.30)
<i>analyst coverage</i>	0.159*** (6.08)	0.534*** (10.54)	-0.154*** (-4.57)	-0.048 (-0.11)
<i>frequency - bundled</i>	0.105*** (9.62)		-0.027* (-1.70)	-0.217* (-1.77)
<i>frequency - unbundled</i>	-0.051*** (-3.59)		0.062*** (4.68)	0.319** (2.14)
<i>firm size</i>	0.021 (1.52)	0.114*** (3.93)	-0.248*** (-7.69)	-1.430*** (-4.04)
<i>stock return</i>	0.001** (2.03)	0.001*** (3.45)	0.000* (1.89)	-0.005 (-1.41)
<i>mb ratio</i>	-0.004 (-0.55)	-0.010 (-0.99)	-0.067*** (-5.94)	-0.417*** (-3.50)
<i>roa ratio</i>	0.626*** (7.56)	0.339*** (4.23)	-0.239*** (-3.67)	-2.096 (-1.00)
<i>years of listing</i>	-0.001 (-1.03)	-0.007*** (-2.80)	0.007*** (4.13)	0.055** (2.37)
<i>stock beta</i>	0.069*** (2.96)	0.035 (1.15)	-0.165*** (-4.55)	-0.838** (-2.41)
<i>stock volatility</i>	0.523** (2.41)	-1.053*** (-3.09)	1.694*** (2.60)	1.574 (0.27)
<i>leverage</i>	-0.020 (-1.51)	-0.000 (-0.01)	0.131*** (4.64)	0.754*** (5.15)
Observations	12,176	12,212	12,206	12,182

R-squared	0.120	0.382	0.452	0.044
Year FE	YES	YES	YES	YES

Table 8 Additional Controls for Robustness Checks**Panel A: Main Tests with Additional Firm Characteristics as Controls**

The table below examines the main test specifications from Tables 3 to 5 using abnormal volatility during earnings announcements in column (1), the frequency of bundled guidance in column (2), the bid-ask spread in column (3) and Amihud's measure of illiquidity in column (4), where additional controls, such as indicator for loss firms, asset turnover, FPS (1994) high-litigation risk industries, counts of words in 10-K filings to describe protected sources of competitive advantage (e.g. patents), industry concentration, number of industrial segment, number of geographical segments, Franco et. al. (2011) accounting comparability, are included in addition to standard firm controls in prior tables. Firm and year clustered t-statistics in parentheses *** p<0.01, ** p<0.05, * p<0.1

VARIABLES (predicted sign)	(1) <i>Abnormal Volatility</i>	(2) <i>Frequency of Bundled Guidance</i>	(3) <i>Spread</i>	(4) <i>Illiquid</i>
<i>txtunique indicator(+)</i>	0.138*** (5.80)	0.137** (2.42)	0.201*** (3.67)	2.014** (2.24)
<i>txtunique indicator* bundled guider (-)</i>			-0.248*** (-4.06)	-2.274** (-2.38)
<i>bundled guider (-)</i>			-0.003 (-0.10)	0.181 (0.67)
<i>Standard Firm Controls</i>	INCLUDED	INCLUDED	INCLUDED	INCLUDED
<u>Additional Controls</u>				
<i>loss firms</i>	-0.211*** (-6.06)	-0.177** (-2.29)	0.118*** (4.24)	1.154** (2.19)
<i>asset turnover</i>	0.125*** (6.42)	0.011 (0.24)	0.107*** (3.23)	0.697** (2.55)
<i>FPS litigation industries</i>	0.188*** (4.33)	0.392*** (5.22)	0.066** (2.26)	0.618 (1.50)
<i>word counts about protected advantage</i>	0.000 (0.10)	-0.002** (-2.34)	-0.001** (-2.34)	0.001 (0.24)
<i>Compustat industry concentration</i>	0.398*** (4.93)	0.127 (0.80)	0.010 (0.15)	-0.174 (-0.33)
<i># industry segments</i>	0.008 (0.68)	0.048* (1.92)	-0.001 (-0.13)	0.024 (0.44)
<i># geo. segments</i>	0.024** (2.23)	0.062*** (4.22)	0.004 (0.71)	0.187 (0.80)
<i>accounting comparability</i>	0.016 (1.43)	0.098*** (3.20)	-0.108*** (-2.89)	-0.649** (-2.11)
Observations	13,879	12,212	13,879	13,879

R-squared	0.085	0.229	0.447	0.028
Year FE	YES	YES	YES	YES

Table 9 Time Series Variation**Panel A: Decreases in Uniqueness due to IPOs of Close Peers**

The table shows the changes in announcement volatility around annual earnings announcement, the presence of regular bundled guidance policy, the bid-ask spread and Amihud's measure of illiquidity (*illiquid*) for a subsample of 44 firms that experience a decrease in information uniqueness after the IPOs of their close peers. *Postevent* is an indicator variable for years 0, 1 and 2 relative to their peer IPO date. Samples only include at most 3 years before and after the event depending on data availability. Sample of columns (2) is post-2002 due to data requirement on guidance. Firm and year clustered t-statistics in parentheses *** p<0.01, ** p<0.05, * p<0.1

VARIABLES (predicted sign)	(1) <i>Announcement Volatility</i>	(2) <i>Regular Bundled Guidance</i>	(3) <i>Spread</i>	(4) <i>Illiquid</i>
<i>postevent</i> (-ve)	-0.004 (-0.47)	-0.023 (-0.26)	-1.407*** (-3.90)	-2.186* (-1.67)
Observations	232	116	233	233
R-squared	0.206	0.578	0.624	0.294
Firm FE	YES	YES	YES	YES

Panel B: Difference-in-Difference with Matched Firms as Controls

The table shows the difference-in-difference in announcement volatility around annual earnings announcement, the presence of regular bundled guidance policy, the bid-ask spread and Amihud's measure of illiquidity (*illiquid*) for a subsample of 44 firms that experience a decrease in information uniqueness after the IPOs of their close peers relative to their size-matched firms with similar prior uniqueness, analyst coverage and profitability. *Postevent* is an indicator variable for years 0, 1 and 2 relative to their peer IPO date. *Treated* is an indicator variable for firms that experience a decrease in information uniqueness due to peer IPOs. Samples only include at most 3 years before and after the event depending on data availability. Sample of columns (2) is post-2002 due to data requirement on guidance. Firm and year clustered t-statistics in parentheses *** p<0.01, ** p<0.05, * p<0.1

VARIABLES (predicted sign)	(1) <i>Announcement Volatility</i>	(2) <i>Regular Bundled Guidance</i>	(3) <i>Spread</i>	(4) <i>Illiquid</i>
<i>postevent*treated</i> (-ve)	-0.010 (-0.86)	-0.030 (-0.37)	-0.903** (-2.13)	-2.199* (-1.72)
<i>postevent</i>	0.008 (0.71)	0.121 (1.15)	-0.034* (0.16)	0.967 (0.84)
Observations	487	236	488	487
R-squared	0.305	0.760	0.610	0.401
Year FE	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES

Figure 1 Uniqueness and Information Environment

Part A: Bid-Ask Spread

Figure 1A below shows the plot of bid-ask spread across different quartiles of uniqueness (*txtunique*). Dotted columns represent the average spread in each quartile of uniqueness without controls. Solid columns represent the estimated spread in each quartile of uniqueness in the model with standard firm controls assuming other firm characteristics being the same as sample median.

Figure 1 Uniqueness and Information Environment (Cont'd)

Part B: Bid-Ask Spread

Figure 1B below shows the plot of bid-ask spread across different quartiles of uniqueness (*txtunique*) conditional on the presence of regular bundled guidance or not. Solid columns are from samples of firms without regular bundled guidance while dotted columns are from samples of firms with regular bundled guidance. Spread in each quartile of uniqueness within each subsample is estimated using the model with standard firm controls assuming other firm characteristics being the same as sample median.

Figure 1 Uniqueness and Information Environment (Cont'd)

Part C: Amihud's measure of illiquidity (2002)

Figure 1C below shows the plot of Amihud's measure of illiquidity (2002) across different quartiles of uniqueness (*txtunique*). Dotted columns represent the average illiquidity in each quartile of uniqueness without controls. Solid columns represent the estimated illiquidity in each quartile of uniqueness in the model with standard firm controls assuming other firm characteristics being the same as sample median.

Figure 1 Uniqueness and Information Environment (Cont'd)

Part D: Amihud's measure of illiquidity (2002)

Figure 1D below shows the plot of Amihud's measure of illiquidity (2002) across different quartiles of uniqueness (*txtunique*) conditional on the presence of regular bundled guidance. Solid columns are from samples of firms without regular bundled guidance while dotted columns are from firms with regular bundled guidance. Illiquidity in each uniqueness quartile within each subsample is estimated using the model with standard firm controls assuming other firm characteristics being the same as sample median.

CHAPTER 2

Peer Dynamics and Discretionary Disclosure

1. Introduction

In this paper, I investigate how multi-firm dynamics influence discretionary disclosure behaviors²⁹. While the presence of comparable public peers should increase investors' overall information endowment on the underlying firm over their common exposures, it also provides investors expanded knowledge about the timing of information arrival for managers. Although prior empirical research has demonstrated the presence of within-industry information transfers from earnings announcements (e.g. Foster 1981) and management forecasts (e.g. Baginski 1987), it is unclear if investors can always perfectly understand the implications of peer disclosures on the underlying firm. If investors can see through the implications, it is possible that the underlying firm can free-ride on the information transfer from peers, therefore negating the need to provide additional discretionary disclosures if they are costly. However, if investors are unable to determine whether the absence of disclosure is due to the absence of new information or strategic information withholding, discretionary disclosures by peer firms may signal the possibility of information arrival to the underlying firm, raising investors' uncertainty and assessed likelihood that non-disclosure is due to deliberate information withholding rather than the absence of new information. Whether such peer dynamics suppress or

²⁹ In this paper, I focus on examining *ex-post* discretionary disclosure behaviors by firms upon signal arrival, instead of *ex-ante* commitment to disclosure.

induce more discretionary disclosures remain largely unanswered by prior empirical literature.

Prior theoretical literature on discretionary disclosure using a single-firm setting has generally shown that a firm's propensity to provide discretionary disclosure increases as disclosure-related cost decreases (e.g. Verrecchia 1983) or as investors' assessed probability that the firm receives information increases (e.g. Dye 1985, Jung and Kwon 1988). The two most closely related papers using a multi-firm setting include Dye and Sridhar (1995) and Jorgensen and Kirschenheiter (2012). Dye and Sridhar (1995) generally predicts "herding" behavior in discretionary disclosure as long as there is positive correlation among firms' receipt of information. It suggests that a firm's equilibrium disclosure policy depends on whether the information the firm receives is "private knowledge" or "industry-wide common knowledge". Jorgensen and Kirschenheiter (2012) consider a sequential discretionary disclosure setting. It predicts a higher disclosure threshold by the follower due to information free-riding if the signals are positively correlated, and a lower disclosure threshold by the follower due to negative information externalities if the signals are negatively correlated. Despite these theoretical works, it remains to be empirically tested whether some of these conditions are more prevalent for firms in general.

Within groups of comparable peers, I expect that both the information arrival processes and signals should be positively correlated on average as prior empirical studies generally find a positive coefficient on within-industry information transfer. While positive correlation in information arrival processes generally predicts a lower disclosure threshold (Dye and Sridhar 1995), positive correlation in information signals

may predict a higher disclosure threshold for the follower due to the possibility of information free-riding if disclosure is costly (Jorgensen and Kirschenheiter 2012). However, the free-riding argument may not hold if discretionary disclosures after peer disclosures are not that costly. If much of the information has already been transferred from peers for free-riding to occur, subsequent discretionary disclosures by the underlying firm is unlikely to bear significant disclosure cost (e.g., proprietary cost), other than small administrative costs in drafting and sending out press releases. In addition, free-riding may be less likely to happen in the first place if investors have difficulty interpreting the implications of peer disclosures on the underlying firm.

Due to possibly correlated signal arrival processes among comparable peer firms, investors may increase their perceived probability of information arrival for the underlying firm whenever any of its close peer firms discloses. This may raise the uncertainty among investors as well as their assessed likelihood that the absence of disclosure by the underlying firm is due to deliberate information withholding by manager for strategic reasons rather than due to the absence of new information, consequently lowering the disclosure threshold for the underlying firm (Jung and Kwon 1988, Dye and Sridhar 1995). In addition, if information transfer is noisy and confirmatory discretionary disclosures are not so costly, investors may have a preference for firms providing discretionary disclosures to resolve the residual uncertainty or even just to confirm prior market expectations in response to peer disclosures. As firms with more comparable public peers are more often exposed to such peer dynamics, I expect that this can induce a higher frequency of discretionary disclosure for these firms.

In the paper, I identify a firm's closest peer by the firm with the highest textual similarity in the business description section of 10-K filings using a vector space model. I measure the number of comparable public peers for a firm by the number of public firms sharing high textual similarity with the underlying firm. An alternative measure based on similarity in business segment distribution is also used as a robustness check. In addition, the analysis on the frequency of discretionary disclosure is also conducted at the industry-level using the number of public firms in each industry to represent the size of peer groups based on three widely used industry classification schemes respectively³⁰. To measure the frequency of discretionary disclosure, I use the annual frequency of "voluntary" 8-K filing under Item 2.02, Item 7.01 or Item 8.01, similar to Cooper, et al. (2015). Alternatively, I also use a measure that excludes Item 2.02, which consists of mostly earnings announcements, to represent the more discretionary disclosures.

I start my analyses by investigating whether the discretionary disclosures by peers may raise the uncertainty among investors over the underlying firm with no contemporaneous disclosure. I examine the changes in the daily closing bid-ask spread of the underlying firm when its closest peer provides discretionary disclosures on major updates of business conditions³¹. Controlling for contemporaneous disclosure by the underlying firm, I find that the bid-ask spread increases slightly, by 9 basis points, on average across the underlying firm after the discretionary disclosure of its closest peer.

³⁰ The three industry classification schemes include North American Industry Classification System (NAICS), Standard Industrial Classification (SIC), and Global Industry Classifications Standard (GICS).

³¹ In my sample, major business updates refer to client announcements, guidance updates, business expansions, business reorganizations, product announcements, strategic alliances, conference calls. These major business updates likely have some implications about the underlying firm either through their common exposures or competitive positions.

However, the presence of subsequent discretionary disclosure by the underlying firm appears to be able to largely restore the bid-ask spread back to its status before its peer discloses. The results suggest that firms may not be able to completely free-ride on the discretionary disclosures from peers as investors may find it difficult to perfectly interpret the implications of peer discretionary disclosures on the underlying firm without its concurrent discretionary disclosure. The results are also consistent with investors possibly increasing their assessed likelihood that non-disclosure is due to strategic information withholding from an increased perceived probability of information arrival from peer disclosure.

Next, I examine whether the presence of comparable peers induces a higher frequency of discretionary disclosure using both industry-level and firm-level analyses. In the industry-level analyses, I find that the number of public firms in each industry defined using any of the SIC, NAICS and GICS industry classifications is positively associated with the average annual frequency of discretionary disclosure per firm in the industry. This relationship is present in both cross-sectional and time-series variations. In the firm-level analyses, I also find that the number of comparable public peers identified using textual analyses on their 10-K business description is positively associated with the annual frequency of discretionary disclosure for each firm. Among the sample of firms that provide management guidance, I also find that the number of comparable public peers is positively associated with the frequency of unbundled guidance, which resembles discretionary disclosure upon information arrival in the context of management guidance. These results provide strong evidence consistent with the presence of comparable peers inducing a higher frequency of discretionary disclosure.

To provide further evidence for such peer dynamics, I examine how the business similarity between the underlying firm and its closest peer relates to the timing of discretionary disclosures between them. Using the Cox proportional hazard model, I find that higher similarity between the underlying firm and its closest peer is associated with a higher likelihood of subsequent discretionary disclosure by the other firm when any one of the firms in the pair provides discretionary disclosure first. Using OLS regression, I also find that the number of days between the closest discretionary disclosures of the underlying firm and its closest peer is smaller if the similarity between the underlying firm and its closest peer is higher. I repeat the tests for the relationships with the second-, third- and fourth-closest peers and find similar results. Thus, when a firm has many close public peers, the presence of peer dynamics from its numerous peers likely induces the firm to provide discretionary disclosure more frequently.

My study contributes to the burgeoning empirical literature on multi-firm dynamics of discretionary disclosure (e.g., Tse and Tucker 2010, Rogers et al. 2014). While Tse and Tucker (2010) find evidence supporting a clustering in the timing of earnings warnings within industries, my results suggest that having more comparable public peers induces a higher frequency of discretionary disclosure due to the presence of peer dynamics. In relation to Rogers et al. (2014), my results may also suggest that having a larger industry size can make tacit collusion to withhold industry-wide news more difficult. In the paper, I also utilize the details of 8-K filings to construct a more comprehensive frequency measure for discretionary disclosure instead of relying on the traditional management forecast frequency measure.

The study also adds to the information transfer literature that what investors can learn from peer discretionary disclosures involve not only the content of news, but also the arrival of information arrival for the underlying firm. The slight increase in the bid-ask spread for non-disclosure firms in the presence of peer discretionary disclosures also suggests that investors may not be able to always perfectly understand the implications of peer disclosures on the underlying firm. The results point to the possibility that peer disclosures can sometimes introduce additional uncertainty over whether the absence of disclosure by the underlying firm is due to strategic information withholding or the absence of new information.

In addition, the results of the study may have implications on the discretionary disclosure behaviors with regard to the developmental stage of capital markets. In emerging markets with few public firms available in the capital markets, investors may have less knowledge on the arrival of information for public firms. Thus in some of the small emerging markets, peer dynamics that induce higher frequency of discretionary disclosure can be much weaker compared to the larger and more developed markets.

The remainder of the paper is organized as follows. Section 2 reviews the prior literature and develops my hypotheses. Section 3 describes my variables measurement and research design. Section 4 presents my results. Section 5 concludes.

2. Prior Literature and Hypotheses Development

2.1 Related Literature

Prior theoretical literature on discretionary disclosure using a single-firm setting has generally shown that a firm's propensity to provide discretionary disclosure increases as

disclosure-related costs decrease (e.g. Verrecchia 1983) and as investors' perceptions on the probability that the firm receives information increase (e.g. Dye 1985, Jung and Kwon 1988). The two most closely related theoretical papers using a multi-firm setting include Dye and Sridhar (1995) and Jorgensen and Kirschenheiter (2012).³²

In Dye and Sridhar (1995), the model assumes that the probability of being informed is correlated but the private signals are uncorrelated. Dye and Sridhar (1995) generally predicts “herding” behavior in discretionary disclosure as long as there is positive correlation among firms' receipt of information and suggests that a firm's equilibrium disclosure policy depends on whether the information the firm receives is private knowledge or industry-wide common knowledge.

In Jorgensen and Kirschenheiter (2012), the model assumes that the private signals are correlated with an exogenous leader and follower. Jorgensen and Kirschenheiter (2012) predict a higher disclosure threshold by the follower due to information free-riding if the signals are positively correlated and a lower disclosure threshold by the follower due to negative information externality if the signals are negatively correlated.

While the theoretical models provide some guidance on discretionary behaviors under certain assumptions and conditions, it remains to be empirically tested whether some of the conditions are more prevalent for firms in general. A key tension is whether firms can free-ride on the information transfer from peer disclosures through non-disclosure assuming that discretionary disclosures are costly or whether firms are induced to

³² Another related paper is by Acharya, et al. (2011) which considers the strategic timing of information releases in a dynamic disclosure model and predicts clustering of negative announcements. However, the paper is more focused on the timing of discretionary disclosure.

disclose more frequently as investors infer a higher probability of information arrival from peer disclosures.

