AN ABSTRACT OF THE DISSERTATION OF

Tonya M. Edinger for the degree of Doctor of Philosophy in Business Administration on June 5, 2020.

Title: Dissemination of Information: Financial and Nonfinancial Disclosure Channels

Abstract approved:

_____

Jared A. Moore

The objective of this research is to examine capital market determinants and implications associated with voluntary sustainability disclosures and the extent to which the informativeness of disclosure innovations differs based on given attributes of the financial disclosure and overall information environment.

In determining the quantity of financial information to disclose, managers face a tradeoff between the benefits of reducing information asymmetry among capital market participants and the costs of aiding potential rivals through revelation of proprietary information. Chapter I examines the firm-level disclosure response as competition among potential rivals differs. I operationalize this construct through use of principal component analysis to capture the competition from potential entrants variable.

In Chapter II, I examine the relation between disclosure of nonfinancial information and information asymmetry. I first employ the issuance of a stand-alone sustainability report as a proxy for disclosure of nonfinancial information. I find that the issuance of nonfinancial

information is associated with reduced information asymmetry in the market. The relation is stronger for firms with internal control weaknesses and financial statement complexity, indicating a complementary relation between nonfinancial and financial disclosure in the reduction of information asymmetries in these contexts. In the context of increased organizational complexity, the signal appears to be less informative. In Chapter II, I empirically treat the nonfinancial disclosures as homogeneous in nature by assigning a binary variable to disclosing and non-disclosing firms. In reality, substantial variation between ESG disclosures exists.

Thus, in Chapter III, I relax the homogeneity assumption and address the evident variation in ESG disclosure content through illustrating that alterations in the overall nonfinancial disclosure content prompt a reduction in information asymmetry. I provide some evidence suggesting that, on average, changes to the overall content of ESG disclosures increase incremental informativeness in a capital market context. By doing so, I build on the empirical conclusions from Chapter II by empirically illustrating that the content of nonfinancial disclosures is informative in a capital market context. I illustrate that, while information disclosure innovations are disseminated through both financial and nonfinancial disclosure channels, nonfinancial disclosure innovations are incrementally informative in their reduction of information asymmetry. The analyses conducted in Chapter III seek to demonstrate that investors heed changes to the textual content of the nonfinancial information in addition to changes in the textual content of financial information.

The analysis advances the literature by empirically demonstrating the extent to which the content of voluntary nonfinancial disclosures enhances the information environment relative to financial disclosure information. I also conduct the analysis in the context of alternative

information environments, namely in the presence of organizational complexity, internal control weaknesses, and financial statement complexity (i.e. diminished readability). I hypothesize that nonfinancial disclosure innovations are more informative in the presence of more organizational complexity, more internal control weaknesses, and less readable financial disclosures.

Taken together, the results are consistent with the argument that capital markets utilize voluntary nonfinancial disclosure information in tandem with financial information. The findings contribute to the understanding of sustainability disclosures and the overarching role of such disclosures.

Dissemination of Information: Financial and Nonfinancial Disclosure Channels

by
Tonya M. Edinger

A DISSERTATION

Submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Presented June 5, 2020
Commencement June 2021

Doctor of Philosophy dissertation of Tonya M. Edinger presented on June 5, 2020.

APPROVED:

_____

Major Professor, representing Business Administration

_____

Dean of the College of Business

_____

Dean of The Graduate School

I understand that my dissertation will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my dissertation to any reader upon request.

_____

Tonya M. Edinger, Author

ACKNOWLEDGEMENTS

I am exceedingly grateful for the guidance, instruction, thoughtful comments, and continual support of my committee chair, Dr. Jared Moore. I would also like to gratefully acknowledge the support and influence of my committee members: Dr. Roger Graham, Dr. Logan Steele, Dr. Jake Rose, and Dr. Christian Langpap. Thank you to my colleagues, for their continual feedback. I would also like to send special thanks to Andy Strat for proofreading and providing guidance. I am truly thankful to the Oregon State University College of Business Faculty for the research support over the past few years and the helpful feedback throughout the process.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

LIST OF TABLES

Table Page

LIST OF TABLES (Continued)

Table Page

Chapter I

Disclosure Choice and the Nature of Competition

1. Introduction

Demand for corporate sustainability and the provision of information through standalone sustainability reports has increased in recent years (Eccles, Perkins, and Serafeim, 2012) to the extent that recent surveys suggest that companies have begun to consider sustainability-related strategies as necessary in order to remain competitive. In determining the quantity of financial information to disclose, managers face a tradeoff between the benefits of reducing information asymmetry among capital market participants and the costs of aiding competitors through revelation of proprietary information. Disclosures intended to convey information to investors is also readily available for observation by a firm's current and potential rivals. Thus, such information can potentially aid rivals in competing with the disclosing firm. This chapter is an exploratory analysis that focuses on determinants of nonfinancial disclosure innovations in the context of a competitive environment. I examine the extent to which competition among potential rivals and existence of proprietary costs influence issuance of nonfinancial disclosure innovations. I also examine whether a tradeoff exist between nonfinancial and financial disclosure innovations.

In this chapter, I examine the extent to which certain aspects of the competitive environment affect environmental and social governance disclosure decisions. I specifically examine whether competition from potential rivals and proprietary costs influence the nonfinancial disclosure choice. In order to do so, I focus on firms that have already made the decision to disclose nonfinancial information in the form of a standalone sustainability report. I conduct the analysis

in the context of nonfinancial disclosure innovation and financial disclosure innovation metrics to quantify the extent to which disclosures have changed over time.

I preface Chapter I analyses with a review of the accounting literature surrounding disclosure choice, competition and proprietary costs. After presenting the current literature, I introduce the Chapter I hypotheses and discuss the sample and data collection process including the measurement of disclosure and nonfinancial disclosure innovations and readability scores. After outlining the empirical methods, I present the empirical results and conclude the chapter.

2. Disclosure Choice, Competition and Proprietary Costs Literature

Disclosure is a manner through which firms reduce information asymmetries between firm managers and capital market participants. Economic theory suggests that, in the absence of costs, rational investors will presume that management has private information and an associated incentive to withhold bad news. In turn, firms will disclose all information to the capital markets (Grossman, 1981; Milgrom, 1981). However, this scenario only holds in the absence of disclosure costs. In practice, many constraints to voluntary disclosure exist. For instance, the potential for harming the firm's competitive position poses a commonly cited constraint on the voluntary disclosure of financial information. The accounting literature refers to this disincentive as a proprietary cost of voluntary disclosure.

The proprietary cost argument is prevalent in the accounting literature. When managers are faced with the prospect of disclosing information regarding their firms' customers, there exists a tradeoff between the cost of aiding rivals (in their potential revelation of proprietary information) and benefits of reducing information asymmetry among capital market participants. Prior disclosure theory (Verrecchia, 1983) predicts a stronger good news bias of disclosure

information in the presence of higher proprietary costs. Nonetheless, relatively little empirical research has focused on the extent to which product market competition affects voluntary disclosure decisions. Prior literature has examined the impacts of product market competition on voluntary disclosure in a financial context. In a financial reporting context, theory generally predicts that competition from potential entrants encourages more voluntary disclosures (Darrough and Stoughton, 1990). Li (2010) observes competition from potential entrants to increase the propensity of financial disclosure.

Early theoretical modeling by Verrecchia (1983) indicates that firms with higher proprietary costs of disclosure will tend to disclose less than those with lower proprietary costs of disclosure. Verrecchia (1990) predicts that there is likely less disclosure in industries containing more intense competition, and therefore higher proprietary costs of disclosure. In a highly competitive industry, one might expect the firm's rivals to aggressively react to the disclosures. Thus, one might expect the incentive structure of highly competitive industries to deter disclosure. Empirical work by Li (2010) present results that support Verrecchia's prediction; the research observes that competition from *existing rivals* tends to decrease financial disclosure quantity. Ali et al. (2014) expand on the prior theoretical modeling to empirically demonstrate a negative relation between industry concentration[1] (a proxy for proprietary costs) and disclosure as well. The authors find both frequency and horizon of management forecasts of annual earnings to be negatively correlated with their U.S. Census industry concentration measure. From the findings, one can infer that firms in the more concentrated industries (as a proxy for higher proprietary costs) tend to disclose less and also tend to do so in a less prompt manner. In the financial context, competition tends to enhance disclosure quality as well. This

[1] The authors calculate the Herfindahl-Hirschman index for each 6-digit NAICS industry in the manufacturing sector. The index is measured by summing the squares of the individual company market shares of the 50 largest public and private companies in the industry (or, all companies in the industry, whichever of these two measures is lower).

observed improvement in quality of financial disclosures in a competitive environment stems mainly from a reduction in profit forecast optimism and a reduction of pessimism in investment forecasts[2].

Dedman and Lennox (2009) further examine the relation between perceived competition and voluntary disclosure in the context of private UK companies. The managers of the UK companies are surveyed regarding the firm's competitive environment in order to construct measures of industry competition. The research observes managers to be more likely to withhold information regarding sales and costs when current or potential competition is perceived to be strong. More profitable firms are also observed to withhold more information.

Ellis, Fee, and Thomas (2012) also contribute to the proprietary costs of disclosure literature by detailing managerial disclosure choice regarding firms' customers over time. Their study tests disclosure theory (Verrecchia, 1983) by hypothesizing the voluntary disclosures of high proprietary cost firms to be significantly larger and also more highly valued than those of low proprietary cost firms. The research hypothesizes and finds proprietary costs to be a significant determinant of the manager's disclosure of customer information. The authors also find higher internal and external auditor quality to increase informativeness of mandatory disclosure, but not that of voluntary disclosures. As the findings in this area appear to be relatively mixed, research by Lang and Sul (2014) builds on the work linking industry concentration, proprietary costs and disclosure as the research in this branch of the literature by citing several prevalent challenges. Two such challenges are the difficulty in quantifying industry concentration and identification of disclosure that is likely to indicate existence of proprietary costs.

[2] The observed associations between competition and disclosure quantity and quality are found to be less pronounced for industry leaders relative to industry followers. This finding is consistent with industry leaders facing less competitive pressure relative to industry followers.

While firm-level voluntary disclosures serve a prominent information conveyance role, firms do not make their disclosure choices independently. When an agent chooses to disclose, non-disclosing agents are also benefactors of the disclosure. In order to make a fully informed disclosure decision, firms must consider the industry consensus behavior as well. Disclosure choice has been evaluated in the context of economic "beauty contests" through use of a game theory approach (Arya and Mittendorf, 2016). In this type of contest, the Arya et al. (2016) shed light on the circumstances in which firms are incentivized to disclose. Using a theoretical game theory approach, the authors suggest that introduction of this type of complementarity brings about a key effect where agents tend to overweight the information that is publicly available. Firms have a tendency to coordinate their disclosure choice as the disclosure imparts information instrumental in the establishment of industry norms. Firms' disclosure choices thereby influence the future disclosure tendencies of other firms as well.

3. Hypothesis Development

3.1. Disclosure Choice and Competition from Potential Entrants

The research has established that firms are sensitive to the existence of public information. I posit that a similar informational cascade exists in the context of nonfinancial disclosure information. Disclosing agents have an opportunity to exert indirect influence on not only their industry peers but also the potential entrants to the industry. Discussions of discretionary disclosure have alluded to how competition might affect the firm's incentives to disclose. With respect to financial information, Darrough and Stoughton (1990) suggest that greater competition from potential entrants can prompt disclosure of bad news. Overall, the

literature suggests that competition through threat of entry likely encourages discretionary disclosure.

The subsequent section builds on the existing literature surrounding proprietary costs of disclosure to incorporate voluntary environmental and social governance (ESG) disclosure decisions based on the premise that manager's disclosure decisions are likely to reflect their firm competitive situation (Darrough and Stoughton, 1990). I examine whether competition from potential rivals influences firms' voluntary ESG disclosure decisions. I focus on firms that have already made the decision to disclose ESG information in a standalone sustainability report. To my knowledge, this is the first study to address proprietary costs of nonfinancial disclosure. I conduct the analysis in the context of the aforementioned nonfinancial disclosure innovation metric.

Disclosing proprietary information is a strategic choice at the firm level. The disclosure literature has generally conceded that firms in possession of favorable market information may display hesitancy to disclose the information; firms may abstain from disclosure in order to alleviate the threat of new rivals that surfaces when the favorable market information has been revealed via disclosure. In the context of disclosure-related class-action securities litigation cases from the time period between 1996 and 2005, Roger and Van Buskirk (2009) examine the disclosure response to litigation. The authors observe a disclosure response post-litigation; there is a negative relation between ligation and disclosure. This result is intuitive in the sense that firms tend towards reduction of the quantity of disclosed information for which they may later be held accountable.

3.2. Competition among Potential Entrants and Nonfinancial Disclosure Innovations

Empirical literature has documented a positive relation between competition from potential entrants and voluntary financial disclosure (Li, 2010). However, the theoretical work surrounding voluntary disclosure in the presence of varying degrees of competition is more nuanced. Theory suggests that when entry costs are low (when threat of entry is high), the incumbent with bad financial news discloses to deter entry. Prior literature also suggests that disclosure of more public information reduces a firm's cost of capital (Easley and O'Hara, 2004). If a firm reveals too much financial information to existing rivals or potential entrants, this can hinder the firm's product market competitiveness. In the case of nonfinancial disclosures, disclosure of ESG-related information generally conveys a benchmark to be met. Thus, we might expect the firm to reveal more nonfinancial information as competition among potential entrants increases. Further, I expect that firms will choose to disseminate more information through a "soft" nonfinancial disclosure channel when proprietary costs of disclosure are high (i.e. high competition among potential rivals). I suspect that competition among potential entrants will prompt nonfinancial disclosure innovations. Therefore, I hypothesize that firms choose to issue nonfinancial disclosures in an attempt to increase the perceived cost of entry, and thereby deter entry. On the other hand, failure to document a significant relation between potential entrants and nonfinancial disclosure innovations will indicate that, beyond the initial strategic decision to disclose the standalone sustainability information, competition from potential entrants bears little influence on nonfinancial disclosure innovations. As competition among potential entrants increases, disclosure innovations will likely increase as well. I formally state the first hypothesis in the alternative form, below.

H1: Ceteris paribus, as competition from potential rivals increases, nonfinancial disclosure innovations increase.

Although I do not make a directional hypothesis surrounding the relation at this point, I also expect nonfinancial disclosure innovations and financial disclosure innovations to be related in some fashion. It is plausible that a tradeoff exists among distinct types of voluntary disclosures. On the other hand, it is also possible that increased financial disclosure innovations may correspond to an increase in nonfinancial disclosure innovations; firms may elect to increase or decrease disclosure homogenously across dissemination channels.

3.3. Proprietary Costs and Disclosure Channels

In the presence of high proprietary costs of disclosure, research has also demonstrated that firms tend to substitute disclosure of nonproprietary information for less disclosure of proprietary information. Glaeser (2017) studies the relation between proprietary information and both corporate transparency and voluntary disclosure. The research employs the use of an external shock to trade secrecy, namely the Uniform Trade Secrets Act. The author also introduces two new measurements to the literature in an attempt to quantify firms' reliance on trade secrecy. The first measure is based on 10-K disclosures and the second is based on litigation outcomes. The study finds that firms with increased reliance on trade secrecy tend to substitute voluntary disclosure of nonproprietary information for less disclosure of proprietary information. Overall, the aggregate implication for trade secrecy is found to be a decrease in corporate transparency. To the extent that firms substitute disclosure of financial information for that of nonfinancial information in the presence of higher proprietary costs, I would expect to observe a significantly negative relation between nonfinancial disclosure innovations and financial disclosure innovations. I state this hypothesis in the alternative form below.

H2: Ceteris paribus, as proprietary costs of disclosure increase, the relation between financial disclosure innovations and nonfinancial disclosure innovations become less positive or more negative.

Although I do not make a directional hypothesis surrounding the relation, I expect that nonfinancial disclosure innovations and financial disclosure innovations to be related in some fashion as well. It is possible that increased financial disclosure innovations may correspond to an increase in nonfinancial disclosure innovations; firms may elect to increase or decrease disclosure homogenously across dissemination channels. On the other hand, it is also plausible that a tradeoff exists among distinct types of voluntary disclosures.