In addition to the theory literature on multi-firm dynamics of discretionary disclosure, there is also a burgeoning empirical literature on multi-firm phenomenon with respect to discretionary disclosure. Brown et al. (2006) provides evidence on the existence of herding behavior in capital expenditure forecasts by firms. Tse and Tucker (2010) focus on the timing of bad-news discretionary disclosure and finds that firms speed up their warnings in response to peer firms' warnings. They argue that this phenomenon is due to managers pooling bad-news disclosure with peers to minimize responsibility for earnings shortfalls. Rogers et al. (2014) investigates whether managers can tacitly collude to withhold industry-wide bad news and seek to understand what factors makes strategic cooperative withholding more likely. It documents some cases of increased intra-industry obfuscation of adverse signals in annual 10-K filings.

Peer dynamics in discretionary disclosure is also related to the literature on within-industry information transfers from earnings announcements (e.g. Foster 1981, Han and Wild 1990, Freeman and Tse 1992) and management forecasts (e.g. Baginski 1987, Han et. al. 1989). While prior empirical literature has generally demonstrated the existence of within-industry information transfers, a few papers have pointed out that the information transfer may not be perfect. For example, Ramnath (2002) find some evidence consistent with investors and analysts not fully incorporating the information from the earnings report of the first announcers in the industry in their revised earnings expectations for subsequent announcers. On the other hand, Thomas and Zhang (2008) find some

evidence consistent with the stock market overestimating the intra-industry implications of early announcers' earnings for late announcers' earnings.

While the above papers suggest that investors may have some trouble interpreting the information content from the earnings announcements from peer firms within an industry, there may be an additional factor for investors to consider when it comes to the discretionary disclosures by peer firms, i.e. the probability of information arrival for the underlying firm. While earnings announcements often follow a relatively fixed schedule, the initiation of discretionary disclosure may often signal information arrival for related firms. This may raise the uncertainty among investors regarding whether non-disclosure by firms is due to strategic withholding of information or the absence of new information.

2.2 Hypotheses Development

2.2.1. Inferences on Non-Disclosure

When a firm's closest peer firm initiates any discretionary disclosure, investors may use the content in peer disclosures to infer the implications on the underlying firm. It is unclear if investors can perfectly make inferences on the news contained in peer disclosures. The news may have a common exposure element as well as a competitive element. In the case of earnings announcements, prior literature presents evidence for both underreaction (e.g. Ramnath 2002) or overreaction (e.g. Thomas and Zhang 2008) by investors on late announcers within an industry in different samples. If investors can perfectly interpret the implications on the underlying firm from peer disclosures, discretionary disclosures by peers likely further improves the information environment of the underlying firm.

On the other hand, if investors cannot perfectly interpret the information contained in peer disclosures, they may want the underlying firm to provide updates and clarifications on the situation. When their close peers provide discretionary disclosures, investors may infer a higher likelihood of information arrival for the underlying firm. If the underlying firm does not provide discretionary disclosure around the same time, investors may incur higher uncertainty in interpreting “non-disclosure” by the underlying firm and may assign an increasing likelihood that non-disclosure is due to strategic information withholding rather than due to the absence of new information.

Thus, there may be an increase in the information asymmetry between insiders and outsiders of the underlying firm when the closest peer provide discretionary disclosures but the underlying firm does not. The rise in perceived information uncertainty by outsiders is likely leading to higher bid-ask spread of the underlying firm (Glosten and Milgrom 1985). For those underlying firms that also provide discretionary disclosures contemporaneously or in the near subsequent periods, I expect that the bid-ask spread of the underlying firm should likely stay flat or drop back to the initial level after disclosures³³. This notion leads to my first hypothesis:

H1: The bid-ask spread of the underlying firm with no contemporaneous discretionary disclosure increases after its closest peer provides business-related discretionary disclosure.

2.2.2. Frequency of Discretionary Disclosure

³³ Parallel to this prediction, in the case of overreaction in intra-information transfer from earnings announcements, Thomas and Zhang (2008) finds that the overestimation is corrected when late announcers disclose their earnings.

In addition to possible information transfers from the discretionary disclosure, the initiation of discretionary disclosures by peer firms may also signal information arrival, which can raise investors' inferred probability that non-disclosure is due to strategic information withholding rather than the absence of new information. While the possibility of free-riding on peer discretionary disclosures may reduce discretionary disclosures by the underlying firm, the increased perceived probability of information arrival may induce more discretionary disclosures by the underlying firm. My second hypothesis is concerned with how the extent of the presence of close public peers relates to the frequency of discretionary disclosure by firms.

On one hand, if discretionary disclosure is costly and information transfer from peer discretionary disclosures is perfect, the underlying firm may incur positive information externality from peer discretionary disclosures (Jorgensen and Kirschenheiter 2012). The underlying firm may thus potentially benefit from free-riding on the discretionary disclosures of peer firms and save on the disclosure cost. However, the free-riding argument may not hold if sequential discretionary disclosure is not that costly. If much of the information has been transferred from peers already for free-riding to occur, discretionary disclosure by the underlying firm is unlikely containing significant disclosure cost (e.g. proprietary cost of disclosure), other than some administrative costs in drafting and sending out press releases.

On the other hand, investors may increase their perceived probability of information arrival for the underlying firm whenever a peer firm initiates any discretionary disclosure due to possibly correlated signal arrival processes among similar firms. Non-disclosure by the underlying firm may be more likely interpreted as due to strategic information

withholding instead of the absence of new information. An increased assessed probability of information arrival likely lowers the disclosure threshold for the underlying firm (Jung and Kwon 1988). For firms with many comparable public peers, they are more often exposed to such group disclosure dynamics whenever some peer firms provide discretionary disclosures (Dye and Sridhar 1995).

Thus, I expect that firms with many comparable public peers should have a higher frequency of discretionary disclosure under the influence of peer dynamics. There is potential positive feedback effect on discretionary disclosure among the groups of similar firms as more peer firms initiate discretionary disclosures and further increase investors' perceived probability of information arrival for the remaining non-disclosure firms in the group. In addition, if information transfer is often imperfect and noisy, the possibility of negative information externality from information transfer may also induce some firms to initiate discretionary disclosures to correct for the imperfect information transfers (Jorgensen and Kirschenheiter 2012). This leads to my second hypothesis:

H2: The number of comparable public peers is associated with a higher frequency of discretionary disclosure.

2.2.3. Timing of Discretionary Disclosure

Naturally, the timing on the incidence of events should be closer among more similar firms. The arrival of information should be more correlated among firms that have more similar business operations. If a firm's closest peer is more similar to the underlying firm, investors are more likely to increase their assessed probability of information arrival when its closest peer initiates any discretionary disclosure. Non-disclosure by the underlying firm in the presence of peer disclosures may be viewed negatively due to the

possibility of strategic information withholding. Thus the underlying firm is more likely to provide discretionary disclosure contemporaneously to avoid the possible adverse interpretation of non-disclosure by investors. As an extension to my second hypothesis, I expect that the timing between discretionary disclosures between firm pairs that are more similar to each other should be on average shorter in duration. In other words, I expect that firms generally do not free-ride on peer discretionary disclosures through non-disclosure themselves as information transfers alone may not be sufficient to fully resolve the information uncertainty by investors. Therefore:

H3: The similarity between a firm and its close peer is associated with closer timing in discretionary disclosures between the two firms.

3. Variable Measurement and Research Design

3.1 Measuring Peer Size

3.1.1 Industry-Level Measure Based on Industry Classification

I use the number of public firms within each industry based on 3 different widely used industry classification schemes as a measure of peer size at the industry-level. The finest industry classification scheme I use is North American Industry Classification System (NAICS) at the 6-digit level, with 1,131 industries in total. In addition, I use the traditional Standard Industrial Classification (SIC) scheme at its 4-digit level, with a size of 441 industries in total. Lastly, I use the Global Industry Classifications Standard (GICS) at its 6-digit level to represent industry boundaries, with a size of 70 industries in total³⁴.

³⁴ GICS was jointly developed by Standard & Poor's (S&P) and Morgan Stanley Capital International (MSCI) in 1999 basing its classification on a firm's principal business activity. Bhojraj et al. (2002)

3.1.2 Firm-Level Measure Based on Textual Business Description

Similar to the textual analysis methodology used in Chapter 1, I compute cosine similarity scores on the business description section in 10-Ks across all public firms using a vector space model with TF-IDF weighting scheme to compare similarities across firms. This measure assumes truthful disclosures by firms in the business description sections of 10-Ks, which are governed by regulation S-K. For my primary measure, the number of peer firms above a similarity threshold that classifies 25% of the firms as having no close comparable peer is used to represent the size of peer groups. I also include measures based on different levels of similarity threshold for counting peers as robustness checks and the results largely stay the same. More details on measure construction are included in Appendix B.

3.2 Measuring Frequency of Discretionary Disclosure

The frequency of discretionary disclosure is the key dependent variable of interest in the paper. Prior empirical literature has used the frequency of management forecasts as one of the common proxies for voluntary disclosure (e.g. Kasznik 1999, Baginski et al. 2002, Ball et al. 2012, Li and Yang 2015). On the other hand, Rogers and Van Buskirk (2013) shows that firms are increasingly bundling their management forecasts with earnings announcements over time. Bundled guidance has evolved to be the most prevalent type of guidance. My first chapter argues that the practice of bundling guidance resembles a tacit *ex-ante* commitment to disclosure and may not highly discretionary. As I focus on understanding the *ex-post* discretionary disclosure behaviors by firms upon the

provides more details on GICS and also shows that GICS classifications are better at explaining stock return comovements and cross-sectional variations in various accounting ratios.

arrival of signals in this paper, I can potentially use the frequency of unbundled guidance as a proxy for the frequency of discretionary disclosure in the context of management guidance³⁵.

However, management guidance is just one out of many types of discretionary disclosure issued by firms. Firms may issue press releases to inform investors upon product launches, client announcements, strategic alliances, business reorganization, etc. One potential way to measure the frequency of discretionary disclosure more comprehensively is by the annual frequency of “voluntary” 8-K filings under “Item 2.02 Results of Operations and Financial Conditions”, “Item 7.01 Regulation FD Disclosure” or “Item 8.01 Other Events”, similar to the 8-K item classification scheme on voluntary disclosure by Lerman and Livant (2010) as well as Cooper et al. (2015)³⁶. Such classification will naturally include both the frequency of bundled and unbundled guidance in the measure³⁷. However, as Item 2.02 often contains a firm’s earnings announcements which are regarded as regular disclosure usually, I also measure the frequency of discretionary disclosure using “Item 7.01 Regulation FD Disclosure” or “Item 8.01 Other Events” from 8-K filings only, in attempt to better capture disclosures that are more discretionary³⁸. The results using both measures are largely similar in the paper.

³⁵ I use this measure in column (4) of Table 5 as additional analysis.

³⁶ There was a major overhaul over 8-K filings in August 2004 but these three items were largely unchanged. However, these three items were numbered differently prior to the overhaul: “Item 12. Results Operations”, “Item 9. Regulation FD Disclosure”, “Item 5. Other Events”. My results are robust to the exclusion of 8-K filings before the overhaul in 2004.

³⁷ Bundled guidance is likely contained in “Item 2.02” of 8-K filings while unbundled guidance is likely contained in “Item 7.01” of 8-K filings.

³⁸ Public companies are required to report material corporate events on a current basis using 8-K forms. While much of the 8-K filings are likely linked to mandatory disclosures upon the occurrence of certain

3.3 Research Design

3.3.1 Inferences on Non-Disclosure

To understand how investors' inferences on the underlying firm changes when its closest peer initiates discretionary disclosure, I examine the changes in the bid-ask spread of the underlying firm when its closest peer initiates discretionary disclosure as well as when the underlying firm also subsequently provides discretionary disclosure within a 7-day window after peer disclosure.

To construct my sample of events, I first include all 8-K filings under Item 2.02, Item 7.01 and 8.01 by each firm's closest peer identified through textual analysis on the 10-K business description. Then I match all these discretionary 8-K filings with the Capital IQ Key Development database using the date of the earliest report in 8-K filings as the event date. I only include matched discretionary disclosures related to the following major corporate event types for my sample of event: client announcements, product announcements, business expansions, business reorganizations, strategic alliances, corporate guidance and conference calls. These major corporate events are chosen because they are likely to have implications on the underlying firm due to either common exposures or competition concerns. Next, I search for the closest discretionary disclosure made by the underlying firm for each event, i.e. discretionary disclosure from its closest peer, and exclude all events in which the underlying firm provides discretionary

material corporate events, firms are also required to file an 8-K filing under either "Item 7.01" or "Item 8.01" for any material discretionary disclosures. For example, if a firm voluntarily releases press releases providing updates on management guidance, product launches, new business developments, etc., it also tends to file an 8-K filing with the press releases as exhibits to satisfy the disclosure requirements. Some examples of 8-K filings under "Item 7.01" or "Item 8.01" are provided in Appendix C. By including 8-Ks containing "Item 7.01" and "Item 8.01" only in counting frequency, this measure will mostly exclude the frequency of bundled guidance but include the frequency of unbundled guidance.

disclosures first, i.e. within 7 days before its closest peer discloses. My final sample consists of 27,867 peer disclosure events and includes observations of daily bid-ask spread for the underlying firm within a [-7 days, +7 days] window around the date of each event.

I measure the bid-ask spread as the difference between daily closing ask price and bid price divided by the mid-point of daily closing bid and ask price. I code *event date* as an indicator for the date of the earliest report of the discretionary disclosure found in the 8-K filings by the closest peer, *after event* as an indicator for dates following the event date [+1 day, +7 days], *post event* as an indicator for dates on and following the event date [0, +7 days]. In addition, I code *disclose date* as an indicator for the date of the earliest report of the discretionary disclosure by the underlying firm within a [0, +7 days] window around the event date, *after disclosure* as an indicator for dates following the disclosure date [disclose date+1, +7 days], and *post disclosure* as an indicator for dates on and following the disclosure date [disclose date, +7 days].

To examine how the market inferences on the underlying firm changes when its closest peer initiates any discretionary disclosure, I regress the daily closing bid-ask spread of the underlying firm on the indicators for days on and after peer disclosure (*post event*) and for days on and after corresponding disclosure by the underlying firm (*post disclosure*), with event fixed effect. To be consistent with *H1*, I expect that the coefficient on *post event* in (1) below should be positive and the coefficient on *post disclosure* in (1) below should be negative, as peer disclosure increases investors' uncertainty on the implications of "non-disclosure" but subsequent disclosure by the underlying firm reduces investors' uncertainty. In another specification, I further break down *post event*

into peer disclosure date (*event date*) and days after peer disclosure (*after event*) and break down *post disclosure* into disclosure date by the underlying firm (*disclose date*) and days after disclosure by the underlying firm (*after disclosure*).

$$Bid-Ask Spread = \alpha + \beta_1 Post Event + \beta_2 Post Disclosure + \beta_4 Event F.E. + \varepsilon \text{ --- (1)}$$

3.3.2 Frequency of Discretionary Disclosure

To examine how the frequency of discretionary disclosure is associated with the number of comparable public peers, I mainly rely on the cross-sectional variations across firms. The number of comparable public peers is likely an inherent and sticky property in relation to all other firms in the universe of public firms. It depends not only on the actions of the firm itself but also other firms. As firms likely cannot change their number of comparable peers easily without substantially change the nature of their businesses, there should be little concern over reverse causality bias in such a cross-sectional research design.

Major identification threats include measurement errors with respect to the uniqueness measures as well as correlated omitted variable bias. To mitigate concerns for measurement errors, I use peer size measured both at the industry-level using 3 different traditional industry classification schemes as well as at the firm-level using identified comparable peer firms by textual analyses on business description in 10-K filings. I have also included various controls for firm characteristics, such as analyst coverage, institutional ownership, firm size, stock returns, stock beta, market-to-book ratio, stock return volatility, years of listing, profitability, leverage, etc. to mitigate concerns for

correlated omitted variables. More control variables are also included in additional analyses for robustness checks.

In the industry-level analysis, I include all U.S. industries with listed public firms from 1996 to 2013 using NAICS 6-digit code, SIC 4-digit code and GIC 6-digit code respectively.³⁹ In firm-level analysis, I include all U.S. public firms from 2002 to 2013 where financial reporting, returns, guidance, industry classification, and valid textual business description data are available from CRSP, COMPUSTAT, IBES, and SEC Edgar filings. I exclude firms in regulated industries such as utilities and banking as there are often different disclosure requirements for these firms. After imposing further data requirement on major control variables, the sample with firm-level data used in most specifications comprises of 33,277 observations.

In my preliminary industry-level analysis, I regress the average frequency of discretionary disclosure per firm within an industry on the number of firms within an industry, with or without year fixed effects. In my main firm-level analysis, I regress the frequency of voluntary 8-K filings on the number of comparable peers identified through textual analysis, and other firm-level controls. To be consistent with *H2*, the sign for the coefficients on *peer size* should be positive in (2) when using either industry-level or firm-level peer size measure. I further separate the frequency of discretionary disclosure into those with positive news and negative news respectively in my analyses⁴⁰.

³⁹ As GICS industry classification is unavailable before 1999 from COMPUSTAT, I assume that the industry assignment of each firm prior to 1999 is the same as that in 1999 or later when data is available.

⁴⁰ I categorize discretionary disclosure as containing positive news or negative news if abnormal returns are above or below 0% within a [-1 day, +1 day] window around its date of first report as identified from the 8-K filings. Based on traditional discretionary disclosure model in which managers try to maximize the share price without considering any litigation costs, “peer dynamics” may manifest slight more in terms of higher frequency of discretionary disclosure with negative news if the discretionary threshold is lower in the

Conditional on firms providing management guidance, I also run a regression with the frequency of unbundled guidance as the dependent variable. Among guiding firms, I expect that the number of comparable peers is positively associated with the frequency of unbundled guidance, which is likely more discretionary in nature.

$$\text{Frequency of Discretionary Disclosure} = \alpha + \beta_1 \text{Peer Size} + \beta_2 \text{Controls} + \beta_4 \text{Year FE} + \varepsilon \quad \text{--- (2)}$$

3.3.3 Timing of Discretionary Disclosure

A commonly used method in analyzing the timing of incidents is duration analysis. I measure the timing of discretionary disclosure as the number of days between the closest discretionary disclosure from the peer and the discretionary disclosure from the underlying firm. For each discretionary disclosure of the underlying firm, I match it with the closest peer's closest discretionary disclosure in terms of disclosure timing. This generates a total 245,202 observations matching discretionary disclosure between the underlying firm and its closest peer from 1996 to 2013, conditional on the duration between discretionary disclosures of the underlying firm and the closest discretionary disclosure of its closet peer below 90 days. As the similarity between the underlying firm and its closest peer firm varies in cross-section, *H3* predicts that higher similarity between the pairs will be associated with shorter duration between their discretionary disclosures.

First, I try to estimate a simple Cox proportional hazard model to conduct the duration analysis in Table 6. In this case, the hazard refers to the probability that its

presence of peer disclosure. However, managers may also have incentives to be “forthcoming” in disclosure for both good and bad news due to litigation or reputational concerns. If information transfer is imperfect, the underlying firm may make confirmatory disclosures for both good and bad news to reduce residual investors' uncertainty.

closest peer provides discretionary disclosure on a particular day either before or after from the day that the underlying firm initiates discretionary disclosure. Thus the hazard function can be interpreted as inversely related to the duration between discretionary disclosure from the closest peer and the underlying firm. Cox proportional hazard model as represented by (3) below specifies a common baseline hazard for all firms $h_0(t_i)$ and allow individual firms' hazard function $h(t_i)$ to differ proportionally with the similarity between the underlying firm and its closest peer, which is measured by the decile ranks of the textual similarity score (*similarity decile*) between the closest peer firm and the underlying firm in each year. To be consistent with *H3*, the coefficient on *similarity decile* should be positive, representing a higher likelihood of disclosure if the peer firm is more similar to the underlying firm.

For the full sample in Panel A, variable t represents the number of days apart from the discretionary disclosure of the underlying firm when the closest peer firm provides discretionary disclosure, either before or after the disclosure of the underlying firm. In Panel B, I restrict the sample to situations in which the peer firm discloses before the underlying firm. Thus variable t represents the number of days before the discretionary disclosure of the underlying firm when its closest firm provides discretionary disclosure. In Panel C, I restrict the sample to situations in which the peer discloses after the underlying firm. Thus variable t represents the number of days after the discretionary disclosure of the underlying firm when its closest firm provides discretionary disclosure. In addition to matching with the discretionary disclosure of the closest peer firm, I also match the discretionary disclosure of the underlying firm with the 2nd closest, 3rd closest and 4th closest firm to show that this phenomenon holds across peer pairs in general.

$$h(t_i) = h_0(t_i) \exp [b_0 \text{Similarity Decile} + \sum b_k \text{Year Dummy}_k] \text{ --- (3)}$$

Second, I run a similar analysis using the more intuitive OLS regression framework with additional control variables in Table 7. Similar to Table 6, the sample is also conditional on the duration between discretionary disclosures of the underlying firm and its closet peer below 90 days. In the regression, my dependent variable uses the number of days apart from the discretionary disclosure of the underlying firm when the closest peer firm provides discretionary disclosure, either before or after the disclosure of the underlying firm. Major independent variable of interest is the similarity between the underlying firm and its closest peer (*similarity decile*). To be consistent with *H3*, I expect the coefficient on *similarity decile* to be negative in (4). I have also included various controls for firm characteristics, such as analyst coverage, institutional ownership, firm size, stock returns, stock beta, market-to-book ratio, stock return volatility, years of listing, profitability, leverage, etc.

$$\text{Number of Days Apart} = \alpha + \beta_1 \text{Similarity Decile} + \beta_2 \text{Controls} + \beta_4 \text{Year FE} + \varepsilon \text{ --- (4)}$$

4. Results

4.1 Summary Statistics

Table 1 displays the summary statistics for the main variables in various tests in the paper. The average annual frequency of 8-K filings which contains either Item 2.02, 7.01 or 9.01 is about 6.1, with about 3.0 filings containing positive news and 3.1 filings contain negative news. The average annual frequency of 8-K filings which contain either Item 7.01 or 9.01 is about 4.3. This suggests that some of 8-Ks contain both Item 2.02 and Item 7.01/Item 9.01 in one filing. The average size of the peer group is 5.06 when

defining comparable peers based on textual analysis on business description using a similarity threshold that defines 25% of firms in each year peerless.⁴¹ Table 2 displays the correlation table for the major variables used in the main test of Table 5. The correlation between peer size and frequency of discretionary disclosure is highly positive, providing preliminary support to my hypothesis in the paper.

4.2 Inferences on Non-Disclosure

In Table 3, I examine how the uncertainty of market inference on the underlying firm changes when its closest peer initiates discretionary disclosure by regressing daily closing bid-ask spread on indicators for days on and after peer disclosures and days on and after disclosures by the underlying firm, with the inclusion of event fixed effect.

The result from column (1) suggest that there is a slight increase in bid-ask spread after the closest peer firm initiates discretionary disclosure but the coefficient on *post event* is not statistically significant, possibly due to some firms also provide contemporaneous discretionary disclosures to remove the hanging uncertainty of “non-disclosure”. In column (2), after adding an additional independent variable *post disclosure* to control for the possible contemporaneous discretionary disclosures by the underlying firm, the positive coefficient on *post event* becomes statistically significant. This suggests that the discretionary disclosure by the closest peer may increase the bid-ask spread of the underlying firm through heightened investors’ uncertainty over the interpretation of “non-disclosure” by the underlying firm. The coefficient on *post*

⁴¹ For robustness checks, I also use the peer size calculated using different similarity thresholds in textual analysis and the results are largely the same. If I use a similarity threshold that defines 15% of the firm as peerless, the average size of the peer group becomes 8.6. If I use a similarity threshold that defines 50% of the firm as peerless, the average size of the peer group becomes 1.6.

disclosure indicator is significantly negative and similar in magnitude with the coefficient on *post event* indicator. This is consistent with discretionary disclosures by the underlying firm serving to remove the heightened investors' uncertainty from peer disclosures. Column (3) further breaks down the indicator variables into the event/disclosure day and the days after the event/disclosure. The results indicate that the effect is generally strongest on the event/disclosure day but also sustains at a slightly decreasing magnitude through the 7-day window.

In column (4), I replace the bid-ask spread with the daily trading volume as the dependent variable in the regression. The result shows that on the day of discretionary disclosure by the closest peer, there is a significant increase in trading volume by about 6% relative to the average trading volume. On the day of discretionary disclosure by the underlying firm, there is a much more significant increase in trading volume by about 69% relative to the average trading volume. If trading volume is related to the information content (e.g. Beaver, 1968), it suggests that the information transfer from peer disclosure is slightly less than 10% of the information content from contemporaneous/subsequent firm-initiated disclosure in this setting.

Overall, the results provide strong support to *H2* that the bid-ask spread of the underlying firm with no contemporaneous discretionary disclosure increases after its closest peer initiates business-related discretionary disclosure. This is consistent with peer disclosures increasing investors' perceived probability of information arrival for the underlying firm and leading to heightened uncertainty over the implications of "non-disclosure" by the underlying firm, and possibly increased assessed probability that "non-

disclosure” by the underlying firm is due to strategic withholding of information rather than the absence of new information.

4.3 Frequency of Discretionary Disclosure

In Table 4, I conduct an industry-level analysis to test *H2* by examining how the average annual frequency of discretionary disclosure of firms varies with the number of public firms in each industry. Using either NAICS 6-digit classification, SIC 4-digit classification or GIC 6-digit classification, I find that the number of public firms in each industry is strongly positively associated with the average annual frequency of discretionary disclosure per firm in the industry in columns (1), (3) and (5). After including year fixed effect in columns (2), (4) and (6) and thus only limiting to examining the cross-sectional variations across industries, I find similar results that the number of public firms in each industry is indeed positively associated with the frequency of discretionary disclosure. In untabulated tests, I include industry fixed effect instead of year fixed effect and thus focus only on the time-series variation within each industry. Similar results are obtained as well.

In Table 5, I conduct a firm-level analysis to test *H2* by examining how the annual frequency of discretionary disclosure of firms varies with the number of comparable public peers identified through textual analysis on business description similarity. Controlling for guidance behaviors and firm characteristics in column (1), I find that the number of comparable public peers for each firm is significantly positively associated with the frequency of voluntary 8-K filings⁴². In columns (2) and (3), I find that the

⁴² In the regression, I control for the frequency of bundled and unbundled guidance. My first chapter shows that more “informationally unique” firms are more likely to provide bundled guidance to compensate for

results in column (1) hold for the frequency of discretionary disclosure with both positive and negative news, with the coefficient on *peer size* for negative news slightly larger in magnitude. In column (4), using the frequency of unbundled guidance as the dependent variable among a sample of guiding firms, I find similar results that the number of comparable public peers for a firm is positively associated with higher frequency of unbundled guidance, which is more discretionary in nature compared to bundled guidance.

Combining Tables 4 and 5, I provide strong evidence that the size of the relevant peer group for a firm is positively associated with higher frequency of discretionary disclosure using both industry-level and firm-level analyses. This is consistent with having more comparable public peers increasing investors' inferred probability of signal arrival for the underlying firm and thus inducing the underlying firm to provide discretionary disclosures more frequently.

4.4 Timing of Discretionary Disclosure

In Table 6, I run a simple Cox proportional hazard model to investigate how the similarity between the underlying firm and its close peer is related to their timing of discretionary disclosure. The hazard function in the model refers to the incidence of disclosure by another firm given that the other firm in the peer pair initiates any discretionary disclosure. Panel A displays the results for days apart between the discretionary disclosure of the underlying firm and the closest discretionary disclosure by

the information deficiency from the lack of comparable public peers. Consistent with the presence of regular bundled guidance possibly preempting the need by firms to provide additional discretionary disclosures. I find that the coefficient on bundled guidance is significantly negative. However, as I control for the presence of bundled guidance in the regression, my results on peer dynamics are on top of the effect from the presence of guidance policy.

the peer firm either before or after that of the underlying firm. Panel B and C displays similar results but conditional on the peer firm discloses before and after the underlying firm respectively. The results across peer pairs from the closest peer in column (1) to the 4th closest peer in column (4) for each firm show that firms in peer pairs of the top similarity decile⁴³ is about 20% more likely to provide discretionary disclosure than firms in peer pairs of the bottom similarity decile once the other firm in the pairs initiates any discretionary disclosure⁴³.

In Table 7, I use the more intuitive OLS regression model to regress the number of days apart between discretionary disclosure of the underlying firm and the closest discretionary disclosure by its close peer either before or after the underlying firm⁴⁴. The table displays the results for the timing of discretionary disclosures between the underlying and its close peer, from the top closest peer in column (1) to the 4th closest peer in column (4). Controlling for standard firm characteristic, the results indicate that the duration between the discretionary disclosures of the underlying firm and its close peer is about 3 to 4 days shorter, moving from peer pairs in the bottom similarity decile⁴³ to peer pairs in the top similarity decile⁴³ for each column in Table 7.

Tables 6 and 7 combined provide strong support to *H3* that the similarity between a firm and its close peer is associated with closer timing in discretionary disclosures between the two firms, consistent with the potential influence of “peer dynamics” on

⁴³ Similarity deciles are calculated with respect to the similarity between the underlying firm and its closest peer for Column (1), between the underlying firm and its 2nd closest peer for Column (2), between the underlying firm and its 2nd closest peer for Column (2), between the underlying firm and its 3rd closest peer for Column (3), between the underlying firm and its 4th closest peer for Column (4).

⁴⁴The results are robust to alternative specification using Poisson regression.

firms' discretionary disclosure behaviors⁴⁵. This also suggests that the informational arrival process between firms that are more similar to each other is more correlated in terms of timing. For a firm that has many similar peers, it may often be subject to such peer dynamics whenever some of its peer firms provide any discretionary disclosure. The potential increase in investors' uncertainty in *HI* may thus induce these firms with many public peers to provide more frequent discretionary disclosures.

4.5 Additional Analyses

4.5.1 Controlling for Frequency of Events

One potential alternative explanation for firms with more comparable public peers providing more frequent discretionary disclosures is that these firms experience more events. However, there does not appear to be any strong economic reasons to expect that firms with more public peers experience more events. For firms with more public peers, it is likely that outsiders have better knowledge on the incidence of events as they happen. Nevertheless, I attempt to control for the frequency of events happening as additional analyses in Table 8.

In column (1) of Table 8, I include a variable that represents the annual frequency of 8-K filings that contain "Item 2.01 Completion of Acquisition or Disposition of Assets" to control for the occurrence of M&A activities. While the frequency of M&A activities is significantly positively associated the frequency of discretionary disclosure, the coefficient on *peer size* continues to be positive and statistically significant. In column (2), I include a variable that represents the frequency of mandatory 8-K filings that do not

⁴⁵The results are robust to alternative samples conditional on shorter duration between discretionary disclosures of the underlying firm and the closest discretionary disclosure from its closet peer, such as 30 days instead of 90 days in the primary sample.

contain Item 2.02, Item 7.01 and Item 9.01 to control for the occurrence of events that trigger mandatory disclosure. While the frequency of mandatory 8-K filings is also significantly positively associated the frequency of discretionary disclosure, the coefficient on *peer size* still continues to be positive and statistically significant. In column (3), I attempt to include proxies for firm complexity by including the number of industrial segments and geographical segments as additional controls. The results still largely remain the same.

4.5.2 Alternative Measure of Discretionary Disclosure

In Table 5, I follow the voluntary 8-K classification scheme from the existing literature (e.g. Cooper et al. 2015) in using “Item 2.02 Results of Operations and Financial Conditions”, “Item 7.01 Regulation FD Disclosure” or “Item 8.01 Other Events” to represent discretionary disclosure. However, as Item 2.02 often contains a firm’s earnings announcements, one may argue that Item 2.02 represents more of a regular disclosure policy and may not contain a large discretionary component. Thus I also compute the frequency of discretionary disclosure using only “Item 7.01 Regulation FD Disclosure” or “Item 8.01 Other Events” from 8-K filings as an alternative measure, in attempt to capture the more discretionary disclosures.

In Table 9, I rerun the main analyses in Table 5 using voluntary 8-K filings that contain Item 7.01 and 8.01 only. The coefficients on *peer size* continue to be positive and statistically significant, similar to Table 5. In column (4), I rerun the analysis using the annual total number of voluntary items under Item 7.01 and Item 9.01 of 8-K filings, instead of the annual frequency of voluntary 8-K filings that contain either Item 7.01 or 9.01 as the dependent variable. The results also largely stay the same.

4.5.3 Alternative Measure of Peer Size

As robustness checks, I use an alternative measure of peer size computed using similarity in business segment distribution across firms to rerun the analyses from Table 5 in Table 10. The results remain largely similar to Table 5 but the statistical significance level generally drops across all four columns. This suggests that similarity in business segment provides a coarser identification of peers than similarity in textual business description. In particular, the coefficient on *peer size* in column (2) becomes no longer significant for the frequency of discretionary disclosure with positive news. But the coefficients on *peer size* in other columns still remain statistically significant.

4.5.4 Additional Controls

In untabulated tests, I include additional control variables in my main test on Table 5 for robustness checks, such as an indicator for loss firms, litigation risks proxied by high-risk industry dummies (Francis et al. 1994), proprietary costs proxied by the frequency of words on protected sources of competitive advantage in 10-Ks (e.g. patents, licenses, copyrights), industry concentration proxied by Herfindahl index, the total number of U.S. public and private firms in the firm's primary NAICS 6-digit industry, etc. The results stay largely the same with the addition of these extra controls

5. Concluding Remarks

In this paper, I find that multi-firm dynamics may have a significant influence on the discretionary disclosure behaviors of firms. While the presence of comparable public peers likely expands investors' information endowment over their related exposures, discretionary disclosures by peer firms may also signal information arrival, raising

investors' assessed likelihood that non-disclosure is due to strategic information withholding rather than the absence of new information.

Consistent with this prediction, I show that the bid-ask spread of the underlying firm slightly increases when its closest peer initiates discretionary disclosures, but subsequently decreases if the underlying firms also disclose. This suggests that firms cannot completely free-ride on the information transfer from the discretionary disclosures by peers through non-disclosure. In addition, I find that the number of comparable public peers is strongly positively associated with the frequency of discretionary disclosure using both industry-level and firm-level analyses. As further evidence of peer dynamics, I also show that the discretionary disclosures between the underlying firm and its closest peer are closer in timing if their businesses are more similar to each other.

Overall, my paper provides strong evidence that the presence of peer dynamics induces firms to provide more frequent discretionary disclosures. It complements the first chapter of my dissertation and focuses on the interdependence among similar public firms in discretionary disclosure behaviors. While firms with many public peers may benefit from investors having an expanded information endowment from information spillover, investors also have better knowledge on information arrival for firms with more public peers. The increased perceived probability of information arrival from peer disclosure may make non-disclosure more costly for these firms. To remove the heightened uncertainty among investors whenever their peer firms provide any discretionary disclosure, firms with more public peers are likely induced to provide discretionary disclosures more frequently.

Appendix A. Variable Definitions

Variable	Definition
Measures of Peer Group Size	
<i>industry size (SIC4)</i>	The number of public firms within each SIC 4-digit code
<i>industry size (NAICS6)</i>	The number of public firms within each NAICS 6-digit code
<i>industry size (GICS6)</i>	The number of public firms within each GICS 6-digit code
<i>peer size (textual)</i>	The number of comparable public peers for a firm identified using similarity in textual business description in 10-K filings
<i>peer size (segment)</i>	The number of comparable public peers for a firm identified using similarity in business segment weighting
Dependent Variables	
<i>bid-ask spread</i>	Daily closing bid-ask spread over closing price
<i>trading volume</i>	Daily trading volume
<i>freq.- 8K w/ Item 2.02, 7.01, 9.01</i>	Frequency of discretionary disclosure in a fiscal year measured through the frequency of 8-K filings under Item 2.02, Item 7.01 and Item 9.01
<i>freq.- voluntary 8K w/ positive news</i>	Frequency of discretionary disclosure with positive news in a fiscal year measured through the frequency of 8-K filings under Item 2.02, Item 7.01 and Item 9.01 with positive excess return in the [-1 day,+1 day] window
<i>freq.- voluntary 8K w/ negative news</i>	Frequency of discretionary disclosure with negative news in a fiscal year measured through the frequency of 8-K filings under Item 2.02, Item 7.01 and Item 9.01 with positive excess return in the [-1 day,+1 day] window
<i>freq.- 8K w/ Item 7.01, 9.01</i>	Frequency of discretionary disclosure in a fiscal year measured through the frequency of 8-K filings under Item 7.01 and Item 9.01 only
<i># days apart</i>	The number of days between the discretionary disclosure by the underlying firm and the closest discretionary disclosure by its closest peer
<i># days apart (peer discloses first)</i>	The number of days between the discretionary disclosure by the underlying firm and the closest discretionary disclosure by its closest peer if the closest peer discloses first
<i># days apart (peer discloses next)</i>	The number of days between the discretionary disclosure by the underlying firm and the closest discretionary disclosure by its closest peer if the underlying firm discloses first
Standard Firm Controls	
<i>mb ratio</i>	Market-to-book ratio
<i>stock return</i>	Cumulative stock returns over the fiscal year
<i>stock volatility</i>	Standard deviation of monthly returns over the fiscal year
<i>% inst ownership</i>	% institutional ownership
<i>stock beta</i>	Stock beta
<i>roa ratio</i>	Income before extraordinary items / total assets at the beginning of the year
<i>firm size</i>	Logarithm of total assets
<i>years of listing</i>	Number of years since initial public offering
<i>coverage</i>	Logarithm of (number of analysts + 1)
<i>leverage</i>	Total debt / total equity

Additional Controls	
<i>frequency – 8K w /Item2.01</i>	Frequency of 8-K filings under “Item 2.01 Completion of Acquisition or Disposition of Assets” to proxy for the frequency of M&A activities
<i>frequency- mandatory 8K</i>	Frequency of 8-K filings that do not contain either Item 2.02, Item 7.01 or Item 9.01 to proxy for the frequency of events that trigger mandatory filings
<i>frequency - bundled</i>	Frequency of bundled guidance in a fiscal year
<i>frequency- unbundled</i>	Frequency of unbundled guidance in a fiscal year

Appendix B. Measuring Peer Group Size and Identifying Closest Peers at Firm-Level

1. *Alternative Firm-Level Measure Based on Business Segment Data*

In addition to using primary industry classification to determine the relevant peer group at the industry level, I can utilize finer details of industry classification at the business segment level to determine the relevant peer group with similar business exposures at the firm level, similar to Chapter 1. Under the assumption that business exposures are similar within each business segment as defined by 6-digit NAICS codes, I can use similarities in business segment distribution to proxy for similarities in business exposures.