4. Data Collection

4.1. Main Variables

4.1.1. Competition from Potential Entrants

In order to operationalize competition from potential entrants, I conduct a principal component analysis on several commonly employed proxies for competition. Prior studies typically employ industry-average size of plant and equipment as a measure of setup costs for a new player to enter the product market and begin operating as an average firm within that industry (Li, 2010; Sutton, 1991). Both industry-average R&D outlays as well as capital expenditures also reflect necessary investments that potential entrants to must make in order to become competitive with existing rivals. As both variables are likely positively correlated with barriers to entry, I employ both industry-average R&D outlay as well as capital expenditures as proxies for competition from potential entrants. Lastly, I employ product market size as the

variable is likely to be negatively associated with potential competition. The logic behind the use of product market size is that entry is less harmful to the incumbent firm that operates in a product market with higher demand. In addition, a large market size is often associated with high entry barriers such as heavy investments in either PP&E (to increase volume) or technology (to increase price) (Li, 2010). The data for the competition variables is obtained from the segments database and fundamentals annual database of Compustat North America[3].

I define industry-average size of plant and equipment as the weighted average of PP&E of all firms in an industry. I use the firm's market share (calculated as the ratio of its segment sales to industry aggregate sales) as its weight. I define industry-average R&D outlays as the weighted average of R&D of all firms in an industry. I use the firm's market share as its weight. I define industry capital expenditures as the weighted average of capital expenditures of all firms in an industry. Consistent with the prior two proxies, the firm's market share is used as its weight. Product market size is measured as the natural log of industry aggregate sales.

I extract the data for computation of the competition variable from the Compustat segments and fundamentals annual database of Compustat North America. I do so for the time frame of 2014-2019. I obtain the data on net sales, research and development, capital expenditures from Compustat segments. I only retain business segments with valid primary four-digit SIC codes. If a business segment has identical SIC codes under the same firm, I merge the segments into one and sum all of the financial items. The segment data is then merged with that of Compustat fundamentals annual data. If the firm is missing segment information, I treat the firm as if there is just one segment. I then calculate the industry-wide variables of industry-average PP&E, industry-average R&D, Industry Capital Expenditures, and product market size.

[3] Note that, as per SAFAS No. 14, multi-industry firms are required to disclose operating profits, identifiable assets, revenues, depreciation and amortization, capital expenditures, and research and development for their significant industry segments.

Descriptive Statistics and Correlation Matrix are included below in Table 18 and Table 19, respectively.

I also incorporate an alternative, firm-level proxy to capture the competition from potential entrants construct (COMPETITIONPCA2). I employ principal component analysis as a means to operationalize the firm-level construct. Utilizing the same logic as the prior proxy, I utilize net property, plant and equipment scaled by prior year total assets (CAPIN), research and development scaled by total assets, and capital expenditures scaled by total assets in the principal component analysis. Appendix C contains Descriptive Statistics and Correlation Matrix for the alternative data set utilizing the aforementioned alternative proxy for competition from potential entrants, COMPETITIONPCA2.

4.1.2. Financial and Nonfinancial Disclosure Innovations

I hand-collect 1,457 ESG disclosures as PDFs from the firm websites and the GRI website over the 2014-2019 time frame. I obtain the disclosure publication year from the GRI website due to inter-firm ESG disclosure name inconsistencies. In order to analyze the content of the PDF documents, I convert the ESG disclosure PDFs to text files. I download the 10-K Annual reports from the SEC company filings that are publicly available on EDGAR. I bulk download the raw text filings via Python by converting html to text files. I match the sustainability reports to the existing data by firm name. After matching the sustainability reports with their respective financial disclosure, Compustat, IBES, CRSP, and Thomson Reuters data, the resulting sample contains 540 observations.

To calculate the degree to which nonfinancial and financial disclosures vary from one year to the next, I use a machine learning approach. After creating a dictionary from the words in all of the documents, I convert each document into a bag of words (words and counts for each

document). Then, I employ word embeddings as a manner to represent the text. In Python, I use the Gensim library to develop Word2Vec embedding from all of the sentences in each document. Word2Vec is an algorithm with the capacity to learn word embedding from a text corpus. The word embedding approach yields a vector representation of words that capture intricacies regarding their meaning. Mikolov et al. (2013) recently introduced the Word2Vec software, and the approach has begun to gain traction for word embeddings, although the approach is not yet prevalent in the accounting literature. The learning models behind the software are described in Mikolov et al. (2013)[4]. To create the similarity score, I train a Word2Vec model on the disclosure data.

Based on the entire corpus of text, the algorithm trains a set of fixed-length dense and continuous-valued vectors. In the embedded space, each word represents a point. Points are learned and then shuffled around based on the words that surround the target word. Therefore, words are defined based on the words that surround them. The premise behind Word2Vec is to train a simple neural network with a hidden layer to perform a certain task, but the neural network will not be used for the task that it has been trained on. Instead, I use the learned weights of the hidden layer.

I load and organize the text into sentences and provide them to the constructor of a new Word2Vec instance. I tokenize each sentence (i.e. divide each sentence into words), remove stop words, non-dictionary words, and lemmatize the body of text as a means of pre-processing the disclosures. The process of lemmatization allows for identification of more complex forms and for the regrouping of the forms that correspond to the same root with different inflexions. Through this process, I change verbs to infinitives, plural words to singular, and group forms that correspond to the same root.

[4] Word2vec is based on the skip-gram model.

I utilize the word embeddings model to create a vector that represents each disclosure document. Then, I compare this vector to that of the prior year through use of the cosine similarity. The cosine similarity is determined by the angle between the vectors and not their lengths (magnitudes). The cosine similarity is robust to length and repetition.

Disclosure innovations represent the extent to which a firm changes the textual content of the disclosure from one year to the next. I focus on the textual narratives rather than the quantitative information in assessing the extent to which each disclosure is boilerplate from one year to the next. The construct captures the extent to which the disclosure is boilerplate from one year to the next. The technique utilizes dictionary words and omits numbers (aside from longhand numbers that are also dictionary words). By focusing on the textual content of the disclosure, I also assume that the quantitative content of the disclosures is captured through quantitative control variables or that the textual content captures the essence of the quantitative content.

In order to ensure validity of the construct, I conduct face validity assessment for some of the disclosures to ensure robustness of the measure. When the algorithm reports low levels of innovation, the disclosures are largely boilerplate from one year to the next. To the extent that the level of innovation increases, I observe more changes to the textual content of the disclosure. At higher levels of innovation, the firm has expended more effort in altering the disclosure content.

4.1.3. Proxies for Proprietary Costs

In accordance with prior proprietary cost literature, I first utilize industry concentration, INDCONC, as an indicator of proprietary information (Ali et al., 2014). Industry concentration is purportedly an adequate proxy for proprietary costs of disclosure if industry concentration

proxies for intensity of industry competition, level of innovation in the industry, or extent to which disclosures by firms in the industry provide more substantive information about future industry demand (Ali et al., 2014).

I also utilize market-to-book ratio, BM, as an indicator of proprietary information (Brockman, Khurana, and Martin, 2008; Bamber and Cheon, 1998). Growth opportunities indicate presence of profitable investments. When faced with prevalent growth opportunities, managers exhibit reluctance surrounding the revelation of information that could potentially dissipate the value of the existing growth opportunities (Bamber and Cheon, 1998).

4.2. Descriptive Statistics

TABLE 1
Chapter I Summary Statistics

VARIABLE	MEAN	25TH PERCENTILE	MEDIAN	75TH PERCENTILE	STANDARD DEVIATION
FININNOVATION	0.0179	0.0029	0.0047	0.0081	0.0469
NFINNOVATION	0.0930	0.0265	0.0502	0.1019	0.1245
INDCONC	0.2765	0.1802	0.2374	0.3669	0.1160
COMPETITIONPCA	0.0001	-0.0666	-0.0666	-0.0665	1.0009
SIZE	10.1650	9.4324	10.1305	10.9866	1.1968
BM	4.0996	2.0881	2.9731	4.6276	2.8537
LEVERAGE	0.5032	0.3670	0.4916	0.6411	0.1696
CONCERNS	0.7711	0.0000	0.5000	1.5000	0.8162
STRENGTHS	1.0678	0.7500	1.1111	1.4167	0.4912
OWNERSHIP	0.0013	0.0006	0.0010	0.0019	0.0009
ROI	0.0898	0.0490	0.0797	0.1247	0.0458
N	540				

5. Empirical Methods

I expect a tradeoff to exist between financial and nonfinancial disclosure innovations and for the functions to be jointly determined. Thus, throughout Chapter I, I utilize three-stage least squares (3SLS) as a means to avoid simultaneity bias. I implement a simultaneous equations approach when estimating equations (i) and (ii) for H1, below. The use of a system of simultaneous equations the relation among the constructs aids in taking potential simultaneity into account. 3SLS estimates systems of structural equations where some equations contain endogenous variables among the explanatory variables. The method utilizes an instrumental-variables approach in order to produce consistent estimates and generalized least squares (GLS) to account for the correlation structure in the disturbances across equations (Greene, 2012).

5.1. Empirical Model: H1

H1: As competition from potential rivals increases, nonfinancial disclosure innovations increase.

$$\text{NFINNOVATION} = \beta_0 + \beta_1 \text{COMPETITIONPCA} + \beta_2 \text{FININNOVATION} + \theta \text{Controls} + \epsilon \quad \textbf{(i)}$$

$$\text{FININNOVATION} = \beta_0 + \beta_1 \text{COMPETITIONPCA} + \beta_2 \text{NFINNOVATION} + \theta \text{Controls} + \epsilon \quad \textbf{(ii)}$$

5.2. Empirical Model: H2

H2: As proprietary costs of disclosure increase, the relation between financial disclosure innovations and nonfinancial disclosure innovations decreases. That is, financial disclosure innovations are decreasing in nonfinancial disclosure innovations in the presence of higher proprietary costs of disclosure.

$$\text{NFINNOVATION} = \beta_0 + \beta_1 \text{FININNOVATION} + \beta_2 \text{FININNOVATION} * \text{INDCONC} + \beta_3 \text{INDCONC} + \theta \text{Controls} + \epsilon \quad \textbf{(iii)}$$

$$\text{FININNOVATION} = \beta_0 + \beta_1 \text{NFINNOVATION} + \beta_2 \text{NFINNOVATION} * \text{INDCONC} + \beta_3 \text{INDCONC} + \theta \text{Controls} + \epsilon \quad \textbf{(iv)}$$

TABLE 2
Correlation Matrix

	FININNOVATION	NFINNOVATION	INDCONC	COMPETITIONPCA	SIZE	BM	LEVERAGE	CONCERNS	STRENGTHS	OWNERSHIP	ROI
FININNOVATION	1.00										
NFINNOVATION	-0.07	1.00									
INDCONC	-0.06	-0.27***	1.00								
COMPETITIONPCA	-0.01	-0.03	0.12***	1.00							
SIZE	0.03	0.36***	-0.20***	-0.10**	1.00						
BM	0.01	-0.10**	-0.10**	-0.04	-0.17***	1.00					
LEVERAGE	-0.05	0.07	-0.02	-0.01	0.20***	0.38***	1.00				
CONCERNS	-0.06	0.37***	-0.21***	-0.02	0.50***	-0.14***	0.25***	1.00			
STRENGTHS	0.09**	0.11**	-0.37***	-0.10**	0.34***	0.27***	0.18***	0.24***	1.00		
OWNERSHIP	-0.14***	-0.26***	0.17***	0.07	-0.61***	-0.23***	-0.16***	-0.32***	-0.60***	1.00	
ROI	0.09**	-0.24***	-0.00	-0.03	-0.30***	0.55***	-0.16***	-0.37***	0.26***	-0.25***	1.00
N	541										

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

6. Empirical Results

6.1.1. H1 Results: Nature of Competition and Nonfinancial Disclosure Innovations

TABLE 3 details the empirical 3SLS results surrounding the first hypothesis. I hypothesize the existence of a significantly positive relation between competition from potential entrants and both nonfinancial disclosure innovations and financial disclosure innovations. I fail to reject the null of no relation between the independent variable of interest, COMPETITIONPCA, and NFINNOVATION. I also fail to reject the null hypothesis of no relation between COMPETITIONPCA and FININNOVATION. The observed results provide evidence suggesting that firms do not issue nonfinancial disclosures as a response to competition among potential entrants. Despite this finding, it is worth noting that the industry concentration variable is observed to be significantly negative. The result indicates that as industry concentration increases, nonfinancial disclosure innovations decrease. This provides preliminary evidence supporting the second hypothesis.

TABLE 4 details the empirical 3SLS results surrounding the first hypothesis with use of an alternative, firm-level proxy for the competition from potential entrants construct (COMPETITIONPCA2). Again, I hypothesize the existence of a significantly positive relation between the firm-level competition from potential entrants variable and both nonfinancial disclosure innovations and financial disclosure innovations. I reject the null hypothesis of no relation between COMPETITIONPCA2 and NFINNOVATION. I observe a significantly negative relation between the firm-level competition from potential entrants variable and nonfinancial disclosure innovations.

In TABLE 4, I also document a tradeoff between financial disclosure innovations and nonfinancial disclosure innovations. I observe a significantly negative relation between financial

disclosure innovations and nonfinancial disclosure innovations. As financial disclosure innovations increase, nonfinancial disclosure innovations are observed to decrease. The variable FININNOVATION is observed to be significant at the 10% level. As nonfinancial disclosure innovations increase, financial disclosure innovations are observed to significantly decrease. The variable NFINNOVATION is observed to be significant at the 5% level.

TABLE 3
Competition Among Potential Entrants and Disclosure Innovations

	NFINNOVATION	FININNOVATION
FININNOVATION	0.478	
	(0.397)	
COMPETITIONPCA	-3.000	-0.101
	(6.098)	(5.390)
SIZE	0.00731***	-0.00389
	(0.00232)	(0.00249)
STRENGTHS	-0.0130**	
	(0.00542)	
CONCERNS	0.0138***	
	(0.00331)	
LEVERAGE	-0.0316**	-0.0122
	(0.0141)	(0.0134)
INDCONC	-0.109***	-0.0290
	(0.0217)	(0.0184)
NFINNOVATION		-0.0311*
		(0.0175)
BM		-0.00104
		(0.000893)
OWNERSHIP		-12.38***
		(3.176)
CONSTANT	-0.157	0.0883
	(0.409)	(0.364)
OBSERVATIONS	540	
R^2	0.145	0.044

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

TABLE 4
Competition Among Potential Entrants (Firm-Level) and Innovations

	NFINNOVATION	FININNOVATION
FININNOVATION	-0.226*	
	(0.126)	
COMPETITIONPCA2	-0.0135*	-0.000398
	(0.00715)	(0.000435)
SIZE	0.0111*	-0.000690
	(0.00629)	(0.000530)
STRENGTHS	-0.0102	
	(0.0129)	
CONCERNS	0.0191*	
	(0.00974)	
LEVERAGE	-0.0229	-0.000769
	(0.0343)	(0.00246)
INDCONC	-0.0533	-0.00312
	(0.0540)	(0.00333)
NFINNOVATION		-0.0189**
		(0.00828)
BM		-0.000181
		(0.000170)
OWNERSHIP		-1.040
		(0.667)
Constant	-0.00775	0.0183***
	(0.0616)	(0.00630)
Observations	246	
R^2	0.059	0.021

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

6.1.2. H2 Results: Proprietary Costs of Disclosure and Disclosure Innovations

Building on the prior results, TABLE 5 presents the empirical results from the estimation of equation (iii) and equation **(**iv**)**. The empirical model intends to address hypothesis two surrounding proprietary costs. I observe a negative relation between financial disclosure innovations and nonfinancial disclosure innovation. That is, as nonfinancial disclosure innovations increase, financial disclosure innovations are observed to decrease. The coefficient on NFINNOVATION is significant at the 10% level. This finding signifies that a tradeoff likely exists between nonfinancial disclosure innovations and financial disclosure innovations. I also observe a significantly negative relation between industry concentration (a proxy for proprietary costs) and both nonfinancial disclosure innovations and financial disclosure innovations. This significantly negative relation is in accordance with the expectation that as proprietary costs of disclosure increase, disclosure decreases. A similar phenomena surrounding financial disclosure has been documented in prior empirical literature (Ali et al., 2014), this is the first study of my knowledge to address the nonfinancial disclosure implications of proprietary costs. The result implies that, as proprietary costs of disclosure increase, firms tend to disclose less nonfinancial information.