Operationally, I compute cosine similarity scores across all publicly listed firms on an annual basis using a vector space model, which represents each firm as a vector in an n-dimensional Euclidean space, where n is the number of unique 6-digit NAICS business segment (~1,200) and the value of each vector element is the proportion of sales contribution by each segment of the firm.. Cosine similarity scores are calculated using the following formula:

$$\text{similarity} = \cos(\theta) = \frac{A \cdot B}{\|A\| \|B\|}$$

The size of peer groups based on business segment information can then be represented by the number of close peers with high cosine similarity score. In counting the number of close peers for a given firm using segment information in the additional analysis in Table 10, a similarity score of 0.867, equivalent to $\cos 30^\circ$, is chosen as the similarity score threshold. While the threshold is arbitrarily chosen, robustness checks are done using a variety of thresholds and the results are largely similar.

2. *Primary Firm-Level Measure Based on Textual Business Description*

Similar to the measure based on business segment information, my peer group size measure calculated based on textual business description replaces business segment weights by the relative term frequency in the business description section of 10-K filings in the vector space model. Similar to Chapter 1, I only include the first 2,000 words in the business description for textual analysis as firms tend to put the most essential parts of business description upfront but may often diverge on the latter part of the business description. For example, some firms may present much more details on items such as regulatory environment, management team profiles, etc., in the latter part of the business description section.

In comparing textual similarities across firms, I apply the commonly used tf-idf (term frequency-inverse document frequency) weighting factor, which weighs down the effects of too frequently occurring terms and weighs up the effects of less frequently occurring terms⁴⁶. This is because

⁴⁶ Brown and Tucker (2010) also uses the tf-idf weight scheme in comparing within-firm MD&A changes. Instead of using tf-idf weighting factor, Hoberg and Philip (2015) only considers the binary occurrence of nouns and pronouns in the whole section of business description as it attempts to focus more on capturing product similarity. Similar to Hoberg and Philip (2015) using a 25% threshold (excluding 75% most common nouns and proper nouns) in calculating product similarities among firms, I also use a stopword threshold of 25% (excluding the 75% most common terms) but exclude extremely rare words/phrases that only occur once across all descriptions in calculating the similarity score in my main measure. As tf-idf weighting scheme is being used, my measure is not sensitive to the choice of stopword threshold. Another

most common terms are unlikely providing much useful information in differentiating among firms. Mathematically, each business description in the corpus d_i is represented by a vector in the word space considering the distinctive power of each word (idf), where tf_{ij} represents the term frequency of word j for firm i , idf_j represents the inverse document frequency of word j and M is the number of unique words in the entire corpus.

$$d_i = \{tf_{i1} * idf_1, \dots, tf_{ij} * idf_j, tf_{iM} * idf_M\}$$

The size of peer groups based on textual business description can then be similarly represented by the number of close peers with high cosine similarity score. In calculating the size of peer groups, I count the number of close peers for a given firm using a similarity score threshold that defines 25% of the firms with no comparable peer. While the threshold is arbitrarily chosen, robustness checks are done using a variety of thresholds and the results are largely similar.

Instead of assuming that all operations within a 6-digit NAICS business segment are homogeneous, this measure attempts to expand the scope of information set in considering the similarities across firms by utilizing the textual business description information in 10-K filings. As business section is part of mandatory disclosure governed by regulation S-K and contains mostly factual information, it is unlikely that certain firms would strategically choose words that are common or uncommon to describe their businesses, causing systematic biases in the measure⁴⁷.

3. *Identifying Closest Peers Based on Textual Business Description*

In Tables 3, 6 & 7, I determine a firm's closest peers by those firms that have the highest cosine similarity scores with the underlying firm in the vector space model using textual business description. Thus a firm's closest peer is identified as the firm which has the highest similarity in 10-K business description with the underlying firm. For example, following my method in 2012, the closest peer of Burger King Worldwide is identified as Carrols Restaurant Group; the closest peer of Las Vegas Sands is identified as Wynn Resorts; and the closest peer of Orbitz is identified as Priceline.

major difference is that I use all pre-processed meaningful words/phrases within the first 2,000 words of business description. The results of the paper are robust to different combinations of parameters in stopword thresholds and the number of words retained from business description section.

⁴⁷ There is fairly detailed framework under regulation S-K on the content of business description in the 10-K filings, with over 3,800 words of extensive guidance, e.g. on products and services, raw materials, seasonality, intangibles, working capital, customers, order backlogs, competition, research and development, regulation, number of employees, etc. However, it is possible that textual analysis based on business description section may suffer some noise in measurement due to differential writing styles by different firms.

Appendix C. Examples of Voluntary 8-K Filings

Microsoft

Item 7.01 Regulation FD Disclosure

On July 8, 2015, Microsoft Corporation (the “Company”) announced plans to restructure the company’s phone hardware business to better focus and align resources. The Company also announced the reduction of up to 7,800 positions, primarily in the phone business. As a result, in the quarter ending June 30, 2015 the Company will record a charge of approximately \$7.6 billion for impairment of goodwill and assets in its Phone Hardware segment related to assets associated with the acquisition of the Nokia Devices and Services business. This charge has no impact on cash flow from operations and is non-deductible for tax purposes. Based on the new plans, the future prospects for the Phone Hardware segment are below original expectations. The Company will record a restructuring charge of approximately \$750 million to \$850 million in connection with the plan. The actions associated with the plan are expected to be substantially completed by the end of the calendar year and fully completed by the end of the 2016 fiscal year.

A copy of Microsoft’s press release is provided as Exhibit 99.1 to this report.

Item 8.01 Other Events.

On May 10, 2011, Microsoft Corporation (the “Company”) and Skype Global S.à r.l. (“Skype”) issued a joint press release announcing that the Company had entered into a purchase agreement to acquire Skype for \$8.5 billion in cash. A copy of the press release is attached hereto as Exhibit 99.1.

Walmart

Item 7.01 Regulation FD Disclosure

Wal-Mart Stores, Inc. (the “company”) today announced that John E. Fleming, executive vice president, chief merchandising officer, Walmart U.S. will be leaving the company on August 1, 2010.

Item 8.01 Other Events

The attached press release from our Seiyu business in Japan discusses business strategies and also reflects that approximately 30 non-performing stores in Japan will be closed in the future. While Wal-Mart Stores, Inc. (“Walmart” or the “company”) is in the process of reviewing and assessing the anticipated charges associated with these store closures, it currently estimates that charges from the closures in the aggregate will range from approximately \$0.04 to \$0.05 of diluted earnings per share from continuing operations. It is currently anticipated that these charges will be recorded over the next several quarters, beginning with the company’s fiscal 2015 fourth quarter. The company’s operations in Japan, like its other international operations, other than Canada, consolidate generally using a one-month lag and are based on a calendar year.

Table 1 Summary Statistics

The table below shows the summary descriptive statistics for the main variables in various tables in the paper.

VARIABLES	N	mean	sd	p25	p50	p75
TABLE 3						
<i>bid-ask spread</i>	401,693	0.828	2.459	0.065	0.187	0.663
<i>trading volume</i>	402,042	2288	8801	64.4	299.7	1210
TABLE 4						
<i>avg. freq.- voluntary 8K(SIC4)</i>	7,682	5.216	3.639	2.48	4.942	6.833
<i>industry size (SIC4)</i>	7,682	13.95	58.25	2	5	11
<i>avg. freq.- voluntary 8K (NAICS6)</i>	13,354	5.169	4.030	2	4.5	7
<i>industry size (NAICS6)</i>	13,354	6.573	20.53	1	2	5
<i>avg. freq.- voluntary 8K (GIC4)</i>	1,327	5.493	3.246	2.697	5.470	7.067
<i>industry size (GICS6)</i>	1,327	65.36	68.41	18	39	98
TABLES 5, 8, 9, 10						
<u><i>Frequency of Discretionary Disclosure</i></u>						
<i>freq.- 8K w/ Items 2.02, 7.01, 9.01</i>	33,277	6.118	4.504	4	5	8
<i>freq.- voluntary 8K w/ positive news</i>	33,277	2.975	2.589	1	3	4
<i>freq.- voluntary 8K w/ negative news</i>	33,277	3.144	2.815	1	3	4
<i>freq.- 8K w/ Items 7.01, 9.01</i>	33,277	4.344	4.925	1	3	6
<u><i>Size of Peer Group</i></u>						
<i>peer size (textual)</i>	33,277	5.061	7.547	0	2	6
<i>peer size (segment)</i>	33,277	58.46	84.74	4	18	81
<u><i>Controls</i></u>						
<i>frequency – 8K w/ item2.01</i>	33,277	0.246	0.639	0	0	0
<i>frequency- mandatory 8K</i>	33,277	3.159	3.101	1	3	5
<i>frequency - bundled</i>	33,277	1.598	1.838	0	0	4
<i>frequency- unbundled</i>	33,277	0.580	1.140	0	0	1
<i>inst. ownership</i>	33,277	0.574	0.302	0.323	0.624	0.823
<i>analyst coverage</i>	33,277	1.475	0.970	0.693	1.609	2.197
<i>years of listing</i>	33,277	17.42	15.01	7.036	13.14	23.00
<i>firm size</i>	33,277	6.199	2.011	4.758	6.141	7.550
<i>stock return</i>	33,277	17.79	69.87	-21.01	7.510	38.47
<i>roa ratio</i>	33,277	0.0160	1.906	-0.0196	0.0331	0.0760
<i>stock beta</i>	33,277	1.351	0.979	0.683	1.186	1.822
<i>mb ratio</i>	33,277	3.125	3.803	1.244	1.998	3.429
<i>stock volatility</i>	33,277	0.133	0.0810	0.0772	0.113	0.164
<i>leverage</i>	33,277	0.910	2.179	0.0076	0.272	0.812
TABLES 6,7						
<i># days apart</i>	259,388	18.58	18.40	5	13	28
<i># days apart (peer discloses first)</i>	121,952	20.06	18.72	6	14	29
<i># days apart (peer discloses next)</i>	137,436	17.26	18.00	4	12	26

Table 2 Correlation Table

The table below shows the correlation table between the frequency of voluntary 8K filings (*freq vol. 8k*) under Item 2.02, Item 7.01 and Item 9.01 and the number of comparable peers identified using textual similarity in business description (*peer size*) as well as other standard firm-level controls.

	<i>freq vol. 8k</i>	<i>peer size</i>	<i>inst. owned</i>	<i>analyst cov.</i>	<i>years of listing</i>	<i>firm size</i>	<i>stock return</i>	<i>roa ratio</i>	<i>stock beta</i>	<i>mb ratio</i>	<i>stock vol.</i>	<i>lev</i>
<i>freq. voluntary 8k</i>	1											
<i>peer size</i>	0.1211	1										
<i>inst. ownership</i>	0.1868	0.0514	1									
<i>analyst coverage</i>	0.2243	0.1556	0.5931	1								
<i>years of listing</i>	0.0443	-0.1097	0.0731	0.0381	1							
<i>firm size</i>	0.2772	0.0994	0.5813	0.6142	0.2718	1						
<i>stock return</i>	0.031	0.0042	-0.0142	-0.0501	-0.0155	-0.0226	1					
<i>roa ratio</i>	-0.0076	-0.0049	0.0014	-0.004	0.0263	0.0002	0.017	1				
<i>stock beta</i>	0.0327	0.0385	0.0337	0.0934	-0.1448	-0.0682	0.0591	-0.0365	1			
<i>mb ratio</i>	0.0338	0.0069	-0.013	0.0899	-0.0291	-0.1128	0.1855	0.0858	0.0305	1		
<i>stock volatility</i>	-0.0486	0.0189	-0.2569	-0.1707	-0.2044	-0.3434	0.1811	-0.0541	0.4354	0.0334	1	
<i>leverage</i>	0.102	0.0352	0.0203	-0.0043	-0.0022	0.2159	-0.0083	-0.01	0.0111	0.3078	0.038	1

Table 3 Changes in Bid-Ask Spread When Closest Peer Discloses

The table below examines how the daily closing bid-ask spread of the underlying firm changes when its closest peer provides discretionary disclosure as well as when the underlying firm also subsequently provides discretionary disclosure within 7 days after its peer disclosure. Sample includes any voluntary 8-K filings by the closest peer (under items 2.02, 7.01 & 8.01) that are identified with the following event types after matching with Capital IQ database: strategic alliances, client announcements, corporate guidance, business expansions, business reorganizations, product announcements, conference calls. The date of the earliest report of the peer disclosure identified from the 8-K filing, instead of the filing date of 8-K, is used as the date of the event. The sample includes observations [-7 days, +7 days] around the event date and contains only peer disclosure events that are not preceded by the disclosures of the underlying firm. I code *event date* as an indicator for the date of the earliest report of the discretionary disclosure by the closest peer, *after event* as an indicator for dates following the event date [+1 day, +7 days], *post event* as an indicator for dates on and following the event date [0, +7 days]. In addition, I code *disclose date (d)* as an indicator for the date of the earliest report of the discretionary disclosure by the underlying firm if it is within 7 days after the event date, *after disclosure* as an indicator for dates following the disclosure date [*d*+1, +7 days], *post event* as an indicator for dates on and following the disclosure date [*d*, +7 days].

VARIABLES (predicted sign)	(1) Bid-Ask Spread	(2) Bid-Ask Spread	(3) Bid-Ask Spread	(4) Trading Volume
<i>post event</i> [0, +7 days] (+)	0.010 (1.18)	0.088* (1.80)		
<i>event date</i> [day 0] (+)			0.091** (2.17)	141.304*** (4.10)
<i>after event</i> [+1 day, +7 days] (+)			0.084 (1.62)	
<i>post disclosure</i> [+ <i>d</i> , +7 days] (-)		-0.083* (-1.79)		
<i>disclose date</i> [day <i>d</i>] (-)			-0.093* (-1.70)	1,580.829*** (3.15)
<i>after disclosure</i> [+ (<i>d</i> +1) day, +7 days] (-)			-0.079* (-1.66)	
Observations	401,693	401,693	401,693	401,693
R-squared	0.615	0.615	0.615	0.787
Event FE	YES	YES	YES	YES

Event and day clustered standard errors in parenthesis *** p<0.01, ** p<0.05, * p<0.1

Table 4 Industry Size and Frequency of Discretionary Disclosure

The table below examines how the average frequency of discretionary disclosure under items 2.02, 7.01 & 8.01 per firm within an industry varies with the number of firms within an industry (*industry size*) using various industry classification schemes. Columns (1) and (2) use the NAICS 6-digit classification scheme. Columns (3) and (4) follow the SIC 4-digit classification scheme. Columns (5) and (6) use the GIC 6-digit classification scheme.

VARIABLE S (predicted sign)	<u>NAICS 6-Digit Classification</u>		<u>SIC 4-Digit Classification</u>		<u>GIC 6-Digit Classification</u>	
	(1)	(2)	(3)	(4)	(5)	(6)
	<i>Avg. Freq. of Discretionary Disclosure Per Firm In Industry</i>	<i>Avg. Freq. of Discretionary Disclosure Per Firm In the Industry</i>	<i>Avg. Freq. of Discretionary Disclosure Per Firm In Industry</i>	<i>Avg. Freq. of Discretionary Disclosure Per Firm In Industry</i>	<i>Avg. Freq. of Discretionary Disclosure Per Firm In Industry</i>	<i>Avg. Freq. of Discretionary Disclosure Per Firm In Industry</i>
<i>industry size</i> (+)	0.023*** (5.08)	0.017*** (4.37)	0.021*** (4.91)	0.015*** (3.70)	0.017*** (4.33)	0.009** (1.97)
Observations	13,354	13,354	7,662	7,662	1,327	1,327
R-squared	0.013	0.267	0.026	0.362	0.123	0.539
Year FE	NO	YES	NO	YES	NO	YES

Industry and year clustered standard errors in parenthesis *** p<0.01, ** p<0.05, * p<0.1

Table 5 Comparable Peer Size and Frequency of Discretionary Disclosure

The table below examines how the frequency of discretionary disclosure of firms varies with the number of comparable public peers (*peer size*). Column (1) uses the frequency of voluntary 8-K filings under items 2.02, 7.01 & 8.01 as dependent variables. Columns (2) and (3) disaggregate total frequency of voluntary 8-K filings into those with positive news and negative news respectively. Column (4) uses the frequency of unbundled guidance as dependent variable among a sample of guiding firms.

VARIABLES (predicted sign)	(1) Frequency of Voluntary 8K Filings	(2) Frequency of Voluntary 8K Filings w/ Positive News	(3) Frequency of Voluntary 8K Filings w/ Negative News	(4) Frequency of Unbundled Guidance
<i>peer size</i> (+)	0.049*** (6.38)	0.023*** (5.79)	0.027*** (5.56)	0.015*** (4.17)
<i>frequency - bundled</i>	-0.117*** (-4.02)	-0.066*** (-4.11)	-0.050*** (-2.82)	
<i>frequency - unbundled</i>	0.203*** (4.14)	0.133*** (4.78)	0.070*** (2.85)	
<i>inst. ownership</i>	0.103 (0.33)	0.108 (0.84)	-0.005 (-0.02)	0.067 (0.89)
<i>analyst coverage</i>	0.317*** (3.76)	0.358*** (5.90)	-0.040 (-0.63)	0.328*** (8.21)
<i>years of listing</i>	-0.003 (-0.55)	-0.001 (-0.17)	-0.002 (-0.82)	0.006*** (3.49)
<i>firm size</i>	0.532*** (10.65)	0.165*** (5.10)	0.368*** (10.57)	0.078*** (3.45)
<i>stock return</i>	-0.000 (-1.28)	0.004*** (3.82)	-0.005*** (-5.23)	0.001*** (5.31)
<i>roa ratio</i>	-0.018* (-1.79)	-0.003 (-0.30)	-0.015*** (-3.32)	0.301*** (2.60)
<i>stock beta</i>	-0.049 (-1.13)	-0.055** (-1.99)	0.005 (0.17)	-0.025* (-1.69)
<i>mb ratio</i>	0.027 (1.59)	0.010 (1.02)	0.017* (1.89)	0.006 (1.22)
<i>stock volatility</i>	4.838*** (6.53)	1.332*** (3.16)	3.506*** (7.21)	0.054 (0.15)
<i>leverage</i>	0.078* (1.70)	0.053** (2.17)	0.025 (1.02)	-0.012 (-1.20)
Observations	33,277	33,277	33,277	16,631
R-squared	0.330	0.249	0.261	0.097
Year FE	YES	YES	YES	YES

Firm and year clustered standard errors in parenthesis *** p<0.01, ** p<0.05, * p<0.1

Table 6 Timing of Discretionary Disclosure Using Cox Hazard Model***Panel A: Peer Similarity and Clustering of Discretionary Disclosures***

The table below presents the Cox proportional hazard model on the number of days apart between discretionary disclosure of the underlying firm and the closest discretionary disclosure by its peer firm (either before or after the underlying firm). Hazard is denoted by peer disclosure. Columns (1) to (4) present the results for the closest peer to the 4th closet peer respectively.

SAMPLES VARIABLES (predicted sign)	(1) <u>Closest Peer</u> Cox Hazard Model Coefficient	(2) <u>2nd Closest Peer</u> Cox Hazard Model Coefficient	(3) <u>3rd Closest Peer</u> Cox Hazard Model Coefficient	(3) <u>4th Closest Peer</u> Cox Hazard Model Coefficient
<i>Similarity Decile</i> (+)	0.159*** (11.26)	0.183*** (13.33)	0.170*** (12.56)	0.173*** (12.64)
Observations	245,202	245,507	245,621	244,379
Year FE	YES	YES	YES	YES

Firm clustered standard errors in parenthesis *** p<0.01, ** p<0.05, * p<0.1

Panel B: Similarity and Clustering of Discretionary Disclosures if Peer Discloses First

The table below presents the Cox proportional hazard model on the number of days apart between discretionary disclosure of the underlying firm and the closest discretionary disclosure by its peer firm, conditional on peer discloses before underlying firm.