Recall that hypothesis two suggests that, as proprietary costs of disclosure increase, the relation between financial disclosure innovations and nonfinancial disclosure innovations decreases. The empirical results indicate the interaction term (NFINNOVATION*INDCONC) and (FININNOVATION*INDCONC) are statistically significant at the 10% level. I reject the null hypothesis. However, the direction of the relation is positive rather than negative. This is an interesting finding. The model documents that, as nonfinancial disclosure innovations increase, financial disclosure innovations decrease. Then, this further result indicates that, as

proprietary costs of disclosure increase, firms are less likely to substitute one form of disclosure for another. That is, financial disclosure innovations are increasing in nonfinancial disclosure innovations in the presence of higher proprietary costs of disclosure.

TABLE 6 reinforces the empirical results for the second hypothesis through use of an alternative proxy for proprietary costs, BM. The model documents a significant tradeoff between nonfinancial disclosures and financial disclosure innovations. Nevertheless, in the presence of higher proprietary costs, the negative relation is attenuated. A tradeoff still remains between nonfinancial disclosure innovations and financial disclosure innovations, but is diminished in the presence of higher proprietary costs. It is interesting to observe that the coefficient on the interaction terms is positive for both of the proxies for proprietary costs.

TABLE 5
Proprietary Costs and Disclosure Innovations

	NFINNOVATION	FININNOVATION
FININNOVATION	-73.04	
	(44.82)	
INDCONC	-4.265*	-0.868**
	(2.543)	(0.442)
FININNOVATION*	270.8	
INDCONC	(168.2)	
SIZE	0.137	0.0174
	(0.0950)	(0.0130)
STRENGTHS	-0.176	
	(0.166)	
CONCERNS	-0.141	
	(0.128)	
LEVERAGE	0.111	0.0446
	(0.325)	(0.0487)
OWNERSHIP	-167.8	-23.90**
	(122.9)	(11.82)
NFINNOVATION		-2.598*
		(1.337)
NFINNOVATION*		10.50*
INDCONC		(5.516)
ROI		-0.454*
		(0.272)
CONSTANT	0.402	0.143
	(0.761)	(0.104)
OBSERVATIONS	541	
R^2	-90.878	-6.925

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

TABLE 6
Proprietary Costs and Disclosure Innovations

	NFINNOVATION	FININNOVATION
FININNOVATION	-41.03***	
	(15.92)	
BM	-0.0861**	-0.0300***
	(0.0348)	(0.00850)
FININNOVATION *	8.445**	
BM	(3.383)	
SIZE	-0.00480	-0.0123*
	(0.0340)	(0.00636)
STRENGTHS	-0.130	
	(0.0798)	
CONCERNS	0.0608	
	(0.0426)	
LEVERAGE	-0.0421	-0.0296
	(0.179)	(0.0311)
OWNERSHIP	-85.10*	-39.17***
	(49.81)	(10.94)
NFINNOVATION		-1.793***
		(0.498)
NFINNOVATION *		0.460***
BM		(0.132)
ROI		-0.405**
		(0.192)
CONSTANT	0.828	0.377***
	(0.507)	(0.116)
OBSERVATIONS	541	
R^2	-21.174	-2.934

Standard errors in parentheses
*$p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

7. Summary

The results in this chapter aim to provide an exploratory analysis surrounding the determinants of nonfinancial disclosure innovations in a competitive environment. I examine whether firms employ ESG disclosure innovations as a strategic means to deter entry, whether a tradeoff exists between nonfinancial disclosure innovations and financial disclosure innovations, and aim to pinpoint how this tradeoff is altered in the presence of proprietary costs. I observe evidence regarding firm utilization of ESG disclosure innovations as a strategic means to deter entry of potential rivals. I observe a significantly negative relation between nonfinancial disclosure innovations and financial disclosure innovations. I also observe that, as proprietary costs increase, the tradeoff is attenuated or eliminated. Overall, Chapter I sheds some light on the potential determinants of nonfinancial disclosure innovations. This chapter acts as a preface to the subsequent chapters that discuss the implications of nonfinancial disclosure.

Chapter II

Nonfinancial Disclosure and Information Asymmetry

1. Introduction

In this chapter, I examine the extent to which dissemination of voluntary nonfinancial disclosures are of relevance to the capital market. While corporate support for various reporting initiatives has varied throughout time, a consistent awareness has persisted regarding the efficacy through which accounting financial reports both capture and accurately present multiple different dimensions of corporate value. In the context of environmental and social governance (ESG) information, traditional financial reports do not always adequately portray a full illustration of the multiple dimensions of firm value, especially with respect to ESG information.

In turn, there exists a growing propensity for firms to issue stand-alone nonfinancial reports[5] to supplement the existing financial information. The quantity of companies that issue sustainability reports has grown incrementally over time, with fewer than 30 firms issuing the reports in the early 1990s to more than 7,000 in 2016. Over the last decade, firms have increasingly opted to follow the Global Reporting Initiative (GRI) guidelines to report sustainability information. In 2006, fewer than 700 firms worldwide followed the GRI guidelines. By 2016, over 4,000 firms were in compliance. A 2016 report, jointly produced by KPMG International, GRI, and the United Nations Environmental Program (UNEP) cites a surge

[5] A sustainability report is a firm-issued non-financial report than provides information to investors, stakeholders and the general public about the firm's activities around social, environmental, and governance issues, either in the form of a stand-alone report or as part of an integrated report. An integrated report is a single document that is intended to present and explain both the company's financial and nonfinancial performance.

in sustainability reporting instruments in place; there are currently almost 400 instruments in 64 countries (Cheng, 2017).

Overall market interest in nonfinancial information is increasing as well. One cited reason for the growth of interest in nonfinancial information has been the increasing percentage of entity market value that may be attributed to intangible assets. Also, there has been growth in assets under management by socially responsible investment (SRI) funds so the nonfinancial information has become increasingly relevant to investment decisions. Although disclosure of ESG information remains voluntary in the United States, an increasing number of companies integrate sustainability information into their capital allocation decisions (Khan, 2016). With the growing use of economic valuation models to quantify environmental implications, environmental information has been of greater interest in the global market than the social information (Eccles, 2011). Specifically, market interest in ESG transparency has been on the rise and it has been suggested that ESG disclosure quality has begun to proxy for management quality and a firm's corresponding ability to grow the business[6].

Based on the trend towards disclosure of sustainability data and the associated implementation of instruments to achieve transparency and accountability, the objective of the research that follows is to examine the relation between voluntary nonfinancial disclosures and information asymmetry for a sample of firms in the United States. By doing so, the ensuing research explores some capital market implications of voluntary nonfinancial disclosure and the characteristics of such disclosures. This dissertation builds on the notion that the information contained in standalone sustainability reports acts to supplement disseminated financial information in a capital market context. The results offer insights as to whether environmental

[6] Brown and Hillegeist (2003) provide empirical evidence indicating that better disclosure quality is negatively associated with information asymmetry in the market.

disclosures provide information that is incrementally useful to investors over financial information, thereby decreasing information asymmetry in the market. In Chapter II, I extend the empirical analysis to explore capital market implications of the nonfinancial information in alternative information environments, namely contexts marked by internal control weaknesses, organizational complexity, and financial disclosure complexity. The research is motivated by the absence of extensive research in this subset of the disclosure literature. The majority of research in this area is based on financial disclosures (as opposed to the nonfinancial disclosures discussed here). Further, Verrecchia (2001) specifically calls for empirical studies that examine the relation between disclosure and information asymmetry and cites that the overall link between disclosure and information asymmetry has proven to be elusive. Very little research has examined the actual content of the ESG disclosures. Thus, the research at hand seeks to further understand the complementarity[7] between financial and nonfinancial disclosures in the context of information asymmetry in the sense that the different disclosure types collectively act to inform the market. There is an interplay such that different dissemination channels improve the overall informational benefits.

Despite their increasing prevalence, environmental disclosures are largely unregulated within the United States; minimum disclosure requirements rarely exist. In general, accounting information can be perceived as a public good. Shareholders pay for information production in an implicit way. However, the shareholders are unable to charge investors for the ensuing use of the information. In this roundabout way, investors free ride on information that is implicitly paid for by shareholders. The end result is an overall underproduction of information (Watts and Zimmerman, 1986). Such is the justification for disclosure regulations. As such, I posit

[7] Please note that I use the word "complementarity" in a loose sense, signifying that information contained within different disclosure types are utilized in tandem.

whether, in the absence of regulations that mandate environmental disclosure, the voluntary disclosures themselves help to reduce an informational gap in the market. I also seek to examine the context under which this proves to be true. Healy et al. (2001) indicate that many fundamental questions about the demand for, and effectiveness of, financial reporting and disclosure regulation in the economy remain unanswered. The research at hand aims to contribute to this ongoing disclosure discussion, amongst other streams of literature.

Consistent with economic theory, I document a negative association between voluntary nonfinancial disclosures and the level of information asymmetry between investors. Nonfinancial disclosure acts to decrease the amount of private information relative to public information. In turn, incentives to seek additional private information decrease. The results are robust to use of three distinct dependent variable measurements, intended to proxy for investor heterogeneity. The results also prove robust to subsequent propensity score matching procedures.

This chapter adds to the literature in the following ways. The research extends recent research on the relation between disclosure and information asymmetry by documenting a negative relation between nonfinancial disclosure and information asymmetry in the market. Perhaps most closely related to this section in terms of data richness and approach is that of Dhaliwal (2014); my research corroborates this recent publication that documents an association between voluntary corporate social responsibility[8] disclosure and analyst forecast accuracy (Dhaliwal, 2014).

The remainder of Chapter II is organized as follows. Section 2 contains a literature review and development of hypotheses. The sample selection procedure, descriptive statistics, and a

[8] Hereby, CSR.

correlation table are provided in Section 3. Section 4 details the empirical methods and overall research design, followed by empirical results in Section 5. Section 6 offers a robustness check in the form of propensity score matching and additional analysis. I conclude in Section 7 with a synopsis of the chapter and some potential future revisions. Detailed variable descriptions are included in Appendix A.

2. Literature Review and Hypothesis Development

2.1. Nonfinancial Disclosures and Firm Value

Neoclassical economic theory explicitly assumes that profit maximization is the core objective of the corporation, subject to capacity constraints. Shareholders are deemed to be the residual claimant and provide financial capital for the firm's operation (Jensen and Meckling, 1976). However, different corporations vary markedly in their pursuit of profit maximization. For instance, corporations choose whether to place an emphasis on the internalization of externalities from their operations. The externalities could impact the environment, other stakeholders, etc. (Eccles et al., 2012). There is a strategic nature to dealing with externalities; sustainability reporting is a known outlet for communicating relevant environmental information. Extant accounting research also reports that capital markets utilize environmental disclosures to inform their assessment regarding the firms' ability to manage environmental risk exposure (Barth and McNichols, 1994).

There are various benefits associated with the internalization of negative environmental externalities. The act of doing so may influence firm value through its effect on firm risk, including supply chain, litigation, regulatory, and product and technology risk (Dhaliwal et al., 2012). Matsumura et al. (2014) find that markets penalize all firms for their carbon emissions but an additional penalty is imposed on firms that do not disclose emissions information; the

research demonstrates that capital markets impound both the negative environmental externality of carbon emissions as well as the associated act of voluntarily disclosing the information in firm valuations.

By disclosing valid nonfinancial information in the form of sustainability reports, the disclosing firm takes a step forward in rectifying inefficiencies that are derived from a failure to fully calculate all social costs and failure to internalize environmental externalities. A short-term focus with exclusively shareholders in mind could potentially lead to both an inter-temporal loss of profit and negative externalities imposed upon the stakeholders[9]. Research has argued that not meeting stakeholder demands can destroy stakeholder value, hinder the company's ability to hire premium talent and potentially lead to punitive fines (Sen et al., 2001). In turn, Eccles et al. (2012) demonstrate that high sustainability companies significantly outperform their counterparts over the long-term, both in terms of stock market and accounting performance.

In a mandated disclosure context, results often differ, and Chen, Hung, and Wang (2018) observe contradictory results. Chen et al. (2018) conduct their research based on China's 2008 mandate that requires for firms to disclose their CSR activity. The research adopts a difference-in-differences research design and observes a negative relation between mandatory CSR reporting and profitability post-mandate. Interestingly, the study cites evidence indicating favorable reduction in environmental externalities as a result of the mandate; areas that were affected by this mandate experience a reduction in industrial wastewater and SO_2 emissions. Based on this research, one can tentatively conclude that the mandate acts to reduce environmental externalities, but also tends to decrease firm profitability. Christensen, Floyd, Liu, and Maffett (2017) also examine the implications of mandatory social responsibility

[9] For instance, this short-term focus may result in neglect of necessary investments in process and product quality and safety.

disclosure. The context for their study is that of mine owners that are required to include the safety records within the financial reports. For the SEC-registered mine owners this safety record required to be contained in the financial report. Nevertheless, this safety record information is available for all mine owners. Thus, the authors exploit this novel context to examine the implications of the social responsibility disclosure information. The authors observe inclusion of safety records within the financial reports to decrease mining-related citations and injuries, but also reduce labor productivity.

The research at hand builds upon the existing literature that links nonfinancial disclosure and firm value. This chapter seeks to further demonstrate that the nonfinancial disclosures send a signal to the market that, in turn, acts to decrease information asymmetry.

2.2. Credibility of Nonfinancial Disclosures

Voluntary environmental disclosures constitute a shift in the firm-level information environment; the related research is burgeoning. Due to the voluntary nature of environmental disclosures, it is unclear as to whether firms are incentivized to provide credible disclosures, and some branches of the literature focus on differentiating between good and bad quality. Based on economic theory, sellers have an incentive to market poor quality merchandise if the returns for good quality accrue to a group, rather than an individual agent. As a result, market participants will tend to experience a reduction in average quality and overall size of the market (Akerlof, 1970). In the economic model at hand, "trust" is relatively important; market participants may not be able to distinguish a credible environmental disclosure from its unreliable counterpart.

Alternative streams of research are more wary of environmental disclosures, cautioning that the nonfinancial information may not be more than "cheap talk" and that environmental disclosures may be employed as a legitimizing tool to influence societal perceptions. The

legitimacy theory of social disclosure suggests that companies with poorer environmental performance would require more extensive environmental disclosures in their financial reports in order to legitimize their role in society. The theory posits that extent of financial report environmental disclosure is a function of exposure to public pressure in social and political contexts. If this theory were to hold true, one would expect firms with poor performance (and, thus, more exposure) to be more likely to issue environmental disclosures. However, much of the previous literature has failed to document a relation, positive or negative, between performance and disclosure (Fekrat et al., 1996).

The adoption of social and environmental policy has also been argued to destroy shareholder wealth (Friedman, 1970) where sustainability is deemed to be an agency cost. This view of social and environmental policy is in line with legitimacy theory in that managers receive private benefits from integrating the environmental and social policies within their organizations. However, it is argued that the act of doing so may have negative financial implications. Jensen (2001) states that "companies that try to do so either will be eliminated by competitors who choose not to be so civic minded, or will survive only by consuming their economic rents in this manner."

The theoretical work surrounding legitimacy theory has indicated that managers may have incentives to make self-serving voluntary disclosures and the validity of such disclosures has been brought into question. For instance, consistent with legitimacy theory, Cho (2007) documents a negative relation between environmental performance and disclosure. However, evidence in the literature is mixed and recent research has also evidenced that voluntary disclosures may be both credible and valuable. For instance, Piotroski (1999a) finds that the expansion of segment reporting disclosures is associated with an increase in analysts' forecast

accuracy and a decline in dispersion. In a nonfinancial disclosure context, Dhaliwal et al. (2014) find that CSR disclosure is associated with an increase in analysts' forecast accuracy.