SAMPLES VARIABLES (predicted sign)	(1) <u>Closest Peer</u> Cox Hazard Model Coefficient	(2) <u>2nd Closest Peer</u> Cox Hazard Model Coefficient	(3) <u>3rd Closest Peer</u> Cox Hazard Model Coefficient	(3) <u>4th Closest Peer</u> Cox Hazard Model Coefficient
<i>Similarity Decile</i> (+)	0.161*** (9.65)	0.199*** (12.29)	0.169*** (10.59)	0.183*** (11.33)
Observations	121,952	122,628	123,000	122,143
Year FE	YES	YES	YES	YES

Firm clustered standard errors in parenthesis *** p<0.01, ** p<0.05, * p<0.1

Panel C: Similarity and Clustering of Discretionary Disclosures if Peer Discloses Last

The table below presents the Cox proportional hazard model on the number of days apart between discretionary disclosure of the underlying firm and the closest discretionary disclosure by its peer firm, conditional on peer discloses after underlying firm.

SAMPLES VARIABLES (predicted sign)	(1) <u>Closest Peer</u> Cox Hazard Model Coefficient	(2) <u>2nd Closest Peer</u> Cox Hazard Model Coefficient	(3) <u>3rd Closest Peer</u> Cox Hazard Model Coefficient	(3) <u>4th Closest Peer</u> Cox Hazard Model Coefficient
<i>Similarity Decile</i> (+)	0.158*** (9.51)	0.167*** (10.38)	0.171*** (10.79)	0.162*** (10.10)
Observations	123,250	122,879	122,621	122,236
Year FE	YES	YES	YES	YES

Firm clustered standard errors in parenthesis *** p<0.01, ** p<0.05, * p<0.1

Table 7 Peer Similarity and Timing of Discretionary Disclosure (OLS)

The table below examines how the similarity with peer relates to the number of days apart between discretionary disclosure of the underlying firm and the closest discretionary disclosure by its peer firm (either before or after the underlying firm) using OLS regression. Columns (1) to (4) present the results from the closest peer to the 4th closet peer for each underlying firm respectively.

SAMPLES	(1) Closest Peer	(1) 2 nd Closest Peer	(1) 3 rd Closest Peer	(1) 4 th Closest Peer
VARIABLES (predicted sign)	#Days Apart From Peer's Closest Discretionary Disclosure	#Days Apart From Peer's Closest Discretionary Disclosure	#Days Apart From Peer's Closest Discretionary Disclosure	#Days Apart From Peer's Closest Discretionary Disclosure
<i>Similarity Decile</i>	-3.525***	-3.494***	-3.455***	-3.613***
(-)	(-12.01)	(-11.99)	(-12.42)	(-13.26)
<i>inst. ownership</i>	0.263	-0.060	0.132	-0.655**
	(0.64)	(-0.18)	(0.40)	(-2.01)
<i>analyst coverage</i>	-0.311**	-0.091	-0.263**	-0.064
	(-2.40)	(-0.70)	(-2.05)	(-1.04)
<i>firm size</i>	-0.290***	-0.337***	-0.203***	-0.257**
	(-4.30)	(-5.30)	(-3.29)	(-2.10)
<i>stock return</i>	-0.003**	-0.000	-0.000	-0.001
	(-2.46)	(-0.41)	(-0.30)	(-0.92)
<i>mb ratio</i>	-0.005	-0.006	0.007	0.017
	(-0.21)	(-0.23)	(0.28)	(0.78)
<i>roa ratio</i>	-0.966***	-0.652***	-0.872***	-0.633**
	(-2.92)	(-3.09)	(-3.65)	(-2.39)
<i>years of listing</i>	0.026***	0.035***	0.023***	0.022***
	(3.94)	(6.03)	(3.98)	(3.91)
<i>stock beta</i>	-0.026	0.011	-0.042	-0.092
	(-0.32)	(0.14)	(-0.52)	(-1.17)
<i>stock volatility</i>	0.487	1.292	1.170	1.886*
	(0.44)	(1.11)	(1.00)	(1.76)
<i>leverage</i>	-0.097**	-0.037	-0.067	-0.115***
	(-2.51)	(-0.92)	(-1.64)	(-3.24)
Observations	213,586	212,386	211,882	210,736
R-squared	0.061	0.060	0.062	0.065
Year FE	YES	YES	YES	YES

Firm clustered standard errors in parenthesis *** p<0.01, ** p<0.05, * p<0.1

Table 8 Controlling for Frequency of Events

The table below re-examines column (1) of Table 5 attempting to control the frequency of events. Column (1) adds an additional control for corporate transactions by including the frequency of 8K filings with “Item 2.01 Completion of Acquisition or Disposition of Assets”. Column (2) adds an additional control for mandatory disclosure by including the frequency of 8K filings that do not contain Items 2.02, 7.01 or 9.01. Column (3) adds additional controls for firm complexity by including the number of industrial segments and the number of geographical segments.

VARIABLES (predicted sign)	(1) Frequency of Voluntary 8K Filings	(2) Frequency of Voluntary 8K Filings	(3) Frequency of Voluntary 8K Filings
<i>peer size (+)</i>	0.049*** (6.43)	0.049*** (6.39)	0.045*** (4.28)
<i>frequency -8Ks w/ item 2.01</i>	0.334*** (4.93)		
<i>frequency- non-voluntary 8Ks</i>		0.080*** (4.36)	
<i># industrial segments</i>			0.040 (0.86)
<i># geographical segments</i>			-0.075*** (-2.58)
<i>frequency - bundled</i>	-0.118*** (-4.08)	-0.115*** (-3.98)	-0.094*** (-3.34)
<i>frequency - unbundled</i>	0.201*** (4.10)	0.199*** (4.05)	0.204*** (3.96)
<i>inst. ownership</i>	0.102 (0.33)	0.087 (0.28)	0.123 (0.41)
<i>analyst coverage</i>	0.331*** (4.01)	0.311*** (3.74)	0.259*** (2.86)
<i>years of listing</i>	-0.002 (-0.35)	-0.002 (-0.40)	-0.001 (-0.18)
<i>firm size</i>	0.525*** (10.61)	0.518*** (10.14)	0.575*** (10.53)
<i>stock return</i>	-0.001 (-1.64)	-0.000 (-0.82)	-0.000 (-0.48)
<i>roa ratio</i>	-0.018* (-1.78)	-0.015 (-1.33)	-0.012 (-1.26)
<i>stock beta</i>	-0.051 (-1.17)	-0.051 (-1.21)	-0.083 (-1.56)
<i>mb ratio</i>	0.029* (1.72)	0.027 (1.64)	-0.000 (-0.01)
<i>stock volatility</i>	4.817*** (6.58)	4.443*** (5.38)	4.945*** (6.03)
<i>leverage</i>	0.075 (1.64)	0.073 (1.58)	0.107* (1.91)

Observations	33,277	33,277	25,384
R-squared	0.157	0.157	0.157
Year FE	YES	YES	YES
Firm and year clustered standard errors in parenthesis *** p<0.01, ** p<0.05, * p<0.1			

Table 9 Using Alternative Measure of Discretionary Disclosure

The table below examines how the frequency of discretionary disclosure of firms varies with the number of their comparable peers (*peersize*). Column (1) uses the frequency of voluntary 8-K filings under items 7.01 & 8.01 as dependent variables. Columns (2) and (3) disaggregate total frequency voluntary 8-K under items 7.01 & 8.01 filings into those with positive news and negative news respectively. Column (4) uses the annual total number of voluntary items (items 7.01 & 8.01) in 8-K filings as dependent variable conditional on a sample of guiding firms.

VARIABLES (predicted sign)	(1) Frequency of Voluntary 8K Filings	(2) Frequency of Voluntary 8K Filings w/ Positive News	(3) Frequency of Voluntary 8K Filings w/ Negative News	(4) Total Number of Voluntary Items in 8-K Filings
<i>peer size</i> (+)	0.056*** (5.87)	0.026*** (5.73)	0.030*** (5.25)	0.051*** (6.23)
<i>frequency - bundled</i>	-0.153*** (-4.74)	-0.087*** (-4.77)	-0.066*** (-3.77)	-0.111*** (-3.18)
<i>frequency - unbundled</i>	0.175*** (3.16)	0.104*** (3.40)	0.071*** (2.60)	0.270*** (5.94)
<i>inst. ownership</i>	-0.416 (-1.19)	-0.151 (-1.04)	-0.265 (-1.01)	0.528 (1.40)
<i>analyst coverage</i>	0.132 (1.37)	0.199*** (3.52)	-0.067 (-1.02)	0.326*** (2.67)
<i>years of listing</i>	-0.007 (-1.16)	-0.002 (-0.50)	-0.005* (-1.70)	-0.016*** (-2.82)
<i>firm size</i>	0.703*** (13.23)	0.274*** (9.76)	0.428*** (12.22)	0.756*** (12.69)
<i>stock return</i>	-0.001 (-1.04)	0.003*** (3.91)	-0.003*** (-8.94)	-0.003*** (-3.63)
<i>roa ratio</i>	-0.029* (-1.75)	-0.013** (-2.08)	-0.017 (-1.55)	-0.063*** (-2.88)
<i>stock beta</i>	-0.095* (-1.90)	-0.052* (-1.85)	-0.044 (-1.58)	-0.020 (-0.33)
<i>mb ratio</i>	0.027 (1.63)	0.009 (1.02)	0.018** (2.06)	0.019 (1.03)
<i>stock volatility</i>	6.834*** (10.01)	2.834*** (10.31)	4.000*** (9.06)	11.146*** (8.09)
<i>leverage</i>	0.117** (2.51)	0.068*** (2.69)	0.050** (2.11)	0.131*** (3.12)
Observations	33,277	33,277	33,277	33,277
R-squared	0.106	0.087	0.089	0.330
Year FE	YES	YES	YES	YES

Firm and year clustered standard errors in parenthesis *** p<0.01, ** p<0.05, * p<0.1

Table 10 Using Alternative Measures of Peer Size

The table below re-examines Table 5 on the relationship between the frequency of discretionary disclosure of firms and the number of their comparable peers identified using business segment data (*segment peer size*) instead of using textual analysis on business description. Column (1) uses the frequency of voluntary 8-K filings under items 2.02, 7.01 & 8.01 as dependent variables. Columns (2) and (3) disaggregate total frequency voluntary 8-K filings into those with positive news and negative news respectively. Column (4) uses the frequency of unbundled guidance as dependent variable conditional on a sample of guiding firms.

VARIABLES (predicted sign)	(1) Frequency of Voluntary 8K Filings	(2) Frequency of Voluntary 8K Filings w/ Positive News	(3) Frequency of Voluntary 8K Filings w/ Negative News	(4) Frequency of Unbundled Guidance
<i>segment peer size</i> (+)	0.001** (2.07)	0.000 (0.16)	0.001** (2.42)	0.001** (1.96)
<i>frequency - bundled</i>	-0.149*** (-4.87)	-0.081*** (-4.81)	-0.068*** (-3.71)	
<i>frequency - unbundled</i>	0.211*** (4.19)	0.137*** (4.83)	0.074*** (2.89)	
<i>inst. ownership</i>	0.086 (0.28)	0.090 (0.69)	-0.003 (-0.01)	0.073 (0.97)
<i>analyst coverage</i>	0.378*** (4.58)	0.392*** (6.48)	-0.014 (-0.23)	0.340*** (8.51)
<i>years of listing</i>	-0.005 (-1.02)	-0.002 (-0.64)	-0.003 (-1.17)	0.006*** (3.28)
<i>firm size</i>	0.553*** (11.15)	0.172*** (5.37)	0.381*** (10.73)	0.083*** (3.41)
<i>stock return</i>	-0.000 (-0.91)	0.004*** (3.89)	-0.005*** (-5.15)	0.001*** (5.40)
<i>roa ratio</i>	-0.016 (-1.58)	-0.003 (-0.28)	-0.014*** (-2.85)	0.273** (2.41)
<i>stock beta</i>	-0.063 (-1.28)	-0.055* (-1.83)	-0.008 (-0.24)	-0.031** (-2.26)
<i>mb ratio</i>	0.022 (1.32)	0.009 (0.91)	0.013 (1.52)	0.003 (0.54)
<i>stock volatility</i>	4.929*** (6.81)	1.394*** (3.32)	3.536*** (7.54)	0.104 (0.29)
<i>leverage</i>	0.084* (1.86)	0.055** (2.24)	0.030 (1.24)	-0.007 (-0.69)
Observations	33,277	33,277	33,277	16,631
R-squared	0.149	0.124	0.112	0.094
Year FE	YES	YES	YES	YES

Firm and year clustered standard errors in parenthesis *** p<0.01, ** p<0.05, * p<0.1

CHAPTER 3

Searching for Close Peers: A Comparison of Different Methods

1. Introduction

Industry practitioners and academic researchers are often facing the task of finding appropriate peer firms for benchmarking or other purposes. Industry classification has long been used as an important input in identifying a firm's peers in the capital market. Some notable ones include the early Standard Industry Classification (SIC) system, Fama-French 48 industry classification (1997), North America Industry Classification (NAIC) system (e.g. Krishnan and Press 2003), and Global Industry Classification (GIC) system (e.g. Bhojraj et al. 2003), textual product-based classification by Hoberg and Philips (2010, 2015), etc. However, industry classification schemes alone often have their limitations in finding close peers as there tends to be heterogeneity within each industry classification and some industries may have rather coarse classification. For example, 4-digit SIC on average has more than 70 firms in each 4-digit classification and the largest number of firms within a 4-digit SIC classification can reach to over 300.

Beyond industry classification, we need some metrics to measure the degree of similarity between firms in our search for close comparable peers. In various contexts, some prior papers have attempted to use historical correlation in returns or earnings as a measure of similarity between firms or homogeneity among a group of firms. For example, Pyo and Lustgarten (1990) used the historical correlation in annual earnings and Parrino (1997) used return-based correlation to measure industry homogeneity within 2-

digit SIC industries. More recently, Lee et al. (2015) even tried to use the intensity of internet co-searches as a measure of similarity between firms. While a variety of methods are being used by different papers, it is not clear how well each method performs and compares against one another.

In the paper, I attempt to compare various methods that can be easily implemented to measure similarity between firms in the search for close comparable peer firms. Besides the method that uses the textual similarity in business description of 10-K filings as described from my earlier two chapters, I also examine various correlation-based methods that utilize key historical market and financial information. These include using correlation in daily returns over the past 1 year, correlation in monthly returns over the past 5 years, correlation in accounting earnings over the past 5 years, as well as correlation in operating cash flows over the past 5 years. In addition, I also include various commonly used industry classifications in my comparison tests, such as SIC, NAIC, GIC, etc.

To compare the performance of each method in identifying close peers, I implement three sets of comparison tests. In my first test, among a subsample of firms with high analyst following, I examine how the top peer firm identified under each method matches with a benchmark sample of highly probable peer pairs identified by using analyst coverage. My second test compares how each method performs in terms of the future comovement in daily returns, monthly returns, accounting earnings and operating cash flows between the top peer firm identified under each method and its base firm. In my third test, I compare the explanatory power of using the key valuation multiples and accounting ratios of the top peer firm identified under each method to explain the

corresponding contemporaneous key valuation multiples and accounting ratios of the base firm.

My results show that my approach in using the textual similarity in business description performs consistently well in various comparison tests. Besides that, the method using historical comovement in daily returns also performs well in most comparison tests. Among others, these two metrics appear to provide the best measures of similarity in business exposure between firms. On the contrary, methods using historical comovement in either earnings or operating cash flows alone perform poorly in various comparison tests, likely due to the lack of sufficient variations with a quarterly data frequency in distinguishing among a large number of firms. But I also find that combining historical comovement in earnings or operating cash flows with industry classification can significantly help improve their performance in identifying close peer firms by limiting the search space to fewer firms.

Not surprisingly, I show that finer industry classification generally performs better than coarser industry classification in identifying close peers. For example, 8-digit GICS generally outperforms 6-digit GICS and 4-digit SIC generally outperforms Fama-French 48 industry classification. Among various industry classification schemes being tested, 8-digit GICS, 8-digit NAICS and 4-digit SIC generally perform better than the others. Across my sample, however, I do not find that GIC performs significantly better than NAICS or SIC as found by Bhojraj et al. (2003). This may be due to continuous improvement in the NAICS system over time as my sample covers a later sample period relative to Bhojraj et al. (2003).

In further analyses, I find that there are indeed large variations in the similarity between the base firms and their closest peers and the degree of similarity is significantly positively associated with the future comovement in returns, earnings and operating cash flows. These translate to large variations in the future comovement in returns, earnings and operating cash flows across the firm-peer pairs. This means that, for some firms that show low comovement in returns, earnings and operating cash flows even with their closest peers, it will be more difficult for investors to use the information from other peer firms to learn about the base firm. This is related to the issue of “informational uniqueness” described in Chapter 1 of this dissertation.

My study contributes to the growing literature on identifying comparable peers. Motivated by earlier studies that use different types of information in measuring similarity between firms, I focus on utilizing and testing several easy-to-implement methods that involve readily accessible data, e.g. correlation-based methods using historical returns, earnings and cash flows as well as textual similarity of business description in 10-K filings. I find that using textual similarity in business description and historical comovement in daily returns are the best two methods in identifying close peer firms.

To the best of my knowledge, the study is also the first to compare how the historical comovement in various market returns and accounting information between close peer firms may translate into future comovement in market returns and accounting information. The results suggest that comovement of historical daily returns works the best in finding peers with the highest future comovement in both monthly returns and daily returns. However, using the firm-pairs’ own historical comovement in earnings /

operating cash flows respectively perform the best in finding peers with the highest future comovement in earnings / operating cash flows.

The remainder of the paper is organized as follows. Section 2 reviews the relevant literature and Section 3 provides an overview of different methodology in identifying peer firms. Section 4 describes my tests of comparison and Section 5 presents the results of comparison and other additional analyses. Section 6 concludes.

2. Literature Review

There has been a lot of demand in both academia and industry to search for peer firms. For example, relative valuation methodologies usually involve finding a set of comparable peer firms. The value of an asset is then compared to the value assessed by the market for these comparable assets. These comparable peer firms are also useful for benchmarking analysis. In most sell-side research reports, we may find analysts using peers in their research reports to compare performance and valuation across firms and estimate the value of the firms they cover (De Franco, et al 2015). In “pairs trading” strategy, traders also need to identify pairs of closely comparable firms whose prices should co-move in the future. In academia, researchers often need to find the most comparable firm to be used as a control firm in certain types of analyses, e.g. difference-in-difference analysis.

Industry classification has been an important input in identifying peers, especially in earlier academic works. It started with the commonly used SIC system, which was introduced in 1937 with the latest modification done in 1987. Later, NAIC system was developed and subsequently adopted by the U.S. Office of Management and Budget for

use by all governmental agencies in 1997. These two industry classification schemes are developed by government agencies mainly based on production and supply chain processes. Academics have also tested and used the GIC system, which is developed by Standard & Poor and Morgan Stanley Capital International on the basis of principal business activity. Bhojraj et al. (2003) find that GICS generally outperform these other two industry classification schemes. In addition to GICS, there are also several other industry classification schemes developed by private companies, such as Industry Classification Benchmark (ICB) by FTSE and Dow Jones Industry Classification System (DJICS), etc.

Academics have also attempted to use different information disclosed by firms to reclassify into alternative groupings. For example, Baginski (1987) used clustering algorithm based on business and financial risk measures to refine groups of firms that are similar within 4-digit SIC industries. In recent years, Hoberg and Philips (2010, 2015) develop an alternative method in grouping firms based on textual analysis on the similarity among firms' self-reported business description in their 10-K filings to infer their product-market peers. In addition, there are several studies that attempt to utilize the judgement by informed parties in determining peer firms. Ramnath (2002) utilizes the choice of firm coverage by analysts in determining peer firms and categorizes firms as peers if at least five analysts covered the same group of firms. De Franco et al. (2015) directly hand-collect the peer firm selection by analysts inside samples of analyst reports. More recently, Lee et al. (2015) uses internet co-searches at the SEC's EDGAR website to aggregate investor perception in identifying peer firms. They find that firms appearing

in chronologically adjacent searches by the same individual appear to be fundamentally similar in a variety of dimensions.

In addition to categorizing firms into groups, in various contexts prior studies have also tried to come up with some metrics to capture homogeneity between two firms or within groups using information from financial reports or stock returns. For example, Pyo and Lustgarten (1990) used the historical correlation in annual earnings and Parrino (1997) developed a proxy for industry homogeneity using return-based correlation within 2-digit SIC industries. Bhorraj and Lee (2002) use a valuation-based approach in the selection of peers by identifying peer firms as those having the closest “warranted multiple” with the underlying firm within the same 2-digit SIC industries. De Franco et al. (2011) develop a measure of financial statement comparability to capture how similar the firms are mapping economic events into financial statement through their accounting system within the same 2-digit SIC industries.

3. Overview of Different Methodology

In this section, I provide an overview and present the strengths and weaknesses of all the major methods that can be used in identifying peer firms, including using existing industry classification, textual similarity in business description, comovement in various return and accounting information, analyst co-coverage, and internet co-searches.

3.1 Industry Classification

As introduced in Section 2, a significant strength of industry classification is that it is widely adopted both in practice and academia. In addition, industry classification should also be quite reliable as it is usually defined by a consistent standard either by a

government agency or private organization. However, industry classification alone may not be able to tell you how similar the firms are to each other. In particular, the size of the groups can be enormous for certain industry classifications. For example, 4-digit SIC on average has more than 70 firms in each 4-digit classification and the largest number of firms within a 4-digit SIC classification can reach to over 300. Thus industry classification may not always be able to provide sufficient fine information to select a few close peers for a firm.

3.2 Textual Similarity in Business Description

As introduced in Chapters 1 and 2, textual information in the business description section in annual 10-K filings provides a rich set of data points for comparing similarities across firms. The much richer textual information set can help to provide a finer comparison across firms. Similar textual methods have also been tested by Hoberg and Philips (2010, 2015) in grouping firms based on product-market peers. However, it may be costly to implement the procedure for textual analysis, especially on parsing out the relevant section of the document. The data may also be subject to some measurement errors due to differential writing styles (e.g. choice of words) by different firms.

3.3 Comovement in Daily Returns

If stock returns can efficiently reflect changes in firm fundamentals on a daily basis, firms with highly correlated daily returns should be similar in their business exposure. The advantage of using stock return data is that it is readily available and stock return on a daily basis may be able to provide sufficiently large number of data points for comparison among firms. However, as the literature on price delay suggests (e.g. Hou and Moskowitz 2005), price of some firms may respond to information with some delays

due to the presence of market frictions. Thus correlation in daily return may contain some components that are unrelated to the common changes in firm fundamentals, potentially introducing some measurement errors.

3.4 Comovement in Monthly Returns

If monthly stock returns can efficiently reflect changes in firm fundamentals, firms with highly correlated monthly returns should likely be similar in their business exposure. Hou and Moskowitz (2005) claim that “at monthly frequencies, there is little dispersion in delay measures since most stocks respond to information within a month’s time”. If this is true, correlation in monthly returns can likely overcome the measurement problem associated with daily returns. However, returns on a monthly basis greatly reduce the data points for comparison as there are only 12 data points of variations available within a 1-year interval. Unless monthly returns are measured over a very long period, correlation in monthly returns may not provide sufficient variations in differentiating among firms.

3.5 Comovement in Earnings

If accounting earnings can efficiently reflect changes in firm fundamentals, firms with highly correlated accounting earnings should likely be similar in their business exposure. While data on earnings are readily available for public firms and some prior research have also attempted to use comovement in earnings as a measure of similarity (e.g. Pyo and Lustgarten 1990), earnings data is available at an even much lower frequency than monthly return. Thus comovement in earnings may provide even fewer variations than monthly returns in differentiating among a large number of firms. In addition, there are also concerns about potential earnings management (e.g. Burgstahler and Dichev 1997) that may create additional measurement errors.

3.6 Comovement in Operating Cash Flows

Similarly, if operating cash flows can efficiently reflect changes in firm fundamentals, firms with highly correlated accounting earnings should likely be similar in their business exposure. While operating cash flows are likely less subject to concerns related to earnings management, it has the same issue as earnings in that there are only 4 observations per firm in a single year, providing limited data points for variations in distinguishing among a large number of firms. On the other hand, operating cash flows may also be affected by differential corporate policy on cash flow management.

3.7 Analyst Co-Coverage

Analysts appear to generally focus on following related firms that belong to the same industry (e.g. Mikhail, et al. 2004, Boni and Womack 2006). Thus if there are substantially overlap in analyst coverage between two firms, these firms are likely peer firms to each other. Kaustia and Rantala (2013) use this analyst-based method in defining peer groups and find that their method produces substantially more homogenous groups of firms compared to common industry classification. However, this method is subject to a substantial limitation that it can only be reliably applied on firms that are covered by many analysts. This approach does not work at all for any firms with fewer than 2 analysts.

3.8 Internet Co-Searches

Lee et al. (2015) uses internet co-searches at the SEC's EDGAR website to aggregate investor perception in identifying peer firms. They find that firms appearing in chronologically adjacent searches by the same individual appear to be fundamentally similar in a variety of dimensions. They also show that this method appear to dominate 6-

digit GICS industry in their out-of-sample tests. However, this method is subject to the limitation that data on Internet traffic at SEC's EDGAR website is not readily available to the public. It may also be subject to some measurement errors, e.g. investors may be searching for firms involved in certain M&A deals or searching for firms which happen to announce their earnings/forecasts at the same time.

4. Comparison Tests

In this section, I implement three sets of tests to compare the performance of each method in identifying close peer firms. These include examining what percentage of the top peer firm identified by each method matches with a sample of highly probable peer pairs identified by using analyst co-coverage, how the top peer firm identified by each method performs in its future comovement in returns, earnings and operating cash flows with the base firm, as well as how much the top peer firm identified by each method explains the base firm's key contemporaneous valuation multiples and accounting ratios.

The various methods being tested include using textual similarity in business description in 10-K filings, correlation in daily returns over the past 1 year, correlation in monthly returns over the past 5 years, correlation in accounting earnings over the past 5 years, and correlation in operating cash flows over the past 5 years. For the tests that involve industry classification schemes, I generally include 8-digit GICS, 8-digit NAICS, 4-digit SIC, 6-digit GICS, textual product-based classification by Hoberg and Philips (2010, 2015) and Fama-French 48 industry classification (1997)⁴⁸.

⁴⁸ My testing sample includes all US public firms from 1996 to 2013 subject to data availability on the variables involved in the test. As GICS code is unavailable before 1999 from Compustat, I back-fill the GICS code of earlier years using the closest available GICS code in later years. I apply the same procedure

4.1 Benchmarking Against Analyst Co-Coverage Peers

Due to knowledge spillover among similar firms, sell-side analysts appear to specialize in an industry or some related industries, presumably to economize on the costs of information gathering and analysis. Institutionally, brokerage houses tend to be organized so that individual analysts focus on a specific industry (Michaely and Womack 1999). Thus firms that are jointly covered by the majority of their analysts are likely peer firms to one another. Kaustia and Rantala (2013) find that this analyst co-coverage method of finding peers produces substantially more homogenous groups of firms compared to common industry classification.

In this exercise, my objective is to come up with a “gold standard” benchmark sample and see how various methodologies perform relative to the “gold standard”. As my goal is to first find some pairs of firms that are highly likely to be peers to each other as the “gold standard”, I try to use the subsample of firms that are well covered by analysts under the analyst co-coverage method. First, I restrict the sample of firms to those that are covered by 5 or more analysts to minimize measurement errors associated with few analysts. Instead of the simulated peer threshold used by Kaustia and Rantala (2013), I require at least 50% of the analysts covering the firm with the least analyst coverage to be covering the other firm in order to count the pair of firms as peers. For example, if company A is covered by 7 analysts and company B is covered by 15 analysts, the two companies are peer firms to each other if they are jointly covered by 4 or more analysts.

for other industry classification system as well, e.g. I assume that the NAICS codes for firms in 1996 are the same as in 1997.

Using this procedure, for each well-covered firm I compile a list of peer firms that also well-covered by analysts.

The list, comprised of a subsample of well-covered firms with their well-covered peers, is subsequently used as a “gold standard” benchmark sample for comparisons among various methods being tested. While the list omits firms and their peers which are not well-covered by the analysts, it is unlikely that the peer pairs identified by this method are wrong. It may not be able to capture all possible peers with the relatively strict criteria imposed, especially for firms and peers that are not so well-covered by analysts. However, this method is likely able to capture the most relevant peers among the well-covered firms. According to this feature, my comparison test of the different methodologies examines how often the top peer firm identified through each method fall into the “gold standard” sample among the same subsample of well-covered firms with at least 5 analysts (for both the base firm as well as potential peer firms).

Next, I try to compare how the top peers identified by each method match with the “gold standard” sample using the same subsample of well covered firms. A better method should be able to identify top peer firms that are more likely to match with the “gold standard” sample. In addition, to help mitigate the problem of lacking variations to differentiate among large number of firms with correlation data of low frequency, I attempt to combine industry classification information with correlation-based data points. In this new set of comparison, I try to limit the search space for peer firms to within the firm’s Fama-French 48 industry classification using correlation data in daily returns, correlation in monthly returns, correlation in accounting earnings, and correlation in operating cash flows. In further comparison tests, I also compare my textual method

using business description with randomly selected peer firms under various industry classification schemes as well as both industry and size-matched peer firms.

4.2 Future Comovement of Key Market and Financial Variables

Similar in spirit to explaining out-of-sample variations in monthly returns in Table 3 of Lee et al. (2015), I examine the future comovement of key market and financial variables between the base firm and its top peer identified under different methodology. While Lee et al. (2015) present the results using the R^2 from univariate regressions of monthly returns, I use the correlation in monthly returns in a rolling window for the next 5 years, which can be viewed as a direct transformation from R^2 in univariate regression. In addition to monthly return as examined in Lee et al. (2015), I also examine the comovement of daily return in the coming year, comovement of earnings in the coming 5 years and the comovement of operating cash flows in the coming 5 years.

More specifically, I attempt to compare the future correlation of the key market and accounting variables between the base firms and their top peers identified under each methodology. Different methodologies being considered include using textual similarity in business description in 10-K filings, correlation in daily returns over the past 1 year, correlation in monthly returns over the past 5 years, correlation in accounting earnings over the past 5 years, and correlation in operating cash flows over the past 5 years, randomly selected peer firms as well as size-matched peer firms under 8-digit GICS and 6-digit GICS. A better method should generate top peer firms that have greater future comovement of key market and financial variables with the base firms.

4.3 Explaining Contemporaneous Valuation Multiples and Financial Ratios

My third comparison test largely follows Table 4 of Lee et al. (2015) in assessing the extent to which the top peers identified under various methodology explain the cross section of base firms' valuation multiples and key financial ratios. Key valuation multiples under consideration include price-to-book multiples (*pb*), enterprise value-to-sales multiples (*evs*), price-to-earnings multiples (*pe*). Key financial ratios under consideration include returns on net operating assets (*rnoa*), asset turnover (*at*), profit margins (*pm*), leverage (*lev*). For each variable, I run the following cross-sectional regression on a rolling quarterly basis where dependent variable $Variable_{i,t}$ is the variable of interest for each base firm *i* and the independent variable $Variable_{p(i),t}$ is the variable of interest for the top peer firm of the base firm *i* identified using different methods:

$$Variable_{i,t} = \alpha_t + \beta_t Variable_{p(i),t} + \varepsilon_{i,t}$$

Then, I try to compare the R^2 of the above regression under different methods in identifying top peer firms. R^2 represents how much the information from the top peer firm can explain the corresponding information of the base firm. A better methodology should generate top peers that have greater explanatory power on the valuation multiples and financial ratios of the base firm. In additional comparison tests, I also compare my textual method using business description with randomly selected peer firms as well as size-matched peer firms under various industry classification schemes, including 8-digit GICS, 8-digit NAICS, 4-digit SIC, 6-digit GICS, textual product-based classification by Hoberg and Philips (2010, 2015) and Fama-French 48 industry classification (1997).

5. Results of Comparison and Additional Analyses

5.1. Benchmarking Against Analyst Co-Coverage Peers

Table 1 displays the results on the percentage of the top peers identified by each method matching with the analyst co-coverage sample among the subsample of firms that are well-covered by analysts. In Panel A, in both the full and common samples of well-covered firms, I find that the method using textual similarity in the business description of 10-K filings performs the best, with its top peer to be found in the analyst co-coverage sample about 60% of the time. The method using historical 1-year correlation in daily returns appears to work the second best, matching the analyst co-coverage sample about 38% and 49% of the time in the full and common samples. The method using 5-year historical correlation in monthly return generally performs worse than the method using correlation in daily returns. On the other hand, the methods using 5-year historical correlations in earnings and operating cash flows perform very poorly, matching the analyst co-coverage sample less than 10% of the time. These results suggest that, for correlation-based method, sample size is very important in distinguishing among a large number of firms. While daily stock returns provides a lot of variations with more than 250 observation points within a single year, earnings and operating cash flows provide few variations with only 20 observation points over the course of 5 years.

In Panel B, I combine the correlation-based methods with industry classification to mitigate the problem from lacking variations to distinguish among a large number of firms. To do this, I limit the search space for the correlation-based methods to firms within the same Fama-French 48 industry classification. As shown in Panel B, the performance of all correlation-based methods improves after combining with the information from industry classification. Methods using correlations in historical earnings and operating cash flows have the most significant improvement, benefiting the most

from the much smaller search space with their top peers reaching 20 to 30% matching with the analyst co-coverage sample. However, the method using textual similarity in the business description of 10-K filings still performs the best, outperforming any correlation-based method combined with Fama-French 48 industry classification.

In Panel C, I compare the method using textual similarity in the business description with randomly selected peers from various commonly used industry classification schemes, including the finest 8-digit GICS, 4-digit SIC, 8-digit NAICS industry classification, 6-digit GICS as used by Bhojraj, et al. (2003), Fama-French 48 industry classification and textual product-based classification (Hoberg and Philips 2010, 2015). Not surprisingly, I find that the top peer selected using textual similarity in the business description outperforms randomly selected peers from any industry classification. This is likely because there are generally some variations in the similarity with the base firm within each industry classification as a result of firm heterogeneity within an industry. I also find that finer classification appears to perform better, as evidenced by the 8-digit GICS performing better than the 6-digit GICS classification as well as the 4-digit SIC performing better than the Fama-French 48 industry classification.

In Panel D, I further compare the method using textual similarity in the business description with size-matched peers from the above industry classification. Consistent with above, I find that the top peer selected using textual similarity in the business description continues to outperform size-matched peers from all the industry classifications. At the same time, there is also significant improvement from combining industry classification with firm size in selecting the top peer firm in matching with the

analyst co-coverage sample. It is possible that analysts may also consider similarity in scale in selecting their portfolio of coverage.

Overall, the results in Table 1 suggest that using textual similarity in business description performs the best among various methods in identifying a firm's closest peer that matches with the analyst co-coverage sample. The results also suggest that having a large number of observation points may be important in differentiating among a large number of firms. As data on earnings and operating cash flows are only available on a quarterly basis, methods using comovement in these financial variables alone perform poorly in differentiating among firms. However, combining them with industry classification may significantly help improve their performance by making the search space much smaller.

5.2 Future Comovement of Key Market and Financial Variables

Table 2 displays the future correlation of monthly returns, daily returns, accounting earnings and operating cash flows between the base firm and its top peer firm identified under various methods. In Panel A, I compare various methods by examining the comovement of monthly returns over the next 5 years between the base firm and the top peer firm selected under each method. I find that using correlation in daily returns over the past year performs the best in selecting the top peer firm with regards to future comovement in monthly returns, with correlation in future monthly returns reaching 0.299. My method using textual similarity in business description narrowly trails behind, with correlation in future monthly returns reaching 0.296.

In Panel B, I compare various methods by examining the comovement of daily returns over the next year between the base firm and the top peer firm selected under each

method. For future comovement in daily returns, I find that correlation in daily returns over the past year dominates the rest, with correlation in daily returns next year reaching 0.244. Size-matched 8-digit GICS performs the second best, with correlation in daily returns next year reaching 0.207. My method using textual similarity in business description also performs well, producing top peers with 0.196 correlation in daily returns next year.

In Panel C, I compare various methods by examining the comovement of accounting earnings over the next 5 years between the base firm and the top peer firm selected under each method. For future comovement in accounting earnings, I find that my method using textual similarity in business description dominates the rest, with correlation in future accounting earnings reaching 0.167. Correlation in past accounting earnings performs the second best, with correlation in future accounting earnings reaching 0.133. Surprisingly, correlation in past operating cash flows performs the worst, with only 0.077 correlation in future accounting earnings.

In Panel D, I compare various methods by examining the comovement of accounting earnings over the next 5 years between the base firm and the top peer firm selected under each method. For future comovement in operating cash flows, I find that the method using correlation in the past operating cash flows performs the best, with correlation in operating cash flows over the next 5 years reaching 0.182. My method using textual similarity in business description performs the second best, with correlation in operating cash flows over the next 5 years reaching 0.128. Interestingly, correlation in past accounting earnings performs the worst. In addition, methods using correlation in monthly returns and daily returns also perform poorly.

The results in Table 2 suggest that using textual similarity in business description performs consistently well across the different dimensions being tested. While using historical correlation in daily returns performs the best for future comovement in returns, using textual similarity in business description consistently outperforms historical correlation in daily returns for future comovement in accounting variables. Interestingly, using the historical comovement of the variable of interest tends to perform well in each future comovement test of that variable. For example, using historical comovement in operating cash flows generates top peers that have the highest future comovement in operating cash flows. This suggests that the comovement of each key market and accounting variable between the base firm and its closest peer tends to be sticky in its own unique way (i.e. not predicted well by other variables).

5.3 Explaining Contemporaneous Valuation Multiples and Financial Ratios

Table 3 displays the R^2 from regressing the base firm's key valuation multiples or accounting ratios on the corresponding contemporaneous variable from its top peer firm identified under various methods across a common sample in which data is available for all methods. In Panel A, I compare my method using textual similarity in business description against other correlation-based methods. I find that my method using textual similarity in business description performs the best in using the top peer firm's ratios in explaining the base firm's ratios across all three valuation multiples and three out of four financial ratios. For profit margin, using correlation of past daily returns performs the best among all the methods. Consistent with the results in my prior tests, correlation based on accounting earnings and operating cash flows perform the worst, likely due to the lack of variations in differentiating firms.