In contrast to legitimacy theory, Al-Tuwaijiri et al. (2004) find a significant, positive association between environmental performance (as measured by the ratio of hazardous waste recycled to hazardous waste generated) and environmental disclosure. Yet another stream of research examines the association between corporate social responsibility (CSR) *performance* and information asymmetry (Cho, 2007). Cho (2007) evidence that the scores themselves are informative, irrespective of whether the performance measure is positive or negative.

2.3. Signaling

Economic theory also predicts that firms with 'good type' have an incentive to separate themselves from firms with 'bad type' in order to avoid an adverse selection problem (Milgrom, 1981) and recent literature has empirically validated that voluntary environmental disclosures signal type to investors (Clarkson, 2013). The good environmental performers signal their type through use of the voluntary disclosures; ease of imitation by the poor environmental performers is low. Dhaliwal et al. (2011) claim that firms with better CSR performance are more likely to signal their long-term focus to the market through use of voluntary disclosure. Transparent voluntary environmental disclosures also act to increase firm value, given that they are perceived as credible by investors and are incrementally informative with respect to the firm's environmental performance. Clarkson et al. (2011) reports a positive relation between voluntary environmental disclosure scores and firm value after controlling for environmental performance. Khan et al. (2016) provide recent evidence that investments in sustainability issues are shareholder-value enhancing. Prior research has also documented that, ceteris paribus, firms whom elect to provide more voluntary disclosures have higher market values (Healy and Palepu

2001). It is, therefore, rational to assume that standalone sustainability disclosures may also provide incremental information with potential relevance to investor decision-making.

2.4. Nonfinancial Disclosure and Earnings Management

In stark contrast to the bulk of the earnings management literature that has been largely dominated by agency-based predictions of managerial opportunism, recent research has empirically demonstrated that firms exhibiting corporate social responsibility also are less likely to (i) manage earnings through discretionary accruals and (ii) manipulate real operating activities (Kim et al., 2012). These findings of Kim et al. (2012) suggest that ethical concerns are likely to drive the managers to produce high-quality financial reports. To the extent that investors also perceive disclosing firms[10] in this socially responsible context, the disclosed information may act as a signal of earnings quality. Kim et al. (2012) evidence that socially responsible firms which expend resources in implementing CSR practices to meet ethical expectations of society are also less likely to manage earnings, thus providing more transparent and reliable financial information.

2.5. Voluntary Disclosure and Information Asymmetry

Information and incentive problems tend to impede the efficient allocation of scarce resources within the capital market. Disclosure has been shown to play a role in the mitigation of these issues. Prior theoretical modeling indicates that firms engage in voluntary disclosure in efforts to mitigate investor uncertainty (Dye, 1985; Lewellen and Shanken, 2002; Pastor and Veronesi, 2003) and decrease information asymmetry (Diamond, 1985; Diamond and Verrecchia, 1991). Kim and Verrecchia (1994) theorize that voluntary disclosures reduce information asymmetries for both informed and uninformed investors alike; the result is an

[10] That is, firms issuing voluntary disclosures in the form of standalone sustainability reports.

increased level of stock liquidity[11]. The overall notion suggests that low levels of disclosure exacerbate information asymmetry between firms and potential investors. Increases in information asymmetry lead to an increase in the cost of capital by introducing 'adverse selection' between buyers and sellers of the firm's shares. Adverse selection decreases liquidity in the firm's shares, thereby leading to issuance of shares at a discount. The discount then limits the funds firms receive from the issue and the result is an increase in the cost of capital.

Diamond et al. (1991) suggest that disclosure of more information is likely to reduce information asymmetry, thereby leading to increased liquidity in the firms' shares and corresponding decrease in the cost of capital. Lang and Lundholm (1996) also find reduction in estimation risk and information asymmetry to be a potential benefit of voluntary disclosure. Hope (2003) documents that financial disclosure quality is positively related to analyst forecast accuracy. In addition, prior literature has empirically proven investor uncertainty to be positively correlated with future stock return volatility. If disclosure acts to reduce uncertainty, one would also expect for disclosure to decrease future stock return volatility.

Nonetheless, the empirical link between disclosure and return volatility is not straightforward. Verrecchia (1983) suggests that traders are unable to interpret any withheld information as definitively bad news. Therefore, traders with rational expectations will discount firm value to the extent that the manager's opportunity cost of not disclosing will become too high, and the manager is better served to disclose what he knows. In this way, managers weigh the perceived costs and benefits of disclosing sustainability information and will choose to disclose only if the perceived benefits outweigh the costs. Milgrom (1981) argues that information asymmetries exist between shareholders and firm managers. He indicates that

[11] Investors may feel more confident when firms have high levels of disclosure due to the potential reduction in information asymmetry; there is increased precision associated with the transaction price.

buyers expect any withheld product information to be unfavorable to his product. This school of thought argues that the divide creates a demand for voluntary disclosure and thereby incentivizes firms to provide information and reduce information asymmetries.

Financial economic theory maintains that there exists a limited role for the idiosyncratic manger-specific influence. To this end, Bamber et al. (2010) empirically examine the role of individual managers in affecting voluntary financial disclosure decisions. The researchers take it upon themselves to temporally track a sample of managers across firms. Overall, the authors find that the top managers do, indeed, influence the firms' voluntary disclosures and do so with unique styles based on their respective backgrounds.

Ceasing to disclose, in general, is found to be associated with an increase in analyst forecast dispersion and decrease in forecast accuracy. Chen, Matsumoto, and Rajgopal (2011) obtain a sample of firms, of which some stop providing earnings guidance (deemed "stoppers). Of these firms, the firm either (i) publicly announces the decision (i.e. "announcers") or (ii) refrains from announcing (i.e. "quiet stoppers"). When compared to the firms that continue providing guidance, the firms that stop guiding tend to have poorer prior performance, more uncertain operating environments, and also fewer informed investors. The announcers experience a negative three-day return around their announcement of non-disclosure. The authors also observe that announcers tend to publicly commit to non-disclosure due to either not expecting good news reports in the future or presence of long-term investors that attenuate the incentive to guide.

In contrast to empirical evidence in the literature linking disclosure to information environment benefits, an alternative stream of research has also provided evidence that guidance may *increase* stock return volatility (Rogers et al, 2009). However, it is unclear as to whether volatility plays a role in the issuance of the guidance. The ensuing research weighs in on this

important debate regarding the link between disclosure and stock return volatility.

Extant empirical research has not thoroughly addressed the breadth of voluntary ESG disclosure implications. There is a lack of cohesiveness in the nonfinancial disclosure literature as to warrant additional research. Based on recent work by Khan (2015) that illustrates the materiality of sustainability issues within given industries, one would expect knowledgeable investors to incorporate environmental considerations into their investment decisions.

Information asymmetry can be expected to exist in the market when market participants have a divergence of opinion. However, the existence of this information asymmetry generally has adverse implications. Divergence of opinion is often marked by an increase in investor trading. Following previous information asymmetry literature, I employ stock return variance as a proxy for information asymmetry as stock return volatility generally increases when market participants are unclear as to the "true" value of the stock (Comprix et al., 2011). I also employ bid-ask spread and annual forecast dispersion as suitable proxies for information asymmetry, based on prior accounting literature (Comprix et al., 2011, Kim et al., 2014). The sample selection portion of the paper provides further details regarding justification and calculation of the chosen proxies[12].

2.6. Voluntary Disclosure and Cost of Capital

It is also possible that the environmental disclosures act to reduce information asymmetry regarding environmental performance. As a result of this relation between disclosure and information asymmetry, one would expect an associated decrease in the firm's cost of capital. As noted in Akerlof (1970), there may exist an incentive for managers to voluntarily disclose in order to reap the ensuing benefits. One such benefit is a reduction in the cost of capital. The

[12] Detailed proxy definitions are also located in Appendix A.

assumption is that firms with higher disclosures are imparting additional information to the market. By doing so, they are also reducing information risk. Since investors will demand benefit for incurring the costs associated with information risk, the firms with lower information risk also are likely to have a lower cost of capital. Bertomeu, J., Beyer, A., & Dye, R. A. (2011) examine firm-level capital structure, voluntary disclosure policy, and cost of capital by presenting a model of financing to describe their joint determination. The authors indicate that both firm-level structure of securities and overall disclosure policy are associated with certain informational advantages on the part of informed traders. The firm's cost of capital and the investors' trading losses are also affected. Overall, the model hypothesizes a negative relation between firm's cost of capital and the amount of firm-level information disclosure[13].

The aforementioned branch of the accounting literature addresses the welfare impacts of disclosure policy. Cost of capital is often employed as a proxy for investor welfare, as investor welfare is not straightforward to measure. Research by Gao urges researchers to consider the empirical limitations of assuming that disclosure quality leads to improved investor welfare through the reduction in cost of capital (Gao, 2010). The research demonstrates that cost of capital can be increasing in disclosure quality if new investment is elastic. There are certain contexts under which disclosure quality may actually act to reduce investor welfare. It is, therefore, important to note that cost of capital is not always an adequate proxy to address investor welfare when examining overall economic implications of disclosure quality.

A related strand of environmental accounting literature focuses on the negative relation between environmental performance and the cost of capital (Silva-Gao et al., 2008). Namely, Dhaliwal et al. (2011) and Plumlee et al. (2010) provide evidence to support a negative

[13] Nonetheless, lengthier (or more detailed) disclosure would not necessarily be expected to reduce the cost of capital as well.

association between both environmental disclosure and the cost of capital as well as CSR disclosure and the cost of capital. As such, research has also established a link between CSR disclosures and the cost of capital. The justification is that the disclosed environmental performance captures an aspect of firm risk; this additional information is taken into account when establishing the cost of capital. Richardson and Welker (2001) attribute this relation to either investor preference for socially responsible and ethical investing or reduced information asymmetry and/or estimation risk. There are also costs associated with information asymmetry. Prior empirical evidence has also established that there are costs associated with information asymmetry. For instance, the research of Easley et al. (2002) evidences the link between less asymmetry and lower cost of equity capital (Easley et al., 2002). This is another avenue through which environmental disclosures act to increase firm value. The research at hand extends this strand of literature by focusing on environmental disclosure and the ensuing implications for information asymmetry in the market.

2.7. Nonfinancial Disclosure and Information Asymmetry

Building on the work of Clarkson et al. (2004), Hope (2003), Dhaliwal (2012) and Plumlee (2010), the subsequent analysis initially concentrates on isolating the link between environmental disclosure and information asymmetry, as proxied by stock return variance, bid-ask spread and annual forecast dispersion. One might anticipate that an increase in reporting behavior is associated with increased transparency surrounding ESG. Recall that Dhaliwal et al. (2011) claim that firms with better CSR performance are more likely to signal their long-term focus to the market through use of voluntary disclosure. Environmental disclosure may also foster a better understanding of how sustainability is incorporated into firm value, thereby leading to a negative relation between environmental disclosure and proxies for information asymmetry.

Based on the complementary premise that voluntary environmental reporting will reduce information asymmetries about corporate performance, one would expect a negative association between disclosure and proxies for information asymmetry, in general.

However, the existence of complementarity between the forms of disclosure is by no means definitive. As previously noted, the existing literature and neoclassical economic theory have questioned whether investment in ESG fulfills the objective of shareholder wealth maximization. If market participants view the disclosed ESG activities as indicative of a failure to maximize shareholder wealth, the direction of the association becomes more ambiguous; if investors differ in their assessment of such disclosures, this may even prompt an increase in investor uncertainty. Neoclassical economic theory generally maintains that CSR engagement would lead to increased costs that would unnecessarily position the firm at a disadvantage vis-à-vis its competitors (Friedman, 1970). Therefore, the first hypothesis of this chapter will explore the relation between sustainability disclosure and volatility, in attempt to disentangle whether disclosure of nonfinancial information is associated with reduced levels of investor uncertainty.

I extend prior literature by focusing specifically on *nonfinancial* information predicting that nonfinancial disclosures (proxied by the issuance of a standalone sustainability report) is negatively associated with proxies for information asymmetry. I directly test the capital market implications of voluntary nonfinancial disclosure by empirically investigating whether information asymmetry decreases when nonfinancial disclosure is present. I would expect that for firms that are more transparent in their sustainability disclosure, current stock returns would incorporate more information about future earnings.

I state this prediction formally and in the alternative form, as the third hypothesis below:

H3: Ceteris paribus, issuance of a standalone sustainability report is negatively associated with information asymmetry.

To test this prediction, I use publication of stand-alone sustainability reports as a proxy for the amount and availability of sustainability-related nonfinancial information to investors. A significantly negative relation between nonfinancial disclosure and information asymmetry will support the stated hypothesis. However, a failure to reject the null hypothesis of no relation would not necessarily indicate that sustainability reports do not have an informative role to play. One could also envision the case that a significant relation only exists within alternative information environments (for instance, in the presence of internal control weaknesses or organizational complexity). The informational role of the nonfinancial disclosure could very well be to improve an uncertain information environment. After testing the prediction that sustainability disclosure reduces information asymmetry in the market, the next research question will address contexts under which the sustainability information is most useful.

2.7.1. Alternative Information Environments: Internal Control Weaknesses

This research setting allows for evaluation of the capital market response to environmental disclosures in the presence of internal control weaknesses in an attempt to capture a degree of financial opacity. I examine the impact of auditor-attested material internal control weaknesses (ICW) over financial reporting under the requirements of Section 404 of the Sarbanes-Oxley Act (SOX).

Empirical research has indicated that investors require compensation for uncertainty regarding a firm's reporting quality (Easley et al., 2004). There is a documented relation between internal control weaknesses and accounting risk. Effective internal controls are expected to reduce accounting risk. As a result of this reduction in risk, *effective* internal

controls are associated with increased accounting quality. The PCAOB defines a material weakness in internal controls as deficiencies that "result in more than a remote likelihood that a material misstatement of the annual or interim financial statements will not be prevented or detected." The presence of internal control weaknesses has been shown to significantly increase the overall likelihood of financial statement misstatements. ICWs are associated with increased likelihood of accounting error, increased scope for earnings management (Ge et al., 2005) and increased likelihood for financial fraud[14]. The increased level of accounting risk due to the presence of ICWs could also adversely affect accounting quality. As such, I use internal control weaknesses as a means to capture financial opacity.

The internal control weakness disclosures specifically address the integrity of each firm's individual financial reporting process. Issuance of weakness disclosures may result in belief revision; the investors perceive weakness disclosures to be indicative of diminished reporting and, therefore, greater risk. Beneish (2008) finds that material weakness disclosures are more informative for smaller firms that are likely to have higher pre-disclosure information uncertainty. Weakness disclosures are associated with adverse stock market responses as investors revise their risk assessments upwards. Ex ante, firms disclosing internal control weaknesses are smaller, younger, riskier and financially weaker (Ashbaugh-Skaife, Collins, and Kinney 2007; Doyle et al. 2007).

For firms with a higher level of internal control weaknesses, I would expect the nonfinancial disclosure to further supplement the financial information and thereby increase the negative relation between sustainability disclosure and information asymmetry. That is, I would expect for the nonfinancial disclosure to be more informative in the presence of this financial

[14] Ogneva et al. (2007) note that the improper segregation of duties may create additional opportunities for employee fraud to occur within the organization.

opacity. I state this prediction formally and in the alternative form, as the fourth hypothesis below:

H4: Ceteris paribus, the negative relation between sustainability report disclosure and information asymmetry is stronger for firms with a higher level of internal control weaknesses.

A rejection of the null hypothesis may indicate that, in spite of the increased internal control weaknesses (and associated opacity), investors may take the disclosure as a signal of earnings quality. For instance, Kim et al. (2012) document a positive relation between CSR and earnings quality and demonstrates that the firms with higher CSR are less likely to manage earnings, manipulate real operating activities, and be the subject of SEC investigations. When viewed in this light, investors may use the nonfinancial information to supplement the financial information; the nonfinancial disclosures improve the information environment. Disclosure of ESG information may be viewed as a signal of the company's degree of transparency. Investors may also perceive less risk in investing in more transparent companies due to less uncertainty regarding the firms' ability to deliver on expected financial performance (Eccles, 2011).