In Panel B, I compare the method using textual similarity in the business description with randomly selected peers from various commonly used industry classification schemes, including the finest 8-digit GICS, 4-digit SIC, 8-digit NAICS industry classification, 6-digit GICS as used by Bhojraj, et al. (2003), Fama-French 48 industry classification and textual product-based classification (Hoberg and Philips 2010, 2015). I find that the top peer under my textual approach performs the best in explaining enterprise value-to-sales ratio, asset turnover ratio and leverage. Although it is not the best in other ratios, my textual approach tends to perform consistently well. For other ratios, 8-digit GICS performs the best in explaining price-to-book ratio and profit margin while 8-digit NAICS performs the best in explaining price-to-earnings ratio and return on net operating asset.

In Panel C, I use size-matched peers from various industry classification schemes in my tests instead of randomly selected peers as in Panel B. Similarly, I find that the top peer under my textual approach performs the best in explaining the price-to-earnings ratio, enterprise value-to-sales ratio and profit margin, and also does well in other ratios. For other ratios, 8-digit NAICS works the best in explaining price-to-book ratio and leverage ratio while 4-digit SIC performs the best in explaining return on net operating asset and asset turnover ratio. Consistent with my prior tables, using finer industry classification generally outperforms, e.g. 8-digit GICS outperforms 6-digit GICS in almost all categories. Contrary to the findings in Bhojraj, et al. (2003), I find that 6-digit GICS firms do not seem to perform significantly better than 8-digit NAICS or 4-digit SIC for my sample of firms during my sample period.

Overall, the results of Table 3 are largely consistent with results in prior tables. Using textual similarity in business description performs consistently well across different valuation multiples and accounting ratios. For some ratios such as price-to-book, return on net operating asset, asset turnover and leverage, certain industry classification schemes may actually perform the best. I also find that a finer industry classification such as 8-digit GICS appears to dominate a coarser industry classification such as 6-digit GICS. In addition, my results also suggest that 8-digit GICS does not significantly outperform 8-digit NAICS or 4-digit SIC in my sample.

5.4 Cross-Sectional Variations in Similarity with the Closest Peer Firms

My prior analyses show that using textual similarity in business description and historical correlation in daily returns generally perform the best in identifying close peer firms. The next natural question is whether there are large cross-sectional variations in the similarity between the base firm and its top peer firm. If so, how are variations in similarity related to the future comovement of key market and financial variables?

In Table 4, I examine the extent of the variations in the similarity between the base firm and its closest peer firm. Since the degree of similarity is measured with different scales for different measures (e.g. my textual similarity measure of similarity is expressed in terms of cosine similarity score while others are expressed in terms of correlation score), I try to use the extent of future comovement in key market and accounting variables to standardize the scale for comparison purposes.

In Panel A of Table 4, I examine how the comovement in monthly returns between the base firm and its closest peer over the next 5 years varies across different quintiles of similarity between the base firm and its closest peer identified using various methods. I

find that the top peer firm identified through historical daily return comovement shows the most significant variations across the quintiles, ranging from 0.18 for the least similar quintile to 0.40 for the most similar quintile. Methods using textual business description and monthly returns are showing a similar scale of slightly less significant variations. However, methods using earnings and operating cash flows show minimal variations across different quintiles.

In Panel B, I examine how the comovement in daily returns between the base firm and its closest peer over the next 1 year varies across different quintiles of similarity between the base firm and its closest peer identified using various methods. The results are largely similar to Panel A. Top peer identified through historical daily return comovement shows the most significant variations across the quintiles, ranging from 0.07 for the least similar quintile to 0.35 for the most similar quintile, a difference of more than 7 times. Methods using textual business description and monthly returns are showing a similar scale of less significant variations while methods using earnings and operating cash flows only show minimal variations across different quintiles.

In Panel C, I examine how the comovement in accounting earnings between the base firm and its closest peer over the next 5 years varies across different quintiles of similarity between the base firm and its closest peer identified using various methods. I find that top peer identified by my textual similarity approach shows the most significant variations across the quintiles, ranging from 0.11 for the least similar quintile to 0.26 for the most similar quintile. The method using historical correlation in daily returns shows slightly less significant variations across quintiles. Other methods using monthly returns,

earnings and operating cash flows also show some variations across the different quintiles, with a difference of 0.05-0.06 between the bottom and top quintile.

In Panel D, I examine how the comovement in operating cash flows between the base firm and its closest peer over the next 5 years varies across different quintiles of similarity between the base firm and its closest peer identified using various methods. In this test, the top peer firm identified through historical daily return comovement shows the most significant variations across the quintiles, ranging from 0.08 for the least similar quintile to 0.19 for the most similar quintile. Methods using textual business description does slightly worse, ranging from 0.09 for the least similar quintile to 0.17 for the most similar quintile. Methods using historical correlation in earnings also show some variations across the quintiles, with a difference of 0.04 between the bottom and top quintile. However, methods using monthly returns only show minimal variations across different quintiles.

In Table 5, I attempt to more directly assess how the similarity between the base firm and its closest peer relates to the future comovement of key market and accounting variables. Instead of using the quintiles as in Table 4, I directly run regression of the future comovement of key market and accounting variables on the raw similarity score between the base firm and its closest peer identified under various methods.

Similar to the results in Table 4, the regression results in Table 5 suggest that the degree of similarity between the base firm and its closest peer is significantly positively associated with the future comovement in their key market and accounting variables. Across Panels A to D, almost all univariate regressions from columns (1) to (5) are statistically significant, except in column (5) of Panel A that similarity measured by

historical comovement in operating cash flow is not significantly associated with future comovement in monthly returns as well as in column (3) of Panel D that similarity measured by historical comovement in monthly return is also not significantly associated with future comovement in operating cash flows. Across all the panels, textual similarity in business description and historical correlation in daily returns show the most significant association with future comovement in key market and accounting variables.

In column (6) across the panels in Table 5, I regress future comovement of market or accounting variables on all the similarity scores computed under different methods as multiple independent variables in one regression. I find that textual similarity in business description and historical correlation in daily returns are the only two methods that are consistently statistically significant in all four panels. I also find that the historical comovement of any particular variable is also significantly associated with the future comovement of that particular variable, even after controlling for other similarity scores computed by other methods.

6. Summary

In the paper, I compare various methods in searching for close comparable peer firms, through examining how the top peer firm identified under each method matches with a sample of highly probable peer pairs identified by using analyst co-coverage, how each method performs in their top peer firms' future comovement in returns, earnings and operating cash flows with the base firm, as well as how much the top peer firm identified under each method can explain their base firm's key valuation multiples and accounting ratios contemporaneously. Besides my method using the textual similarity in business

description from the earlier two chapters, I also include various easy-to-implement correlation-based methods using daily returns, monthly returns, accounting earnings and operating cash flows, along with the commonly used industry classification in my comparison tests.

The results show that my approach using the textual similarity in business description consistently performs well in various comparison tests. The method using historical comovement in daily returns also performs well in most comparison tests. Among others, these two metrics appear to provide the best measures of similarity in business exposure between firms. On the contrary, methods using historical comovement in either earnings or operating cash flows alone perform poorly in various comparison tests, likely due to the lack of variations in distinguishing among a large number of firms. However, combining historical comovement in earnings or operating cash flows with industry classification can significantly help improve their performance by limiting the search space to much fewer firms.

In further analyses, I find that there are indeed large variations in the similarity between the base firms and their closest peers and the degree of similarity is significantly positively associated with the future comovement in returns, earnings and operating cash flows. These translate to large variations in the future comovement in returns, earnings and operating cash flows across different firm-peer pairs. This means that, for some firms that show low comovement in returns, earnings and operating cash flows even with their closest peers, it will be more difficult for investors to use the information from other peer firms to learn about the base firm. This is related to the issue of “informational uniqueness” described in my Chapter 1.

While I find that textual similarity in business description and historical correlation in daily returns appear to be the best two measures of similarity between firms, future research may investigate if some of these methods can be combined in some ways to minimize the measurement errors associated with each method. In deciding which method to choose as a measurement of similarity between firms, researchers may need to think about how the construction of the measure is related to other variables of interests. In my previous two chapters, I choose the textual similarity in business description in 10-K filings as my main measure because correlation in daily returns may be mechanically related to my other variables of interests, e.g. stock returns may be affected by a firm's guidance policy and discretionary disclosure.

Table 1 Comparison of % Top Peer in Analyst Co-Coverage Sample

The tables below display what percentage of the top peers identified by each methodology coincides with the analyst co-coverage sample among the subsample of firms that are well-covered by analysts. Panel A shows the results for top peers identified by textual similarity in business description, correlation in daily returns, correlation in monthly returns, correlation in accounting earnings, and correlation in operating cash flows without imposing any restriction on industry membership. Panel B shows the results for the same methodology in A, but restricting the search space for peers to well-covered firms within the same Fama-French 48 industry classification for correlation in daily returns, correlation in monthly returns, correlation in accounting earnings, and correlation in operating cash flows (not for textual similarity in business description). Panel C compare the results for top peers identified by textual similarity in business description to randomly selected peers from various industry classification, including 8-digit GICS, 8-digit NAICS, 4-digit SIC, 6-digit GICS, textual product-based classification by Hoberg and Philips (2010, 2015) and Fama-French 48 industry classification. Panel D compares the results for top peers identified by textual similarity in business description to the size-matched peers from various industry classification, including 8-digit GICS, 8-digit NAICS, 4-digit SIC, 6-digit GICS, textual product-based classification by Hoberg and Philips (2010, 2015) and Fama-French 48 industry classification.

Panel A: Textual Method and Correlation-Based Method

	Textual Similarity in Business Description	Correlation in Daily Returns Over Past 1 Year	Correlation in Monthly Returns Over Past 5 Years	Correlation in Acct. Earnings Over 5 Years	Correlation in Operating Cash Flows Over 5 Years
Full Sample					
% Top Peer in Analyst Co-Coverage Sample	0.579	0.382	0.281	0.074	0.048
No. of Observation	15419	17294	14433	12555	12162
Common Sample					
% Top Peer in Analyst Co-Coverage Sample	0.617	0.487	0.347	0.092	0.057
No. of Observation	4277	4277	4277	4277	4277

Panel B: Textual Method and Correlation-Based Method Combined with Industry

	Textual Similarity in Business Description	Correlation in Daily Returns within FF48	Correlation in Monthly Returns within FF48	Correlation in Acct. Earnings within FF48	Correlation in Op. Cash Flows within FF48
Full Sample					
% Top Peer in Analyst Co-Coverage Sample	0.579	0.471	0.425	0.254	0.230
No. of Observation	15419	17569	15969	12790	12713
Common Sample					
% Top Peer in Analyst Co-Coverage Sample	0.607	0.563	0.498	0.286	0.271
No. of Observation	5278	5278	5278	5278	5278

Panel C: Textual Method and Industry Classification

	Textual Similarity	Randomly Selected From					
		GIC8	NAICS8	SIC4	GIC6	TNIC	FF48
Full Sample							
% Top Peer in Analyst Co-Coverage Sample	0.579	0.286	0.363	0.327	0.184	0.239	0.132
No. of Observation	15419	13286	13054	13473	13088	13373	13328
Common Sample							
% Top Peer in Analyst Co-Coverage Sample	0.620	0.325	0.397	0.374	0.215	0.257	0.142
No. of Observation	2216	2216	2216	2216	2216	2216	2216

Panel D: Textual Method and Industry Classification Combined with Firm Size

	Textual Similarity	Size- matched from GIC8	Size- matched from NAICS8	Size- matched from SIC4	Size- matched from GIC6	Size- matched from TNIC	Size- matched from FF48
Full Sample							
% Top Peer in Analyst Co- Coverage Sample	0.579	0.435	0.454	0.434	0.312	0.313	0.219
No. of Observation	15419	15645	15085	15797	15424	15758	15420
Common Sample							
% Top Peer in Analyst Co- Coverage Sample	0.609	0.494	0.494	0.480	0.365	0.362	0.264
No. of Observation	6465	6465	6465	6465	6465	6465	6465

Table 2 Future Comovement in Key Market and Financial Variables

The tables below display the future correlation of key market and financial variables between the base firm and its top peer firm identified under various methodologies. Panel A shows the results for the future correlation in monthly returns between the base firm and its closest peer firm. Panel B shows the results for the future correlation in daily returns between the base firm and its closest peer firm. Panel C shows the results for the future correlation in earnings between the base firm and its closest peer firm. Panel D shows the results for the future correlation in operating cash flows between the base firm and its closest peer firm.

Panel A: Comovement in Monthly Returns in the next 5 Years

	Mean	S.D.	N
Textual Similarity in Business Description	0.296	0.221	26,445
Correlation in Daily Returns	0.299	0.235	31,175
Correlation in Monthly Returns	0.271	0.271	26,661
Correlation in Earnings	0.211	0.192	20,369
Correlation in Operating cash flows	0.199	0.187	19,877
Random from GIC6	0.243	0.220	24,674
Size-matched from GIC6	0.271	0.213	25,178
Random from GIC8	0.238	0.203	24,779
Size-matched from GIC8	0.285	0.220	25,549

Panel B: Comovement in Daily Returns in the next 1 Year

	Mean	S.D.	N
Textual Similarity in Business Description	0.196	0.212	57,514
Correlation in Daily Returns	0.244	0.240	65,059
Correlation in Monthly Returns	0.196	0.224	59,799
Correlation in Earnings	0.161	0.182	41,670
Correlation in Operating cash flows	0.154	0.172	41,602
Random from GIC6	0.166	0.208	53,235
Size-matched from GIC6	0.197	0.211	53,271
Random from GIC8	0.161	0.190	53,228
Size-matched from GIC8	0.207	0.219	53,583

Panel C: Comovement in Earnings in the Next 5 Years

	Mean	S.D.	N
Textual Similarity in Business Description	0.167	0.338	23,695
Correlation in Daily Returns	0.119	0.335	25,400
Correlation in Monthly Returns	0.111	0.333	21,466
Correlation in Earnings	0.133	0.335	19,632
Correlation in Operating cash flows	0.077	0.316	18,968
Random from GIC6	0.110	0.337	22,180
Size-matched from GIC6	0.105	0.324	22,758
Random from GIC8	0.119	0.329	22,957
Size-matched from GIC8	0.098	0.320	22,106

Panel D: Comovement in Operating Cash Flows in the Next 5 Years

	Mean	S.D.	N
Textual Similarity in Business Description	0.128	0.319	23,359
Correlation in Daily Returns	0.085	0.311	24,582
Correlation in Monthly Returns	0.076	0.308	20,631
Correlation in Earnings	0.074	0.308	18,797
Correlation in Operating cash flows	0.182	0.336	18,823
Random from GIC6	0.084	0.319	21,889
Size-matched from GIC6	0.078	0.306	22,475
Random from GIC8	0.074	0.300	21,821
Size-matched from GIC8	0.090	0.314	22,707

Table 3 Explaining Valuation Multiple and Accounting Ratios

The tables below display the R^2 from regressing the base firm's variable about key valuation multiples or accounting ratios on the corresponding contemporaneous variable from its top peer firm identified under various methods. These variables include price-to-book multiples (PB), enterprise value-to-sales multiples (EVS), price-to-earnings multiples (PE), returns on net operating assets (RNOA), asset turnover (AT), profit margins (PM), leverage (LEV). Panel A shows the results for top peers identified by textual similarity in business description, correlation in daily returns, correlation in monthly returns, correlation in accounting earnings, and correlation in operating cash flows. Panel B compares the textual method to randomly selected peer from various industry classification, including 6-digit GICS, 8-digit GICS, 8-digit NAICS, 4-digit SIC, Fama-French 48 industry classification (FF48), and textual product-based classification (TNIC) by Hoberg and Philips (2010, 2015). Panel C compares the textual method to size-matched peer from various industry classification, including 6-digit GICS, 8-digit GICS, 8-digit NAICS, 4-digit SIC, Fama-French 48 industry classification (FF48), and textual product-based classification (TNIC) by Hoberg and Philips (2010, 2015).

Panel A: Textual Method and Correlation-Based Methods

	N	Textual Description	Daily Returns	Monthly Returns	Accounting Earnings	Operating Cash Flow
PB	15,957	0.060	0.026	0.025	0.008	0.005
PE	8,818	0.025	0.015	0.025	0.019	0.006
EVS	15,957	0.281	0.236	0.230	0.031	0.043
RNOA	15,976	0.052	0.019	0.016	0.007	0.000
AT	16,224	0.478	0.388	0.249	0.147	0.105
PM	16,224	0.162	0.184	0.141	0.082	0.048
LEV	16,224	0.041	0.028	0.025	0.020	0.016

Panel B: Randomly Selected Peer from Industry Classification⁴⁹

	N	Textual	GIC6	GIC8	NAICS	SIC	FF48	TNIC
PB	3,688	0.024	0.011	0.044	0.003	0.006	0.011	0.008
PE	1,365	0.011	0.000	0.002	0.015	0.001	0.002	0.011
EVS	3,684	0.342	0.040	0.216	0.202	0.273	0.055	0.125
RNOA	3,684	0.082	0.015	0.068	0.128	0.115	0.021	0.028
AT	3,700	0.371	0.057	0.204	0.245	0.252	0.031	0.138
PM	3,700	0.163	0.018	0.171	0.167	0.146	0.066	0.024
LEV	3,700	0.074	0.027	0.033	0.062	0.041	0.002	0.028

Panel C: Size-Matched Peer from Industry Classification⁴⁹

	N	Textual	GIC6	GIC8	NAICS	SIC	FF48	TNIC
PB	48,660	0.054	0.019	0.029	0.067	0.059	0.015	0.029
PE	25,181	0.017	0.004	0.007	0.007	0.007	0.005	0.004
EVS	48,660	0.340	0.122	0.232	0.241	0.272	0.111	0.215
RNOA	48,576	0.068	0.044	0.068	0.094	0.099	0.028	0.053
AT	49,056	0.367	0.115	0.347	0.365	0.401	0.141	0.213
PM	49,056	0.212	0.062	0.159	0.175	0.171	0.048	0.113
LEV	49,056	0.036	0.017	0.054	0.027	0.028	0.009	0.018

⁴⁹ In Panel B, I do not require the availability of data to calculate the multiples and ratios before randomly assigning a peer firm. After requiring the availability of data for all 7 industry classification, sample size drops significantly from 208,645 firm-quarter observations for textual similarity method, 162,217 firm-quarter observations for 6-digit GICS, 165,424 firm-quarter observations for 8-digit GICS, 158,037 firm-quarter observations for 8-digit NAICS, 169,875 firm-quarter observations for 4-digit SIC, 162,754 firm-quarter observations for Fama-French 48 industry classification and 180,507 observations for textual product-market based classification by Hoberg and Philips (2010, 2015). This is out of a total of 520,219 firm-quarter observations. In Panel C, since there is an overlap in peers selected under various methods due to size-matching, there is a less significant drop in sample size after requiring data availability for all 7 industry classification methods.

Table 4 Variations in Similarity with the Closest Peer

The tables below display the variations in the similarity with the closest peer under different methods, including using textual similarity in business description, correlation in daily returns, correlation in monthly returns, correlation in accounting earnings, and correlation in operating cash flows. Panel A shows how different quintiles of similarity relate to the comovement in monthly returns in the next 5 years. Panel B shows how different quintiles of similarity relate to the comovement in daily returns in the next 1 year. Panel C shows how different quintiles of similarity relate to the comovement in accounting earnings in the next 5 years. Panel D shows how different quintiles of similarity relate to the comovement in operating cash flows in the next 5 years.