On the other hand, a failure to reject the null hypothesis may indicate that, in the presence of more internal control weaknesses, investors are less inclined to reference the nonfinancial information. If so, the nonfinancial disclosures do not act to supplement the existing financial information. One reason for this result could be that investors are aware of the questionable incentive structure behind voluntary disclosures. The presence of financial opacity might lead investors to infer that the nonfinancial information could be "cheap talk". Investors may have a tendency to use simplifying heuristics and could also deem the firm to be less credible and more likely to be obfuscating or legitimizing their role in a contract with society rather than providing

information that is valuable to their investment decision. In this case, the presence of a nonfinancial disclosure would not improve the information environment and thus would not play a role in the reduction of information asymmetry in the presence of uncertainty.

2.7.2. Alternative Information Environments: Organizational Complexity

Disclosures in the form of sustainability reports represent just one part of a firms' overall information environment. In order to further explore the relation, I identify organizational complexity as an additional information environment. I would anticipate the relative importance of the nonfinancial disclosure information in explaining information asymmetry to vary based on organizational complexity. I would expect for the importance of disclosure in the communication process to be greater for firms with greater organizational complexity. I state this prediction formally and in the alternative form, as the fifth hypothesis below:

H5: Ceteris paribus, the negative relation between sustainability report disclosure and information asymmetry is stronger for firms with a higher level of organizational complexity.

2.7.1. Alternative Information Environments: Financial Statement Complexity

H6: Ceteris paribus, the negative relation between sustainability report disclosure and information asymmetry is stronger for firms with a higher level of financial statement complexity.

I would expect that the sustainability report disclosure will impart relatively more information for firms with greater financial statement complexity, thereby resulting in a stronger relation between disclosure and information asymmetry.

3. Sample Selection and Data Collection

3.1. Sample

I gather a sample of United States firms that have issued sustainability reports, based on the Global Reporting Initiative List. In total, I identify 1,457 standalone sustainability reports, covering the period from 2014-2018. I merge the data via SAS to obtain a final sample size of 926 observations, including both disclosing and non-disclosing firms. The primary observation unit is the firm. Two of the dependent variables used to proxy for information asymmetry are gathered via Center for Research in Security Prices (CRSP); these include bid-ask spread and stock return variance. The third proxy for information asymmetry (annual dispersion) is obtained from the Institutional Brokers' Estimate System (I/B/E/S). In addition, information regarding the number of forecasts is obtained from I/B/E/S. I obtain data from Compustat North America for the following variables: total assets, total market value, total liabilities, operating income after depreciation, total long-term debt in current liabilities, total common/ordinary equity, total property, plant and equipment, and total revenue. Segment and sales data is also obtained via Compustat in order to calculate the revenue-based Hirfindahl index for industry concentration, and then transformed to represent a measure organizational complexity. The MSCI KLD database is used for data in respect to environmental strengths and weaknesses. Restatement and ICW data are collected via Audit Analytics. Institutional Ownership data is obtained via the Thomson Reuters database. All firms with adequate data are included within the analysis.

3.2. Main Variables

The main variables of interest are proxies for information asymmetry as the dependent variables, nonfinancial disclosure, and internal control weaknesses and organizational complexity as interaction terms. I describe the measurement of each of these variables below.

3.2.1. Information Asymmetry Proxies

As information asymmetry is not directly observable, I employ three market-based proxies based on prior literature (Comprix et al., 2011, Kim et al., 2014). I include bid-ask spread, stock return variance, and annual forecast dispersion. Leuz and Verrecchia (2000) suggest bid-ask spread as a suitable proxy for information asymmetry. This variable is measured as the average bid-ask spread. The average bid-ask spread is calculated using the absolute spread scaled by the average of bid and ask (bid-ask). Data are obtained from the CRSP database.

3.2.2. Sustainability Disclosure Data

The Global Reporting Initiative has provided sustainability disclosure data, upon request. The GRI Report List includes detailed sustainability reporting information for firms worldwide. The research incorporates the U.S. firms that produce a standalone sustainability report[15]. In the later robustness check, firms will be matched with non-disclosing firms from the Compustat database through use of the propensity score matching method.

[15] As a proxy for nonfinancial disclosure, the sustainability disclosures are treated as homogeneous end goods. This is likely an oversimplification of the true disclosure product, albeit adequate in the context of the current research question. Although delving into the components of this disclosure would likely prove to be quite interesting, it is beyond the scope of the current analysis.

3.2.3. Information Environment Proxies

Material Internal Control Weaknesses (ICW)

The first empirical test employs the presence of auditor-attested material internal control weaknesses (ICW) over financial reporting under the requirements of Section 404 of the Sarbanes-Oxley Act (SOX). For the purposes of this research, I am interested in evaluating the firm-level characteristics, and not the event-specific characteristics. As a result, I construct the measure through use of long-window regressions analysis instead of event studies (Kim et al., 2014). Section 404 of SOX requires firms to maintain adequate internal controls over financial reporting and also to provide auditor-attested evaluations of their effectiveness. The internal controls are intended to offer reasonable assurances as to the reliability of financial reporting and the preparation of financial statements for external purposes in accordance with generally accepted accounting principles (Kim, 2014). Feng et al. (2009) find that the quality of internal controls also affects the quality of internal financial reports and management earnings forecast quality. Ashbaugh-Skaife et al. (2008) indicate that ICW is associated with low accrual quality. Overall, prior research has shown earnings guidance for ICW firms to be less accurate. I obtain internal control weaknesses data from the Audit Analytics database via WRDS.

Organizational Complexity

To further explore the intricacies of the relation between nonfinancial disclosure and information asymmetry, I also employ a measure of organizational complexity. I compute within-firm industry concentration through use of a revenue-based Hirfindahl-index (Chen, 2002). I calculate the industry concentration as the sum of the squares of each segment's sales as a percentage of the total firm sales. I transform the measure to represent organizational

complexity by multiplying industry concentration by negative one and then adding one. By doing so, an increase in the variable *ORGCOMPLEXITY* indicates that the organizational complexity is increasing.

Financial Statement Readability

I employ several measures of readability based on the prior literature. Several common readability measures are employed across accounting, finance, and computational linguistic disciplines. I use the Natural Language Tool Kit (NLTK) in Python to calculate each of the following readability measures: (1) Flesch Reading Ease formula, (2) Flesch-Kincaid Grade Level, (3) SMOG Index (4) Automated Readability Index, (5) The Coleman-Liau Index, (6) Linsear Write Formula, (7) Dale-Chall Readability Score, and (8) Fog Index. As I expect the readability measures to be highly correlated, I estimate a consensus readability score that incorporates measures (1) through (8) to ensure a robust readability score.

The Flesch Reading Ease and Flesch-Kincaid Grade Level formulas use the same components as Fog (the percentage complex words and average number of words per sentence). The Flesch Reading Ease formula and Flesch-Kincaid Grade Level differ in that the measures employ an explicit count of complex word syllables rather than the binary classification used for the Fog index. Both the Fog Index and the Flesch-Kincaid Grade Level produce numeric estimates of grade level. For instance, a Flesch-Kincaid Grade Level of 9.3 indicates that a ninth grader would be able to read the document with ease. The Flesch Reading Ease formula differs in that it returns the Flesch Reading Ease Score based on a scale of 0-100 where score ranges represent levels of difficulty. For instance, a Flesch Reading Ease Score in the range of 90-100 indicates a "Very Easy" text whereas a score of 0-29 is deemed to be "Very Confusing"[16].

[16] A more thorough explanation of the Flesch Reading Ease Score can be found in the Appendix section where I enumerate each score range accompanied by the associated interpretation.

The SMOG index measure is also comparable in the sense that the interpretation is that of grade levels[17]. The following SMOG formula yields the grade level at which the text can be interpreted with ease.

$$Smog = 1.0430 \sqrt{number\ of\ polysyllables\ x \frac{30}{number\ of\ sentences}} + 3.1291$$

The Automated Readability Index (ARI), Linsear Write equations, and Dale-Chall Readability Score also yield grade levels. The Dale-Chall Readability Score differs in that the Dale-Chall Readability Score requires a lookup table of the most commonly used 3000 English words.

The SEC provides clear guidance in their recommendation that managers employ certain Plain English attributes. Managers can do so by avoiding writing constructs such as passive voice, weak or hidden verbs, legal and financial jargon, numerous defined terms, abstract words, unnecessary details, lengthy sentences, and unreadable design and layout in financial disclosure (SEC, 1998b; Bonsall, Leone, Miller, and Rennekamp, 2017).

In addition to the aforementioned readability measures obtained from Python textstat[18], I also incorporate the Bog Index as a proxy for financial disclosure complexity. Bonsall, Leone, Miller, and Rennekamp (2017) introduce the Bog Index, a new multi-faceted measure of readability based on Plain English attributes[19]. With their new Bog Index measure, the authors intend to capture plain English attributes of the disclosures. Then, the researchers proceed to define their own readability measurement.

---

[17] A thorough explanation of the original SMOG calculation process as detailed by McLaughlin (1969) is included in the Appendix.

[18] Textstat is a Python package used to calculate statistics from text to determine readability, complexity, and grade level of a particular corpus.

[19] In conjunction with the proposal of this new measure, the authors ensure the validity of the measurement by running a series of experiments, and also subjecting the measure to empirical verification as compared to the alternate readability measures contained in the literature to date.

In the empirical analysis at hand, I employ the Bog Index as an additional measure of financial disclosure linguistic complexity. The Bog Index is constructed via a software program called StyleWriter. The program identifies attributes that are contained within the SEC Plain English Handbook. For instance, the software captures qualitative attributes such as sentence length, passive voice, weak verbs, overused words, complex words, and jargon. The software does not immediately conclude that every multi-syllabic word is complex, a criticism of prior measurements. Instead, the word complexity is measured based on a lengthy word list. The Bog Index provides a summary of the writing attributes that tend to bog readers down. The index is computed as the sum of Sentence Bog, Word Bog and Pep as follows:

Bog Index=Sentence Bog + Word Bog – Pep

A higher Bog Index indicates diminished readability. Sentence Bog includes readability issues stemming from sentence length. Longer sentences imply lower readability. StyleWriter identifies average sentence length across the entire document, and the average sentence length is squared and scaled by 35 words per sentence. Word Bog is comprised of plain English style problems as well as word difficulty. The plain English style problems and word difficulty are summed, multiplied by 250 and then divided by the number of words. Word difficulty is computed based on a proprietary list of over 200,000 words. Based on this proprietary list, certain abstract words receive higher scores relative to less abstract words.

The last component of the Bog Index is Pep. Pep includes writing attributes that aid in understanding of texts. It is a sum of names and interesting words that tend to make writing more interesting. Pep is calculated to be the sum of these components that aid in interpretation multiplied by 25 and scaled by the number of words in the document plus sentence variety

(standard deviation of sentence length multiplied by ten and scaled by the average sentence length).

3.3. Sustainability Performance Data

MSCI KLD sustainability data is collected to control for firm environmental performance. The MSCI KLD sustainability data contains a large number of U.S. firms over time. The KLD historical ratings data set is a dichotomous system that incorporates both strengths and weaknesses of the firms. The data set includes seven issue areas as follows: Community, Corporate Governance, Diversity, Employee Relations, Product, Environment and Human Rights. The binary nature of the data includes a "1" to indicate presence of the criterion and "0" to indicate its absence. Each of the broad categories includes several subsets of related data items. I define the environmental *STRENGTHS* variable as the average number of proactive ratings for a firm identified in KLD. Conversely, I define environmental *CONCERNS* as the average number of damaging ratings for a firm identified in KLD.

3.4. Other Independent Variables

Based on the collected data, I calculate *BM* as the firm's market-to-book ratio, *MKVALT* as the log of market value of a firm's equity. *ROI* represents the firm's return on invested capital, defined as operating profit scaled by total assets. This variable is included as a control in the disclosure choice model because firms may increase disclosure when they are performing well. *INSTOWN* is defined as the percentage of total shares outstanding held by institutional investors. I define the variable *size* as the log of the firm's total assets at the end of the fiscal year. Each firm's leverage is also included as *LEVERAGE* and measured as (DLTT + DLC)/(DLTT +DLC+CEQ) in Compustat.

3.5. Heteroskedasticity

I employ the Breusch-Pagan test and White's test for heteroskedasticity and fail to reject the null hypothesis of constant variance. That is, heteroskedasticity may be present in the dataset. As a result, the heteroskedasticity-robust standard errors are reported. The heteroskedasticity-robust standard errors allow for computation of t-statistics that are asymptotically t-distributed whether or not heteroskedasticity is present. The two sets of standard errors are only minimally different. However, due to the likely presence of heteroskedasticity, inference through use of the robust standard errors is more appropriate. Note that important conclusions are not overturned through use of the robust standard errors.

3.6. Descriptive Statistics

TABLE 7 presents summary statistics of key variables for all observations. I winsorize all continuous firm-level variables at the 95% level in order to ensure that the prevalence of outliers within the dataset does not drive the results.

TABLE 7

Chapter II Summary Statistics

Variable	Mean	25th Percentile	Median	75th Percentile	Standard Deviation
BIDASKSPREAD	0.0129	0.0102	0.0112	0.0139	0.0039
RETVAR	45.2998	7.5459	21.8914	67.3718	50.7759
DISPERSION	0.4518	0.2489	0.4054	0.6232	0.2668
NFDISC	0.7797	1.0000	1.0000	1.0000	0.4147
LOGASSET	10.1273	9.2860	10.1045	10.9914	1.1487
BM	4.0799	2.0985	3.1410	4.6987	2.8023
LEVERAGE	0.4877	0.3448	0.4730	0.6128	0.1709
ICW	0.0097	0.0000	0.0000	0.0000	0.0982
STRENGTHS	1.0539	0.6667	1.0385	1.4167	0.5038
CONCERNS	0.7184	0.0000	0.4444	1.3000	0.7695
OWNERSHIP	0.0013	0.0006	0.0010	0.0018	0.0009
ROI	0.0930	0.0527	0.0839	0.1258	0.0461
FOLLOWING	4.2171	3.0000	4.0000	5.0000	1.6873
ORGCOMPLEXITY	0.7270	0.6433	0.7685	0.8227	0.1159
FINREADABILITY	-29.1803	-28.0000	-23.0000	-21.0000	15.0017
BOGINDEX	88.1350	83.0000	87.0000	93.0000	5.8007
N	926				

This table presents the summary statistics for the variables used in Chapter II.

TABLE 8 represents the correlation matrix for the full sample. Consistent with H3, at least one of the proxies for information asymmetry is negatively related with the disclosure. It is reassuring to note that all of the proxies for information asymmetry are positively correlated with one another, albeit not as strongly as one might initially anticipate. A strong correlation exists between firms with both environmental strengths and environmental concerns and the binary disclosure variable. That is, it appears that firms with no information to disclose are disclosing what they know. If the firms were withholding the information, we might anticipate a negative correlation. Firm size and disclosure is positively correlated. One would anticipate the existence of this correlation, as larger firms may be more likely to voluntarily disclose, in general. Lang and Lundholm (1993) document that disclosure levels are positively correlated with firm size. It is also reassuring to note that average analyst following (following) and institutional ownership (Ownership) are both positively correlated with disclosure as both of the measures proxy for overall information environment.

TABLE 8

Correlation Matrix

	BIDASKSPREAD	RETVAR	DISPERSION	NFDISC	SIZE	BM	LEVERAGE	ICW	STRENGTHS	CONCERNS	OWNERSHIP	ROI	FOLLOWING	ORGCOMPLEXITY	FINREADABILITY	BOGINDEX
BIDASKSPREAD	1.00															
RETVAR	0.59***	1.00														
DISPERSION	0.00	0.30***	1.00													
NFDISC	-0.09***	-0.10***	0.01	1.00												
SIZE	-0.08**	0.10***	0.37***	0.22***	1.00											
BM	0.19***	0.16***	-0.10***	-0.01	-0.14***	1.00										
LEVERAGE	-0.13***	-0.07**	-0.01	0.11***	0.18***	0.36***	1.00									
ICW	0.07**	-0.07**	-0.06*	-0.03	-0.09***	-0.02	-0.01	1.00								
STRENGTHS	-0.04	-0.03	0.02	0.08**	0.33***	0.29***	0.22***	-0.04	1.00							
CONCERNS	-0.16***	-0.01	0.10***	0.17***	0.38***	-0.09***	0.22***	-0.08***	0.27***	1.00						
OWNERSHIP	0.01	-0.20***	-0.27***	-0.09***	-0.59***	-0.25***	-0.18***	0.11***	-0.60***	-0.27***	1.00					
ROI	0.12***	0.12***	-0.15***	-0.17***	-0.26***	0.54***	-0.14***	-0.02	0.27***	-0.29***	-0.28***	1.00				
FOLLOWING	0.07**	0.08**	-0.08**	0.04	0.10***	0.17***	-0.06*	0.02	0.12***	-0.13***	-0.18***	0.20***	1.00			
ORGCOMPLEXITY	-0.17***	-0.22***	-0.07**	0.14***	0.12***	0.15***	0.06*	0.03	0.36***	0.21***	-0.12***	0.07**	-0.00	1.00		
FIN-READABILITY	0.19***	0.13***	0.06*	0.01	0.03	0.19***	-0.04	0.02	0.12***	-0.16***	-0.05	0.13***	0.11***	0.18***	1.00	
BOG	-0.04	-0.10***	-0.08**	-0.00	0.13***	-0.25***	0.05	0.01	-0.13***	0.05	0.06*	-0.26***	-0.06*	0.03	-0.02	1.00
N	926															

This table presents Pearson correlation coefficients for the variables of interest in Chapter II.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

4. Empirical Methods

4.1. Empirical Model: H3

Equation **(v)**, below, details the empirical model used to test the third hypothesis.

$$\text{Information Asymmetry} = \beta_0 + \beta_1 \text{DISCLOSER} + \theta \text{Controls} \quad \text{(v)}$$

The key parameter of interest in the regression model is the coefficient pertaining to nonfinancial disclosure. The model is completed by a set of additional controls. I control for various factors that are likely to confound the relation between information asymmetry and nonfinancial disclosure. Specifically, I control for firm size, book-to-market ratio, average KLD strengths, average KLD concerns, leverage, annual institutional ownership and average analyst following. For additional detail, all variables are defined in Appendix A.

4.2. Empirical Model: H4

$$\text{Information asymmetry} = \beta_0 + \beta_1 \text{Discloser} + \beta_2 \text{Discloser} * \text{ICW} + \beta_3 \text{ICW} + \theta \text{Controls} + \epsilon \quad \text{(vi)}$$

The key parameter of interest in the regression model is the coefficient pertaining to the nonfinancial disclosure term interacted with internal control weaknesses. Controls include firm size, book-to-market ratio, average KLD strengths, average KLD concerns, leverage, annual institutional ownership and average analyst following.

4.3. Empirical Model: H5

$$\begin{aligned}\text{Information asymmetry} &= \beta_0 + \beta_1 \text{DISCLOSER} + \beta_2 \text{DISCLOSER} * \text{ORGCOMPLEXITY} \\ &+ \beta_3 \text{ORGCOMPLEXITY} + \theta \text{Controls} + \epsilon\end{aligned} \quad \text{(vii)}$$

The key parameter of interest in the regression model is the coefficient pertaining to the interacted nonfinancial disclosure term with organizational complexity. The vector of control

variables remains the same and includes firm size, book-to-market ratio, average KLD strengths, average KLD concerns, leverage, annual institutional ownership and average analyst following.

4.4. Empirical Model: H6

$$\text{Information asymmetry} = \beta_0 + \beta_1 \text{DISCLOSER} + \beta_2 \text{DISCLOSER} * \text{FINREADABILITY} + \beta_3 \text{FINREADABILITY} + \theta \text{Controls} + \epsilon \quad \text{(viii)}$$

$$\text{Information asymmetry} = \beta_0 + \beta_1 \text{DISCLOSER} + \beta_2 \text{DISCLOSER} * \text{BOG} + \beta_3 \text{BOG} + \theta \text{Controls} + \epsilon \quad \text{(ix)}$$

The key variables of interest in equations **(viii)**and **(ix)** are FINREADABILITY and BOG, respectively. The vector of control variables remains unaltered.

5. Results

5.1. H3 Results

TABLE 9 presents the empirical results of the tests of the third hypothesis for all three of the information asymmetry proxies. For all three information asymmetry proxies, the main variable of interest, *DISCLOSER*, has a significantly negative coefficient at the 10%, 1%, and 5% significance levels for the BIDASKSPREAD, RETVAR, and DISPERSION proxies, respectively. The findings signify that sustainability disclosure is significantly negatively correlated with proxies for information asymmetry in the baseline model. In all three models, I reject the null hypothesis of no relation between nonfinancial disclosure and information asymmetry in favor of the alternative hypothesis that there is a negative relation between disclosure and information asymmetry.

TABLE 9

Chapter II Base Model

Independent Variable	Dependent Variables		
	BIDASKSPREAD	RETVAR	DISPERSION
DISCLOSER	-0.000621*	-14.08***	-0.0436**
	(0.000350)	(4.425)	(0.0200)
SIZE	0.000103	4.153**	0.0662***
	(0.000141)	(1.915)	(0.0111)
BM	0.000501***	4.941***	0.00622
	(0.0000599)	(0.765)	(0.00393)
STRENGTHS	-0.000332	-26.11***	-0.0869***
	(0.000360)	(4.477)	(0.0188)
CONCERNS	-0.000535***	1.478	-0.0226*
	(0.000188)	(2.493)	(0.0123)
LEVERAGE	-0.00571***	-53.99***	-0.176***
	(0.000753)	(10.21)	(0.0570)
OWNERSHIP	-0.102	-16298.7***	-89.60***
	(0.205)	(2503.2)	(15.27)
FOLLOWING	0.00000502	0.200	-0.0215***
	(0.0000719)	(0.971)	(0.00526)
ROI	-0.0115***	-57.37	-0.988***
	(0.00424)	(56.49)	(0.276)
Intercept	0.0151***	73.12***	0.286*
	(0.00185)	(24.85)	(0.147)
Observations	926	926	926
R^2	0.107	0.138	0.206

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

5.1. H4 ICW Results

TABLE 10 presents the regression results for testing H4. The coefficient on the interaction term *Discloser*ICW* is observed to be significant for two of the three information asymmetry proxies. The hypothesis suggests that, in the presence of ICWs, nonfinancial disclosure and annual dispersion are *more negatively correlated.* The coefficient on the interaction term *DISCLOSER*ICW* is observed to be significantly negative in the model with the proxy DISPERSION as the dependent variable. I observe mixed results as the interaction term is observed to be significantly positive in the model with the proxy BIDASKSPREAD as the dependent variable. It is also worth noting that the interaction term is not observed to be statistically significant when the RETVAR proxy is utilized.

TABLE 10

Interaction with Internal Control Weaknesses

Independent Variable	Dependent Variables		
	BIDASKSPREAD	RETVAR	DISPERSION
DISCLOSER	-0.000685*	53.62	-0.0618*
	(0.000354)	(69.10)	(0.0363)
ICW	-0.000911	63.27	0.133
	(0.000554)	(177.5)	(0.0964)
DISCLOSER*ICW	0.00515**	-39.50	-0.337***
	(0.00213)	(196.0)	(0.109)
ASSET	0.000119	246.4	0.0734***
	(0.000141)	(154.3)	(0.0193)
BM	0.000502***	28.53***	0.0105*
	(0.0000598)	(8.656)	(0.00548)
STRENGTHS	-0.000369	-9.363	-0.154***
	(0.000357)	(64.28)	(0.0319)
CONCERNS	-0.000517***	14.76	-0.0175
	(0.000188)	(22.88)	(0.0203)
LEVERAGE	-0.00569***	-711.9	-0.223**
	(0.000755)	(447.5)	(0.0994)
OWNERSHIP	-0.147	137841.2	-123.2***
	(0.203)	(117979.1)	(30.76)
FOLLOWING	-0.00000454	37.04	-0.0416***
	(0.0000724)	(29.01)	(0.00961)
ROI	-0.0113***	4070.8	-2.097***
	(0.00423)	(2936.4)	(0.517)
Constant	0.0150***	-2879.3	0.582**
	(0.00185)	(1963.5)	(0.272)
Observations	926	926	926
R^2	0.114	0.031	0.161

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

5.1. H5 Organizational Complexity Results

TABLE 11 presents the empirical results for H5, with the interaction of *DISCLOSER and ORGCOMPLEXITY*. The correlation is observed to be positive and statistically significant for all three of the information asymmetry proxies. For organizationally complex firms, the relation between discloser and information asymmetry is more positively correlated across all information asymmetry proxies. The results are contrary to the original hypothesis that the relative importance of the nonfinancial disclosure information would increase in an organizationally complex environment. Despite the unanticipated positive correlation, the results lend themselves to an alternative interpretation. I interpret the result as indicative of a decrease in the relative importance of nonfinancial disclosure in the reduction of investor uncertainty within organizationally complex environments.

TABLE 11

Interaction with Organizational Complexity

Independent Variable	Dependent Variables		
	BIDASKSPREAD	RETVAR	DISPERSION
DISCLOSER	-0.00952***	-128.2***	-0.515***
	(0.00210)	(26.50)	(0.121)
ORGCOMPLEXITY	-0.0164***	-225.1***	-0.655***
	(0.00246)	(31.40)	(0.140)
DISCLOSER*	0.0130***	167.1***	0.679***
ORGCOMPLEXITY	(0.00277)	(35.04)	(0.164)
SIZE	0.000154	4.965***	0.0666***
	(0.000136)	(1.866)	(0.0110)
BM	0.000513***	5.183***	0.00563
	(0.0000610)	(0.765)	(0.00386)
STRENGTHS	0.000350	-15.99***	-0.0701***
	(0.000358)	(4.448)	(0.0206)
CONCERNS	-0.000332*	4.420*	-0.0165
	(0.000178)	(2.389)	(0.0123)
LEVERAGE	-0.00565***	-54.10***	-0.160***
	(0.000787)	(10.31)	(0.0553)
OWNERSHIP	0.251	-11252.7***	-78.02***
	(0.209)	(2509.7)	(15.38)
FOLLOWING	-0.0000303	-0.289	-0.0229***
	(0.0000697)	(0.919)	(0.00512)
ROI	-0.00671	5.887	-0.752***
	(0.00426)	(56.79)	(0.281)
Intercept	0.0242***	197.4***	0.677***
	(0.00234)	(31.90)	(0.167)
Observations	926	926	926
R^2	0.162	0.204	0.223

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

5.1. H6 Readability Results

TABLE 12 presents the empirical results for H6, with the interaction of DISCLOSER and FIN-READABILITY. The correlation is observed to be negative for two of the three information asymmetry proxies and significant at the 5% level for two of the proxies. In the presence of financial complexity, the relation between disclosure and information asymmetry is more negatively correlated across two of the three information asymmetry proxies. The results lend themselves to the interpretation that nonfinancial disclosure issuance becomes more informative in an environment marked by financial statement complexity.

TABLE 12

Interaction with Readability

Independent Variable	Dependent Variables		
	BIDASKSPREAD	RETVAR	DISPERSION
DISCLOSER	0.00267*	95.59	0.0359
	(0.00146)	(98.69)	(0.0454)
FIN-READABILITY	0.0000609	1.354	0.00158
	(0.0000426)	(1.462)	(0.00118)
DISCLOSER*	-0.0000865**	-1.954	-0.00293**
FIN-READABILITY	(0.0000423)	(1.495)	(0.00122)
SIZE	-0.000981***	150.2	0.0862***
	(0.000342)	(95.52)	(0.0110)
BM	0.0000117**	1.522*	0.000138
	(0.00000571)	(0.892)	(0.000134)
STRENGTHS	0.0000799	-29.40	-0.0522**
	(0.000489)	(34.49)	(0.0233)
CONCERNS	-0.000506*	-44.78	-0.0190
	(0.000278)	(46.72)	(0.0183)
LEVERAGE	-0.00285***	-387.4	-0.0480
	(0.000937)	(326.5)	(0.0582)
OWNERSHIP	-0.102	6446.9	-20.91***
	(0.0968)	(12898.1)	(7.679)
FOLLOWING	-0.0000544	37.90	-0.0179***
	(0.0000709)	(28.25)	(0.00581)
ROI	-0.00182	1912.5	-1.411***
	(0.00274)	(1454.3)	(0.199)
Intercept	0.0242***	-1554.9	-0.0310
	(0.00416)	(1072.6)	(0.123)
Observations	926	926	926
R^2	0.097	0.029	0.190

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

5.1. H6 Bog Results

TABLE 13 examines the same research question with a novel readability proxy that is deemed to be more specific to accounting research based on Plain English attributes, the Bog Index (BOG). TABLE 13 presents the empirical results for H, with the interaction of DISCLOSER and the BOG index. In TABLE 13, I observe a significantly negative relation between the interaction of DISCLOSER and BOG and information asymmetry for one of the chosen proxies (DISPERSION). In the presence of financial complexity, the relation between disclosure and information asymmetry is more negatively correlated with DISPERSION.

TABLE 13

Interaction with BOG

Independent Variable	Dependent Variables		
	BIDASKSPREAD	RETVAR	DISPERSION
DISCLOSER	-0.00350	-75.47	0.978**
	(0.00712)	(230.7)	(0.486)
BOG	-0.0000288	0.210	0.000909
	(0.0000784)	(3.076)	(0.00531)
DISCLOSER *	0.0000364	1.193	-0.0119**
BOG	(0.0000817)	(2.744)	(0.00564)
SIZE	-0.000981***	149.1	0.0981***
	(0.000346)	(92.81)	(0.0110)
BM	0.0000123**	1.541*	0.000125
	(0.00000578)	(0.898)	(0.000128)
STRENGTHS	0.000177	-25.48	-0.0593***
	(0.000496)	(33.94)	(0.0229)
CONCERNS	-0.000754***	-51.29	-0.0244
	(0.000263)	(47.77)	(0.0174)
LEVERAGE	-0.00289***	-390.3	-0.0267
	(0.000961)	(330.0)	(0.0579)
OWNERSHIP	-0.0606	7116.4	-16.53**
	(0.120)	(13191.4)	(7.350)
FOLLOWING	-0.0000404	38.38	-0.0184***
	(0.0000721)	(28.35)	(0.00571)
ROI	-0.00237	1911.5	-1.495***
	(0.00290)	(1470.8)	(0.205)
Intercept	0.0288***	-1518.7	-0.168
	(0.00775)	(1164.9)	(0.449)
Observations	926	926	926
R^2	0.074	0.029	0.206

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

6. Additional Analysis and Robustness Checks

6.1. Propensity Score Matching

While the specification of the baseline regressions capture the essence of the underlying theory, they almost certainly suffer from conceptual limitations. In order to explicitly address endogeneity concerns, I propensity score match the disclosing firms to Compustat non-disclosing firms for a final sample of 1290 observations. It is important to note that the decision to disclose the nonfinancial information is a voluntary choice on the part of managers. Managers weigh the perceived costs and benefits of disclosing sustainability information and will choose to disclose only if the perceived benefits outweigh the costs.

Thus, the following logit model is used to examine the disclosure choice.

$$Disclose = \beta_0 + \beta_1 STRENGTHS + \beta_2 CONCERNS + \beta_3 Size + \beta_4 BM + \beta_5 LEVERAGE + \beta_6 FOLLOWING + \beta_7 ROI + \epsilon \quad \text{(x)}$$

where

STRENGTHS=number of proactive ratings for firm in KLD.

CONCERNS=number of damaging ratings for firm in KLD.

SIZE=log of firm's total assets at end of fiscal year.

BM=firm's market-to-book ratio

LEVERAGE=firm's leverage (DLTT+DLC)/(DLTT+DLC+CEQ)

FOLLOWING= Percentage of total shares outstanding held by institutional investors, from the Thomson Reuters 13-F database.