Panel A: Comovement in Monthly Returns in the Next 5 Years

Quintiles of Similarity with Closest Peer	Textual Business Description	Daily Returns	Monthly Returns	Earnings	Operating cash flows
1 (most similar)	0.352	0.404	0.345	0.305	0.262
2	0.307	0.308	0.314	0.265	0.259
3	0.273	0.268	0.300	0.263	0.278
4	0.252	0.216	0.265	0.245	0.279
5 (least similar)	0.239	0.182	0.224	0.264	0.261

Panel B: Comovement in Daily Returns in the Next 1 Year

Quintiles of Similarity with Closest Peer	Textual Business Description	Daily Returns	Monthly Returns	Earnings	Operating cash flows
1 (most similar)	0.235	0.346	0.239	0.201	0.216
2	0.208	0.233	0.220	0.226	0.221
3	0.191	0.177	0.199	0.208	0.207
4	0.160	0.109	0.161	0.198	0.178
5 (least similar)	0.145	0.065	0.142	0.179	0.186

Panel C: Comovement in Earnings in the Next 5 Years

Quintiles of Similarity with Closest Peer	Textual Business Description	Daily Returns	Monthly Returns	Earnings	Operating cash flows
1 (most similar)	0.256	0.246	0.196	0.192	0.210
2	0.178	0.169	0.171	0.199	0.169
3	0.149	0.143	0.166	0.179	0.162
4	0.134	0.115	0.160	0.166	0.161
5 (least similar)	0.109	0.115	0.135	0.136	0.160

Panel D: Comovement in Operating Cash Flows in the Next 5 Years

Quintiles of Similarity with Closest Peer	Textual Business Description	Daily Returns	Monthly Returns	Earnings	Operating cash flows
1 (most similar)	0.167	0.190	0.133	0.152	0.217
2	0.131	0.127	0.122	0.144	0.125
3	0.127	0.108	0.132	0.131	0.110
4	0.114	0.096	0.128	0.126	0.104
5 (least similar)	0.092	0.085	0.121	0.114	0.102

Table 5 Similarity with the Closest Peers and Comovement in Future Periods

The tables below display how the similarity with the closest peer under different methods relates to the future comovement of key market and accounting variables using regression analysis. Panel A regresses the comovement in monthly returns in the next 5 years on the similarity measure under various methods. Panel B regresses the comovement in daily returns in the next 1 year on the similarity measure under various methods. Panel C regresses the comovement in accounting earnings in the next 5 years on the similarity measure under various methods. Panel D regresses the comovement in operating cash flows in the next 5 years on the similarity measure under various methods. Firm clustered t-statistics in parentheses *** p<0.01, ** p<0.05, * p<0.1

Panel A: Comovement in Monthly Returns in the Next 5 Years

VARIABLES	(1) Future Monthly Return Comovement	(2) Future Monthly Return Comovement	(3) Future Monthly Return Comovement	(4) Future Monthly Return Comovement	(5) Future Monthly Return Comovement	(6) Future Monthly Return Comovement
Textual Similarity	0.457*** (8.11)					0.268*** (3.13)
Daily Return Comovement		0.503*** (5.59)				0.476*** (5.61)
Monthly Return Comovement			0.345*** (7.68)			0.169** (2.16)
Earnings Comovement				0.211*** (3.35)		0.099 (1.19)
Op. Cash Flow Comovement					0.002 (0.03)	-0.228** (-2.32)
Year FE	YES	YES	YES	YES	YES	YES
Observations	32,582	32,109	30,153	26,521	26,640	23,888
R-squared	0.086	0.163	0.087	0.074	0.070	0.203

Panel B: Comovement in Daily Returns in the Next 1 Year

VARIABLES	(1) Future Daily Return Comovement	(2) Future Daily Return Comovement	(3) Future Daily Return Comovement	(4) Future Daily Return Comovement	(5) Future Daily Return Comovement	(6) Future Daily Return Comovement
Textual	0.395***					0.242***
Similarity	(12.41)					(4.90)
Daily Return		0.775***				0.764***
Comovement		(17.49)				(13.94)
Monthly Return			0.299***			0.042
Comovement			(8.26)			(1.18)
Earnings				0.199***		0.052
Comovement				(4.89)		(1.27)
Op. Cash Flow					0.275***	0.044
Comovement					(4.24)	(0.87)
Year FE	YES	YES	YES	YES	YES	YES
Observations	63,584	62,723	58,616	47,798	48,252	42,709
R-squared	0.211	0.459	0.220	0.192	0.195	0.480

Panel C: Comovement in Earnings in the Next 5 Years

VARIABLES	(1) Future Earnings Comovement	(2) Future Earnings Comovement	(3) Future Earnings Comovement	(4) Future Earnings Comovement	(5) Future Earnings Comovement	(6) Future Earnings Comovement
Textual Similarity	0.634*** (12.11)					0.616*** (9.97)
Daily Return Comovement		0.426*** (16.27)				0.364*** (12.23)
Monthly Return Comovement			0.171*** (5.07)			0.051 (1.30)
Earnings Comovement				0.277*** (6.13)		0.151*** (3.30)
Op. Cash Flow Comovement					0.265*** (4.63)	0.210*** (3.63)
Year FE	YES	YES	YES	YES	YES	YES
Observations	22,199	20,327	20,366	16,970	17,008	15,244
R-squared	0.028	0.032	0.006	0.006	0.005	0.059

Panel D: Comovement in Operating Cash Flows in the Next 5 Years

VARIABLES	(1) Future Operating Cash Flow Comovement	(2) Future Operating Cash Flow Comovement	(3) Future Operating Cash Flow Comovement	(4) Future Operating Cash Flow Comovement	(5) Future Operating Cash Flow Comovement	(6) Future Operating Cash Flow Comovement
Textual	0.347***					0.368***
Similarity	(6.56)					(6.18)
Daily Return		0.306***				0.285***
Comovement		(12.96)				(10.37)
Monthly Return			0.014			-0.078*
Comovement			(0.43)			(-1.96)
Earnings				0.178***		0.066
Comovement				(3.93)		(1.45)
Op. Cash Flow					0.613***	0.581***
Comovement					(9.44)	(8.55)
Year FE	YES	YES	YES	YES	YES	YES
Observations	21,864	20,140	20,182	16,758	16,857	15,129
R-squared	0.009	0.018	0.001	0.002	0.019	0.045

Bibliography

- Acharya, V. V., DeMarzo, P., & Kremer, I. (2011). Endogenous information flows and the clustering of announcements. *American Economic Review*, 101(7), 2955-2979.
- Ali, A., Klasa, S., & Yeung, E. (2009). The limitations of industry concentration measures constructed with compustat data: Implications for finance research. *Review of Financial Studies*, 22(10), 3839-3871.
- Amihud, Yakov, (2002), Illiquidity and stock returns: cross-section and time-series effects, *Journal of Financial Markets*, 5, issue 1, p. 31-56.
- Badertscher, B., Shroff, N., & White, H. D. (2013). Externalities of public firm presence: Evidence from private firms' investment decisions. *Journal of Financial Economics*, 109(3), 682-706.
- Baginski, S. P. (1987). Intraindustry information transfers associated with management forecasts of earnings. *Journal of Accounting Research*, 25(2), 196-216.
- Baginski, S. P., Hassell, J. M., & Kimbrough, M. D. (2002). The effect of legal environment on voluntary disclosure: Evidence from management earnings forecasts issued in U.S. and canadian markets. *The Accounting Review*, 77(1), 25-50.
- Baiman, S., & Verrecchia, R. E. (1996). The relation among capital markets, financial disclosure, production efficiency, and insider trading. *Journal of Accounting Research*, 34(1), 1-22.
- BALAKRISHNAN, K., BILLINGS, M. B., KELLY, B., & LJUNGQVIST, A. (2014). Shaping liquidity: On the causal effects of voluntary disclosure. *The Journal of Finance*, 69(5), 2237-2278.
- Ball, R., and L. Shivakumar. (2008). How much new information is there in earnings? *Journal of Accounting Research* 46: 975-1016.
- Ball, R., Jayaraman, S., & Shivakumar, L. (2012). Audited financial reporting and voluntary disclosure as complements: A test of the confirmation hypothesis. *Journal of Accounting and Economics*, 53(1-2), 136-166.

- Barney, J. B. (1986). Strategic factor markets: Expectations, luck, and business strategy. *Management Science*, 32(10), 1231-1241.
- Beaver, W., 1968, The information content of annual earnings announcements, *Journal of Accounting Research*, Supplement, 67-92.
- Beyer, A., Cohen, D. A., Lys, T. Z., & Walther, B. R. (2010). The financial reporting environment: Review of the recent literature. *Journal of Accounting and Economics*, 50(2-3), 296-343.
- Beyer, A., & Dye, R. (2012). Reputation management and the disclosure of earnings forecasts. *Review of Accounting Studies*, 17(4), 877-912.
- Bhojraj, S., & Lee, C. M. C. (2002). Who is my peer? A valuation-based approach to the selection of comparable firms. *Journal of Accounting Research*, 40(2), 407-439.
- Bhojraj, S., Lee, C. M. C., & Oler, D. K. (2003). What's my line? A comparison of industry classification schemes for capital market research. *Journal of Accounting Research*, 41(5), 745-774.
- Billings, M. B., Jennings, R., & Lev, B. (2015). On guidance and volatility. *Journal of Accounting and Economics*, 60(2-3), 161-180.
- Boni, L., & Womack, K. L. (2006). Analysts, industries, and price momentum. *The Journal of Financial and Quantitative Analysis*, 41(1), 85-109.
- Brown, N., & Kimbrough, M. (2011). Intangible investment and the importance of firm-specific factors in the determination of earnings. *Review of Accounting Studies*, 16(3), 539-573.
- Brown, Nerissa C. and Gordon, Lawrence A. and Wermers, Russ. (2006). Herd Behavior in Voluntary Disclosure Decisions: An Examination of Capital Expenditure Forecasts. Working Paper.
- BROWN, S. V., & TUCKER, J. W. (2011). Large-sample evidence on firms' year-over-year MD&A modifications. *Journal of Accounting Research*, 49(2), 309-346.
- Burgstahler, D., & Dichev, I. (1997). Earnings management to avoid earnings decreases and losses. *Journal of Accounting and Economics*, 24(1), 99-126.

- Chen, S., Matsumoto, D., & Rajgopal, S. (2011). Is silence golden? an empirical analysis of firms that stop giving quarterly earnings guidance. *Journal of Accounting and Economics*, 51(1–2), 134-150.
- Chuk, E., Matsumoto, D., & Miller, G. S. (2013). Assessing methods of identifying management forecasts: CIG vs. researcher collected. *Journal of Accounting and Economics*, 55(1), 23-42.
- Coller, M., & Yohn, T. L. (1997). Management forecasts and information asymmetry: An examination of bid-ask spreads. *Journal of Accounting Research*, 35(2), 181-191.
- Cooper, Michael J. and He, Jing and Plumlee, Marlene (2015), Voluntary Disclosure and Investor Sentiment. Working Paper.
- DE FRANCO, G., KOTHARI, S. P., & VERDI, R. S. (2011). The benefits of financial statement comparability. *Journal of Accounting Research*, 49(4), 895-931.
- Dechow, P. M., & Dichev, I. D. (2002). The quality of accruals and earnings: The role of accrual estimation errors. *Accounting Review*, 77(4), 35.
- Demirakos, E. G., Strong, N. C., & Walker, M. (2004). What valuation models do analysts use? *Accounting Horizons*, 18(4), 221-240.
- DIAMOND, D. W., & VERRECCHIA, R. E. (1991). Disclosure, liquidity, and the cost of capital. *The Journal of Finance*, 46(4), 1325-1359.
- Dye, R. A. (1985). Disclosure of nonproprietary information. *Journal of Accounting Research*, 23(1), 123-145.
- Dye, R. A., & Sridhar, S. S. (1995). Industry-wide disclosure dynamics. *Journal of Accounting Research*, 33(1), 157-174.
- Eberhart, A. C. (2001). Comparable firms and the precision of equity valuations. *Journal of Banking & Finance*, 25(7), 1367-1400.
- EINHORN, E., & ZIV, A. (2008). Intertemporal dynamics of corporate voluntary disclosures. *Journal of Accounting Research*, 46(3), 567-589.
- Fama, E. F., & French, K. R. (1997). Industry costs of equity. *Journal of Financial Economics*, 43(2), 153-193.

- Fischer, P. E., & Verrecchia, R. E. (2000). Reporting bias. *The Accounting Review*, 75(2), 229-245.
- Foster, G. (1981). Intra-industry information transfers associated with earnings releases. *Journal of Accounting & Economics*, 3(3), 201-232.
- Francis, J., Philbrick, D., & Schipper, K. (1994). Shareholder litigation and corporate disclosures. *Journal of Accounting Research*, 32(2), 137-164.
- Franco, G., Hope, O., & Larocque, S. (2015). Analysts' choice of peer companies. *Review of Accounting Studies*, 20(1), 82-109. doi:10.1007/s11142-014-9294-7
- Freeman, R., & Tse, S. (1992). An earnings prediction approach to examining intercompany information transfers. *Journal of Accounting and Economics*, 15(4), 509-523.
- Glosten, L. R., & Milgrom, P. R. (1985). Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics*, 14(1), 71-100.
- Gong, G., Li, L. Y., & Zhou, L. (2013). Earnings non-synchronicity and voluntary disclosure. *Contemporary Accounting Research*, 30(4), 1560-1589.
- HAIL, L. (2011). Discussion of consequences and institutional determinants of unregulated corporate financial statements: Evidence from embedded value reporting. *Journal of Accounting Research*, 49(2), 573-594.
- Han, J. C. Y., Wild, J. J., & Ramesh, K. (1989). Managers' earnings forecasts and intra-industry information transfers. *Journal of Accounting and Economics*, 11(1), 3-33.
- Heinle, Mirko, Verrecchia, Robert E (2015), Bias and Commitment to Disclosure. *Management Science*, Forthcoming.
- Hoberg, G., & Phillips, G. (2010). Product market synergies and competition in mergers and acquisitions: A text-based analysis. *Review of Financial Studies*, 23(10), 3773-3811.
- Hoberg, Gerard and Phillips, Gordon M. (2014), Product Market Uniqueness, Organizational Form and Stock Market Valuations. Working Paper.

- Hoberg, Gerard and Phillips, Gordon M. (2015), Text-Based Network Industries and Endogenous Product Differentiation. *Journal of Political Economy*, Forthcoming
- Hou, K., & Moskowitz, T. J. (2005). Market frictions, price delay, and the cross-section of expected returns. *Review of Financial Studies*, 18(3), 981-1020.
- HOUSTON, J. F., LEV, B., & TUCKER, J. W. (2010). To guide or not to guide? causes and consequences of stopping quarterly earnings guidance. *Contemporary Accounting Research*, 27(1), 143-185.
- Jerry, C. Y. H., & Wild, J. J. (1990). Unexpected earnings and intraindustry information transfers: Further evidence. *Journal of Accounting Research*, 28(1), 211-219.
- JORGENSEN, B. N., & KIRSCHENHEITER, M. T. (2012). Interactive discretionary disclosures*. *Contemporary Accounting Research*, 29(2), 382-397.
- Kasznik, R.. (1999). On the Association between Voluntary Disclosure and Earnings Management. *Journal of Accounting Research*, 37(1), 57–81.
- Kaustia, Markku and Rantala, Ville (2013), Common Analyst - Based Method for Defining Peer Firms Working paper (December 17, 2013).
- KRISHNAN, J., & PRESS, E. (2003). The north american industry classification system and its implications for accounting research. *Contemporary Accounting Research*, 20(4), 685-717.
- Kim, I., & Skinner, D. J. (2012). Measuring securities litigation risk. *Journal of Accounting and Economics*, 53(1–2), 290-310.
- LAMBERT, R. A., LEUZ, C., & VERRECCHIA, R. E. (2012). Information asymmetry, information precision, and the cost of capital. *Review of Finance*, 16(1), 1-29.
- Landsman, W. R., Maydew, E. L., & Thornock, J. R. (2012). The information content of annual earnings announcements and mandatory adoption of IFRS. *Journal of Accounting and Economics*, 53(1–2), 34-54.
- Lee, C. M. C., Ma, P., & Wang, C. C. Y. (2015). Search-based peer firms: Aggregating investor perceptions through internet co-searches. *Journal of Financial Economics*, 116(2), 410-431.

- Lee, S. , Matsunaga, S. R., & Park, C. W. (2012). Management forecast accuracy and CEO turnover. *The Accounting Review*, 87(6), 2095-2122.
- Lerman, A., & Livnat, J. (2010). The new form 8-K disclosures. *Review of Accounting Studies*, 15(4), 752-778.
- Leuz, C., & Verrecchia, R. E. (2000). The economic consequences of increased disclosure. *Journal of Accounting Research*, 38, 91-124.
- Li, Xi and Yang, Holly, Mandatory Financial Reporting and Voluntary Disclosure: The Effect of Mandatory IFRS Adoption on Management Forecasts (August 2015). *The Accounting Review*, Forthcoming.
- Litov, L. P., Moreton, P., & Zenger, T. R. (2012). Corporate strategy, analyst coverage, and the uniqueness paradox. *Management Science*, 58(10), 1797-1815.
- Matsumoto, D., Pronk, M., & Roelofsen, E. (2011). What makes conference calls useful? the information content of managers' presentations and analysts' discussion sessions. *The Accounting Review*, 86(4), 1383-1414.
- Michaely, R., & Womack, K. L. (1999). Conflict of interest and the credibility of underwriter analyst recommendations. *Review of Financial Studies*, 12(4), 653-686.
- Mikhail, M. B., Walther, B. R., & Willis, R. H. (2004). Do security analysts exhibit persistent differences in stock picking ability? *Journal of Financial Economics*, 74(1), 67-91.
- Morck, R., Yeung, B., & Yu, W. (2000). The information content of stock markets: Why do emerging markets have synchronous stock price movement? *Journal of Financial Economics*, 58(1), 215-260.
- Parrino, R. (1997). CEO turnover and outside succession A cross-sectional analysis. *Journal of Financial Economics*, 46(2), 165-197.
- Pownall, G., & Waymire, G. (1989). Voluntary disclosure choice and earnings information transfer. *Journal of Accounting Research*, 27(Current Studies on The Information Content of Accounting Earnings), 85-105.

- Pyo, Y., & Lustgarten, S. (1990). Differential intra-industry information transfer associated with management earnings forecasts. *Journal of Accounting and Economics*, 13(4), 365-379.
- Ramnath, S. (2002). Investor and analyst reactions to earnings announcements of related firms: An empirical analysis. *Journal of Accounting Research*, 40(5), 1351-1376.
- Rogers, Jonathan L. and Schrand, Catherine M. and Zechman, Sarah L. C., Do Managers Tacitly Collude to Withhold Industry-Wide Bad News? (October 30, 2014). Chicago Booth Research Paper No. 13-12.
- Rogers, J. L., & Van Buskirk, A. (2013). Bundled forecasts in empirical accounting research. *Journal of Accounting and Economics*, 55(1), 43-65.
- Rogers, J. L., Skinner, D. J., & Van Buskirk, A. (2009). Earnings guidance and market uncertainty. *Journal of Accounting and Economics*, 48(1), 90-109.
- Skinner, Douglas J. (1994). Why firms voluntarily disclose bad news? *Journal of Accounting Research* 32 (1) (Spring): 38-60.
- THOMAS, J., & ZHANG, F. (2008). Overreaction to intra-industry information transfers? *Journal of Accounting Research*, 46(4), 909-940.
- Tse, S., & Tucker, J. (2010). Within-industry timing of earnings warnings: Do managers herd? *Review of Accounting Studies*, 15(4), 879-914.
- Verrecchia, R. E. (1983). Discretionary disclosure. *Journal of Accounting and Economics*, 5, 179-194.
- Woon-Oh Jung, & Kwon, Y. K. (1988). Disclosure when the market is unsure of information endowment of managers. *Journal of Accounting Research*, 26(1), 146-153.