ROI=return on invested capital

Recall that Dhaliwal et al. (2011) claim that firms with better CSR performance are more likely to signal their long-term focus to the market through use of voluntary disclosure. For this reason, I predict a positive coefficient on *STRENGTHS*, the measure of environmental proactive initiative. In regards to the *CONCERNS* variable, economic theory would likely predict a negative coefficient. The environmentally damaging firms might have incentives not to disclose. By refraining from disclosure, they may pool with the other non-disclosing firms of the average type (Healy and Palepu, 2001). There are many hypotheses for determinants of voluntary disclosure and economic theory does not offer a definitive directional prediction. In the environmental disclosure context, it is possible that firms are incentivized to issue disclosures in order to reveal their type. Proprietary cost hypothesis suggests that firms disclose information based on whether the information will hinder their competitive position within the product markets (Verrecchia, 2001). This hypothesis suggests that firms have an incentive to withhold (i.e. not disclose) information that will reduce their competitive position, even if doing so results in more costly equity.

I include a size variable, as measured by the log of the firm's total assets, and expect a positive coefficient on this variable. I also control for growth by including book-to-market ratio (*BM*) of the firm. I predict that the leverage variable (*LEVERAGE)* will be positive as I imagine higher-leverage firms will provide higher-quality disclosures. *LEVERAGE* captures informational demand by debtholders who are concerned about downside risk (Simnett et al., 2009).

Prior research provides evidence of systematic firm-level characteristics that may increase the likelihood that firms will voluntarily disclose sustainability information. Most of the research in the area of voluntary disclosure is plagued by the problem of self-selection bias. The

decision to disclose sustainability reports is a choice and, therefore, the initial sample may suffer from self-selection bias. Propensity score matching addresses the issue of self-selection bias and allows one to disentangle the treatment effect on the outcome (Rosenbaum and Rubin, 1983). The propensity score matching procedure helps to mitigate the likelihood that the sample suffers from systematic bias. The technique is commonly used to establish effects of a treatment or program when a randomized controlled experiment does not exist. The basic idea is to create a new control group. For each observation in the treatment group, I select the most comparable control observation based on the selection variables. Applying the propensity score matching technique, I compare the information asymmetry values for the firms that choose to issue standalone sustainability reports with a matched sample of firms that choose not to disclose. I calculate the propensity score by running a logit model using the disclosure choice model. The propensity score is the conditional probability that the act of disclosing will be chosen by a firm with the defined characteristics. Then, I match each disclosing firm to the closest non-disclosing firm(s) using the nearest-neighbor matching algorithm. There is a clear tradeoff between inclusion of the full breadth of characteristics that distinguish between disclosers and non-disclosers and finding a non-disclosing firm that matches on all the chosen characteristics. I choose the independent variables based on economic theory (Akerlof 1970) and also based on prior environmental disclosure research (Clarkson et al., 2008).

Upon estimation of the disclosure choice model, I find *SIZE, STRENGTHS* and *LEVERAGE* to be positive and statistically significant at the 1% level, as predicted. The coefficient on *BM* also proves to be in line with expectations as it is positive and statistically significant at the 5% level. *CONCERNS* is statistically insignificant in the disclosure choice model.

After matching the disclosing and non-disclosing firms using the calculated propensity scores, I estimate hypotheses three through six again.

6.2. H3 Propensity-Score Matched Results

TABLE 14 presents the propensity-score matched results for the baseline model and the third hypothesis. Once again, each column in the table depicts a different information asymmetry proxy; the disclosure coefficients may be interpreted in the same manner across each of the different model specifications. The baseline results are observed to be robust. The coefficient on the main variable of interest (DISCLOSER) is observed to be negative for all three of the information asymmetry proxies and significant at the 1% level for two of the three proxies. The *DISCLOSER* variable is negative and statistically significant at the 1% level when the variables *RETVAR and DISPERSION* are used as the information asymmetry proxies. I estimate the models again using the propensity-score matched dataset, and the full breadth of tabulated alternative model specifications are included below.

TABLE 14
Base Model
Propensity Score Matched

Independent Variable	Dependent Variables		
	BIDASKSPREAD	RETVAR	DISPERSION
DISCLOSER	-0.000258	-7.712***	-0.0750***
	(0.000214)	(2.821)	(0.0136)
SIZE	-0.000537***	-4.235**	0.0206**
	(0.000131)	(1.705)	(0.0103)
GROWTH	0.000238***	-0.354	0.00204
	(0.0000577)	(0.767)	(0.00309)
STRENGTHS	0.0000780	-19.70***	-0.0716***
	(0.000324)	(4.299)	(0.0153)
CONCERNS	-0.00130***	-6.504***	-0.0125
	(0.000184)	(2.445)	(0.0117)
LEVERAGE	-0.00386***	-34.31***	-0.330***
	(0.000669)	(9.304)	(0.0485)
OWNERSHIP	-0.509***	-24883.9***	-110.4***
	(0.158)	(2134.7)	(13.57)
FOLLOWING	0.00000233	-0.192	-0.00364
	(0.0000642)	(0.856)	(0.00437)
ROI	-0.0117***	-8.558	-1.305***
	(0.00375)	(47.03)	(0.212)
Intercept	0.0220***	172.3***	0.817***
	(0.00176)	(22.65)	(0.132)
Observations	1290	1290	1290
R^2	0.089	0.110	0.195

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

6.3. H4 Propensity-Score Matched Internal Control Weaknesses Results

TABLE 15 presents the propensity-score matched regression results for testing H4. The coefficient on the interaction term *DISCLOSER*ICW* is observed to be significant for two of the three information asymmetry proxies. The hypothesis suggests that, in the presence of ICWs, nonfinancial disclosure and annual dispersion are *more negatively correlated.* The coefficient on the interaction term *DISCLOSER*ICW* is observed to be significantly negative in the model with the proxy DISPERSION as the dependent variable. The results are mixed across proxies as the coefficient is observed to be positive for the BIDASKSPREAD proxy and negative for the DISPERSION proxy. It is also worth noting that the interaction term is not observed to be statistically significant when the RETVAR proxy is utilized. The results in this table are consistent with the empirical findings reported in TABLE 10 (i.e. in the absence of propensity-score matching procedures).

TABLE 15
Interaction with Internal Control Weaknesses
Propensity Score Matched

Independent Variable	Dependent Variables		
	BIDASKSPREAD	RETVAR	DISPERSION
DISCLOSER	-0.000348	-7.895***	-0.0760***
	(0.000219)	(2.872)	(0.0137)
ICW	-0.00110**	-34.52***	0.239**
	(0.000450)	(9.325)	(0.0939)
DISCLOSER*ICW	0.00467**	14.93	-0.427***
	(0.00204)	(10.10)	(0.101)
SIZE	-0.000364***	-4.228***	0.0387***
	(0.000113)	(1.522)	(0.00919)
GROWTH	0.000159***	-0.463	-0.00660**
	(0.0000551)	(0.761)	(0.00291)
STRENGTHS	-0.0000603	-19.31***	-0.0874***
	(0.000321)	(4.201)	(0.0157)
CONCERNS	-0.00114***	-6.728***	0.00421
	(0.000165)	(2.295)	(0.0114)
LEVERAGE	-0.00315***	-32.75***	-0.262***
	(0.000640)	(9.015)	(0.0496)
OWNERSHIP	-0.318**	-24473.9***	-84.87***
	(0.142)	(2009.5)	(12.41)
FOLLOWING	0.00000925	-0.167	-0.00172
	(0.0000653)	(0.860)	(0.00448)
Intercept	0.0189***	170.5***	0.480***
	(0.00126)	(18.39)	(0.109)
Observations	1290	1290	1290
R^2	0.084	0.112	0.177

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

6.4. H5 Propensity-Score Matched Organizational Complexity Results

TABLE 16 presents the empirical results for H5, with the interaction of *DISCLOSER* and *ORGCOMPLEXITY*. The correlation is observed to be positive and statistically significant for all three of the information asymmetry proxies. For organizationally complex firms, the relation between discloser and information asymmetry is more positively correlated across all information asymmetry proxies. As in TABLE 11, the results are at odds with the original hypothesis that the relative importance of the nonfinancial disclosure information increases in an organizationally complex environment. I interpret the result as indicative of a decrease in the relative importance of nonfinancial disclosure issuance in the reduction of investor uncertainty within organizationally complex environments. In the presence of organizational complexity, I also posit that nonfinancial disclosure issuance is not, on average, beneficial in the reduction of information asymmetry. That is, in this particular context, the issuance of the nonfinancial disclosure does appear to be sending a signal to investors. In this uncertain information environment, I expect that changes to the *content* of the nonfinancial disclosure may be more informative. For firms with information environments marked by organizational complexity, *issuance* of the nonfinancial disclosure significantly increases proxies for information asymmetry.

TABLE 16
Interaction with Organizational Complexity
Propensity Score Matched

Independent Variable	Dependent Variables		
	BIDASKSPREAD	RETVAR	DISPERSION
DISCLOSER	-0.00737***	-98.87***	-0.603***
	(0.00125)	(15.95)	(0.0785)
ORGCOMPLEXITY	-0.0142***	-198.8***	-0.751***
	(0.00106)	(14.75)	(0.0699)
DISCLOSER*	0.0112***	145.7***	0.798***
ORGCOMPLEXITY	(0.00167)	(21.68)	(0.109)
SIZE	-0.000263**	-2.464*	0.0426***
	(0.000104)	(1.457)	(0.00860)
GROWTH	0.000226***	0.549	-0.00355
	(0.0000523)	(0.714)	(0.00277)
STRENGTHS	0.00107***	-3.520	-0.0377**
	(0.000330)	(4.231)	(0.0177)
CONCERNS	-0.000923***	-3.095	0.0140
	(0.000154)	(2.184)	(0.0112)
LEVERAGE	-0.00318***	-34.07***	-0.249***
	(0.000639)	(8.431)	(0.0472)
OWNERSHIP	0.155	-18301.7***	-65.84***
	(0.151)	(2022.4)	(11.44)
FOLLOWING	-0.00000927	-0.552	-0.00312
	(0.0000610)	(0.783)	(0.00418)
Intercept	0.0250***	253.6***	0.836***
	(0.00138)	(19.98)	(0.107)
Observations	1290	1290	1290
R^2	0.164	0.204	0.228

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

6.5. H6 Propensity Score Matched Bog Results

TABLE 17 presents the empirical results for H6, with the interaction of DISCLOSER and the BOG index. The interaction term is observed to be significantly positive for all three information asymmetry proxies. The observed finding indicates that, in the presence of increased financial complexity, issuance of the nonfinancial disclosure corresponds to an increase in information asymmetry. Tabulated results are reported below in TABLE 17, below.

TABLE 17
Interaction with Bog
Propensity Score Matched

Independent Variable	Dependent Variables		
	BIDASKSPREAD	RETVAR	DISPERSION
DISCLOSER	-0.0160***	-247.9***	-0.769***
	(0.00292)	(37.77)	(0.187)
BOG	-0.000127***	-3.136***	-0.0148***
	(0.0000250)	(0.298)	(0.00138)
DISCLOSER*	0.000177***	2.687***	0.00769***
BOG	(0.0000328)	(0.421)	(0.00211)
SIZE	-0.000522***	-3.399**	0.0260***
	(0.000136)	(1.637)	(0.00908)
GROWTH	0.000271***	0.0656	0.00281
	(0.0000569)	(0.745)	(0.00329)
STRENGTHS	0.000142	-18.93***	-0.0704***
	(0.000325)	(4.224)	(0.0157)
CONCERNS	-0.00142***	-9.670***	-0.0282***
	(0.000191)	(2.504)	(0.0109)
LEVERAGE	-0.00397***	-32.84***	-0.310***
	(0.000649)	(8.911)	(0.0498)
OWNERSHIP	-0.513***	-24916.4***	-110.3***
	(0.162)	(2202.7)	(11.87)
FOLLOWING	-0.0000360	-1.122	-0.00799*
	(0.0000628)	(0.853)	(0.00419)
ROI	-0.0145***	-84.53*	-1.688***
	(0.00375)	(45.60)	(0.215)
Intercept	0.0335***	453.4***	2.133***
	(0.00260)	(34.79)	(0.174)
Observations	1290	1290	1290
R^2	0.111	0.175	0.266

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

6.6. Existing Limitations, and Future Research Extensions

As previously indicated, self-selection bias is prevalent within the disclosure literature. The disclosure changes are unlikely to be random events and are likely to coincide with changes in firm-level economics and governance (Healy et al., 2002). In order to further control for the effects of potential self-selection bias, it may be beneficial to estimate a Heckman two-stage model[20] in future robustness checks. Doing so will lend additional credence to the analysis and perhaps shed additional light on the observed phenomena. Firms with more transparent financial disclosure policies could also have better nonfinancial disclosures, thereby inducing a self-selection problem when analyzing sustainability reporting. Although the sample is certainly large enough to use instrumentation, I have not implemented the analysis in this fashion as of now. Future drafts of the research could conceivably incorporate this method as a sensitivity analysis; instrumentation would be empirically advantageous. In the meantime, propensity score matching is a sufficient and appropriate alternative method to address possible empirical shortcomings of the research design. It is worth noting that I observed some mixed results after conducting the propensity score matching procedure. In the subsequent chapter, I further examine the research questions by conducting a more thorough analysis surrounding the implications of nonfinancial disclosure content.

6.7. Conclusions

This chapter provides concrete evidence suggesting that the issuance of the nonfinancial disclosure acts as a signal to market participants. I also examine the effectiveness of the disclosure in reducing information asymmetry in the context of uncertain information environments. In some environments (ICW and diminished readability), the signal is

[20] The existing propensity score matching procedure is likely adequate in the absence of an appropriate instrument.

documented to be more informative while, in others (namely, contexts marked by organizational complexity) the signal is observed to be less informative.

Chapter III
Nonfinancial Disclosure Innovations and Information Asymmetry

1. Introduction

Chapter III directly builds on the empirical conclusions from Chapter II. In this chapter, I consider the dissemination of information across multiple different disclosure channels, and the extent to which the textual attributes of each disclosure channel affect the incremental informativeness of the nonfinancial disclosures relative to the financial disclosures in a capital market context. The previous chapter attests to the capital market implications of nonfinancial disclosure and establishes that the issuance of nonfinancial information acts to reduce information asymmetry in multiple different information environments. While it is clear that the firm-level decision to issue the reports sends a signal to the market relative to the firms that do not disclose ESG information at all, it still remains unclear as to whether investors heed changes to the textual content of the nonfinancial information over time. In order to assess the extent to which the information contained in the disclosed reports is informative from one year to the next, I introduce a measure of nonfinancial disclosure innovation. In order to more effectively demonstrate that the nonfinancial disclosure innovations are incrementally informative relative to financial information, I also develop a textual measure of financial disclosure innovation.

I seek to advance the literature by empirically assessing the extent to which the content of voluntary nonfinancial disclosures enhance the information environment relative to financial disclosure information. In doing so, I build on a large body of textual analytics literature and

further contribute to the burgeoning Natural Language Processing (NLP) subset of the accounting literature.

Prior research has suggested that ceasing to disclose entirely tends to increase analyst forecast dispersion and decrease forecast accuracy (Chen et al., 2011). On a similar note, one might expect that when firms refrain from updating the information in the nonfinancial disclosures from one year to the next, the lack of new information will yield less informational benefits relative to new disclosure innovations. That is, to the extent that the disclosure information is boilerplate from one year to the next, I expect the information to be less informative whereas new nonfinancial disclosure content will likely yield increased informational benefits. The research at hand is both relevant and important as prior literature has suggested that disclosure type is of significance, and highlights the importance of considering multiple different disclosure channels in assessing the capital market implications of disclosure levels (Botosan et al., 2002[21]). Prior literature also suggests that aggregating across different disclosure types results in a loss of information (Botosan and Plumlee, 2002); thus, the incorporation of additional detail regarding both the financial and nonfinancial disclosures aids in verifying the validity of the prior chapter's analysis.

I begin Chapter III with a thorough review of the textual analytics literature as it relates to accounting, and I introduce the Chapter III hypotheses. I discuss the sample and data collection process including the measurement of disclosure and nonfinancial disclosure innovations and readability scores. Then, I incorporate the new measurements in the subsequent empirical analysis. Lastly, I present the empirical results and conclude the chapter.

[21] Botosan et al. (2002) find that greater total disclosure is not associated with lower cost of equity capital. However, the relation between disclosure level and cost of equity capital is observed to vary by type of disclosure.

2. Textual Analytics Literature

Researchers have long recognized the importance of corporate textual disclosures, and a large stream of the accounting literature focuses on qualitative attributes of firm communications. Attributes such as tone and readability have been demonstrated to have a variety of implications for both investors and information intermediaries. Recent research highlights the overall relevance of ESG disclosures as well, with specific emphasis on the textual content of such disclosures. Despite the importance of textual analytics of the aforementioned disclosures, the task of quantifying large bodies of unstructured text is nontrivial. Few studies have sought to examine the content of ESG disclosures. The research at hand seeks to examine the interplay between the content of financial disclosures (i.e. 10-K annual report) and that of nonfinancial disclosures. To my knowledge, this is the first study that seeks to quantify the interplay between financial and nonfinancial disclosures through use of both textual analytics and content analysis. I preface the empirical analyses that follow with a review of the textual analytics literature.

The subsequent component of this research can be categorized under the well-known topic of automated text analysis or textual analytics. The primary aim of automated text analysis is to quantify textual information (Bao and Datta, 2014). Automated text analysis techniques are widely used as a quantitative method within the social sciences, and such techniques draw on machine learning, text mining, and natural language processing. Natural language processing methods are gaining traction in the accounting discipline as well. For the purposes of the subsequent research, I quantify textual information from multiple channels of information dissemination.

Text analytics have been historically implemented in accounting research to further understand the accounting information conveyed to the market. Textual analytics refers to the extraction of meaningful information from text. The technique ranges from extracting specific words, phrases, or sentences to extracting latent pattern structures that are contained within the text of interest. Text analytics is often deemed textual analysis in the context of accounting research[22]. Textual analysis research draws on a variety of different disciplines such as computational linguistics natural language processing, and computer science; the technique yields constructs such as sentiment, content, readability and emotion.

Manual content analysis has gained traction in the accounting discipline since the 1980s and 1990s. Early research in the area primarily focuses on readability with relatively small quantities of text, and manually records the selected aspects of the text. Research by Dorrell and Darsey (1991) is the first to automate the content analysis in accounting. However, most of the early studies tend to base their results on small samples within select industries (Jones and Shoemaker, 1994).

Coinciding with the increase in computer power over the next several years, research has begun to implement text analytic techniques at a larger scale. Li (2008) examines the relation between annual report readability and both firm performance and earnings persistence. The research finds that annual reports of firms with lower earnings are harder to read, and firms with annual reports that are easier to read have more persistent positive earnings[23]. Along with automation, research now incorporates tone and sentiment classification with Naïve Bayes and other statistical classifiers (Antweiler and Frank 2004; Das and Chen 2007; Li 2010).

[22] The computer science discipline often refers to text analytics as text mining.
[23] Li (2008) measures readability through use of a measure originating in computational linguistics, the Fog index, and length of the document. In the subsequent section, I will discuss the Fog index in more detail.

While prior research often employs general purpose word lists (e.g. Harvard IV), Loughran and McDonald (2011) bring attention to potential misspecification through use of dictionaries borrowed from other contexts aside from business or accounting. Their publication introduces the Loughran and McDonald Sentiment Word Lists. This accounting-specific list contains sets of positive, negative, modal strong and weak, litigious, and constraining words for use in textual analysis. The Loughran and McDonald dictionaries have gained traction in accounting research over the last several years.

The topic of readability comprises an important subset of the textual analysis literature as it relates to firm-level disclosures. Readability refers to the ease with which readers can process and comprehend a written body of text. In a variety of contexts, empirical research has shown that the quantity of textual reporting has temporally increased (Lang and Stice-Lawrence, 2015). As of late, excessive, unnecessarily complex, and redundant disclosures have been a topic of interest for the SEC. In 2013, the SEC began to review regulation to identify these potentially unfavorable types of disclosures. FASB is also involved in an agenda project (the Disclosure Framework) aimed to assess overall effectiveness of textual disclosures. Motivated by prior investor and regulator concerns that corporate disclosures are potentially becoming lengthier, more redundant, and less readable over time, Dyer, Lang, and Stice-Lawrence (2017) illustrate that a temporal trend exists in 10-K disclosure from 1996 to 2013. In a way, this article is an extension of their prior work that contests the validity of the Guay et al. (2016) article. During this time period, length, boilerplate, and stickiness are observed to increase. Specificity, readability, and relative amount of hard information are observed to decrease over this time period. The authors indicate that the main drivers for the increasing trend in 10-K disclosure length is likely due to new FASB and SEC requirements. More specifically, the authors note that

fair value, internal controls, and risk factor disclosures constitute the majority of the increase in disclosure length.

The quantity of studies that focus on readability of financial reports has increased significantly over the course of the last decade[24]. Recent research provides significant evidence that the market responds to the readability of financial reports. A study by Lawrence (2013) empirically demonstrates that retail investors are more likely to invest in firms that have shorter, more readable financial reports. Research by De Franco et al. (2015) indicates that trading volume reactions are increasing in the overall readability of analysts' reports. You and Zhang (2009) observes that the firms with longer 10-K filings have a larger delay in the market reaction to the 10-K filings.

Another stream of the disclosure readability literature suggests that financial reporting readability is likely related to future investment and profitability. For instance, Biddle et al. (2009) demonstrates that more readable financial reports are associated with lower over-investment and under-investment. In addition, Li (2008) observe that less readable and longer reports are associated with lower profitability as well as lower earnings persistence. Bonsall and Miller (2017) also observe that less readable financial disclosures are associated with less favorable ratings, greater bond rating agency disagreement, and a higher cost of debt.

Asay, Libby, and Rennekamp (2018) examine the effect of reporting goals and firm performance on language choices, in an experimental setting. Asay et al. (2018) employ the use of two experiments and a survey of experienced managers. By doing so, the authors evaluate determinants of disclosure readability. Bad news disclosures are found to be less readable than their good news counterparts; participants tend to provide less readable reports when

[24] Some of such literature likely responds to a call by Core (2001) citing a need for more computational linguistic techniques to evaluate disclosure quality.

performance is bad. The effect is magnified when participants are given stronger incentives for self-enhancement (i.e. a reporting goal that alters the incentive structure such that the participants seek to cast the firm in a more favorable light.) Lo, Ramos and Rogo (2017) also demonstrate that readability of annual reports is decreasing in earnings management measures. Specifically, Lo et al. (2017) follow Li (2008) in their use of the fox index to measure disclosure readability. The authors document a significantly positive relation between management of earnings to beat prior year's earnings and complexity of the management discussion and analysis section of the annual report (MD&A).

Readability represents an important differentiating factor between disclosures, and the literature has also identified a variety of implications of disclosure readability. A stream of accounting research indicates that adverse information environment implications may arise as a result of financial statement complexity. Guay, Samuels, and Taylor (2016) examine the relation between financial statement complexity and voluntary disclosure. That is, the authors hypothesize that managers use voluntary disclosure to mitigate adverse effects associated with financial statement complexity, and find this result to hold true in their empirical setting. The relation is observed to be stronger when liquidity is lower and in the presence of additional outside monitors. The presence of poor performance and greater earnings management attenuates the relation between financial statement complexity and voluntary disclosure.

Dyer, Lang, and Stice-Lawrence (2016) build off of prior research by Guay et al. (2016) to examine the validity of Guay's empirical findings. Guay et al. (2016) demonstrates that the firms with longer and more complex 10-Ks also tend to issue more voluntary disclosures; the primary measure for voluntary disclosure is management forecasts. A monotonic relation is shown to exist between frequency of management forecasts and quintiles of annual report length

and readability. Management forecast frequency is also shown to be significantly associated with annual report readability and length. However, the Dyer et al. (2016) article questions the overall validity of the study by questioning the causal claims that are drawn in the previous research. The authors indicate that correlated omitted variables are likely an issue and that an upward temporal trend also exists in annual report length, complexity, and voluntary disclosure. The reader is left with the impression that more research is in order for future examination with respect to economic drivers of voluntary disclosure.

As previously noted, the SEC has emphasized use of plain English with the intent to ensure that disclosure informativeness and readability. Readability of disclosures has been experimentally shown to prompt investor's use of simplifying heuristics such that investors rely more heavily on the disclosure as it is perceived to be more credible than the less readable counterparts. In an experimental setting, Rennekamp (2012) finds more readable disclosures to be positively associated with small investor reactions. Changes in valuation judgments are amplified in the sense that such changes are more positive in the presence of good news and more negative in the presence of bad news. When the investors are explicitly told of potential changes in disclosure readability, the simplifying heuristic finding no longer holds.

Lawrence (2013) attests to the relevance of readable disclosure practices and highlights some potential benefits to investors. On average, Lawrence (2013) finds that investors invest more in firms that have clear, concise financial disclosure. The relation is attenuated in the presence of high frequency trading and financial literate individuals. One would anticipate that readability of disclosures would lend itself to mitigation of informational disadvantages. The authors find returns to be increasing in the clarity and conciseness of disclosure.

While most recent studies place emphasis on an individual or investor's interpretation of the disclosure, a recent study by Allee, Deangelis, and Moon (2018) assesses the relative ease with which a computer program or programmer might be able to transform the unstructured disclosure data into usable information through a new measure of disclosure scriptability.

Natural Language processing (hereby, NLP) is also a burgeoning technique in the textual analysis subset of accounting research. The field of NLP is devoted to understanding the interpretation of the human language. The accounting discipline is in the initial phases of incorporating NLP into accounting research. In NLP, topic models aid researchers in defining the latent structure (i.e. topics) of a collection of documents. The topic models are algorithms used to uncover the prevalent themes that pervade a large, unstructured collection of documents.

Latent Dirichlet Allocation (hereby, LDA) methods are beginning to surface in the literature as well. LDA is a Bayesian method to assess the content of the document at hand, and the method marks one of the most prevalent methods for conducting topic modeling. LDA is based on the assumption a set of specific topics exist within each document; the set follows a Dirichlet prior for each document. The words contained within topics also have a Dirichlet prior (Blei, Ng, and Jordan, 2003). Bao and Datta (2014) present a novel application of LDA by quantifying risk disclosures. The authors study the extent to which the risk disclosures in 10-Ks affect risk perceptions of investors, and find that the majority of risk types lack informativeness while those that are informative do not necessarily increase investors' risk perceptions. Bird, Karolyi, and Ma (2018) also implement the LDA technique in the context of 8-K categorization mismatches to demonstrate that such misclassified disclosures are associated with less investor attention[25].

[25] Misclassification is also found to be more likely in the context of negative news and when market attention is high.

Brown Crowley and Elliott (2018) also use latent dirichlet allocation to assess whether thematic content of financial statement disclosures is incrementally informative in predicting intentional misreporting, and find several meaningful topics that are predictive of financial misreporting. The current literature has also implemented content-based analyses to detect changes in 10-Ks over time (Dyer, Lang and Stice-Lawrence, 2017) and classify content of disclosures through multiple different dissemination channels (Crowley, 2016).

3. Nonfinancial Disclosure Innovations and Information Asymmetry

3.1. Hypothesis Development

Voluntary disclosure research is firmly rooted within the financial disclosure literature. In addition to the wide breadth of voluntary disclosure theory research in the financial reporting arena, the theory has also been adapted to suit research surrounding environmental and social governance disclosure. Within the scope of a firm's disclosure strategy, ESG disclosure has become more prominent over the last decade. The foundation for voluntary disclosure theory's application in the realm of environmental and social disclosure builds on the work of Verrecchia (1983) and Dye (1985) in that both studies implement formal analytical modeling to examine the choice to voluntary disclosure (or withhold) information. Verrecchia demonstrates that, due to the existence of proprietary costs associated with information disclosure, non-disclosure will not be unequivocally interpreted as bad news. In the absence of disclosure, Dye (1985) also suggests that "investors may be uncertain about the nature of the information a manager possesses." Therefore, Lang and Lundhom (1993) conclude that, in the presence of adverse selection, firms whose performance exceeds a certain threshold will disclose. On the other hand, those firms below the threshold will not. Based on the prior

research, Clarkson et al. (2008) introduces the notion that firms with better environmental performance will want to signal this via disclosure while worse performing companies will redact the information in an attempt to fall into the 'average type' category.

The signaling argument is prevalent in the nonfinancial disclosure literature. The signaling argument applies if certain conditions hold. The conditions include (1) information asymmetry between managers and potential users of the environmental information, and (2) potential proprietary costs associated with the use of the data. Further, Guidry and Patten (2012) assess the validity of the voluntary disclosure theory logic within the context of environmental disclosure, and suggest that application of Verrecchia and Dye's theoretical work to the environmental disclosure decision would be valid if the conditions (1) and (2) are met.

Based on the prior literature, it is evident that both financial and nonfinancial disclosures are tools that managers can use to inform investors. A large body of accounting literature attests to the informativeness of the content of financial disclosures and the respective cost of capital implications. However, despite its growing importance to investors and stakeholders alike, extant research has not thoroughly addressed implications that stem from content of nonfinancial disclosures. I seek to do so by first assessing whether differences exist in the informational provisions of sustainability disclosures. I then examine whether such differences in the informational provisions of sustainability disclosures impact the extent to which the disclosures tend to reduce information asymmetry.

In order to more thoroughly demonstrate the interplay between multiple channels of information dissemination (i.e. financial disclosure and ESG disclosure), I build on the baseline analysis in Chapter II through incorporation of textual analysis. While the Chapter II results

provide clear indication of the capital market response to issuance of nonfinancial information, the prior analysis does not address the extent to which investors respond to the content contained therein. Thus, the subsequent hypotheses and related empirical analyses seek to further address the content of the disclosures. I first do so by examining the extent to which the disclosures, both financial and nonfinancial, vary in content from one year to the next. I denote the inherent changes in the content of the disclosures (both nonfinancial and financial) as disclosure innovations. The Chapter III analysis remains similar in format to that of Chapter II. However, this analysis focuses on the subset of firms that elect to issue standalone sustainability reports. For the disclosing firms, I introduce a new disclosure innovation measurement. I also introduce a financial disclosure innovation measurement for all firms. The nonfinancial disclosure innovation and financial innovation variables will be described in additional detail in the subsequent Sample and Data Collection section.

I draw on the first Chapter II hypothesis in my suggestion that additional nonfinancial disclosure content is likely to impart additional information to the capital market, thereby reducing information asymmetry. As previously noted, accounting research asserts that voluntary disclosures are often boilerplate. To the extent that innovations in the textual content of the disclosures at hand act to inform the capital market, I expect to observe a negative relation between nonfinancial disclosure innovations and information asymmetry after controlling for financial disclosure innovations. That is, market participants tend to incorporate information from multiple different dissemination channels, and the variation in the content provided via the nonfinancial disclosure dissemination channel is incrementally informative beyond that of the financial disclosure innovations. The decision to issue additional nonfinancial disclosure information rather than retain a boilerplate disclosure from

one year to the next is a strategic choice at the firm level. Thus, I anticipate that capital market implications will ensue if the information is, indeed, informative. I expect information asymmetry to be decreasing in the degree to which firms alter their nonfinancial disclosures. I state this prediction formally and in the alternative form, as the seventh hypothesis below

H7: Ceteris paribus, information asymmetry is decreasing in nonfinancial disclosure innovations.

I build on the Chapter II analysis and logic in the expectation of a stronger relation between nonfinancial disclosure innovations and information asymmetry in alternative information environments. As in the prior chapter, I begin by testing the baseline hypothesis in the context of organizational complexity. That is, as organizational complexity increases, I expect to observe a stronger relation between nonfinancial disclosure innovations and information asymmetry. I state this prediction formally and in the alternative form, as the eighth hypothesis below.

H8: Ceteris paribus, the negative relation between nonfinancial disclosure innovations and information asymmetry is stronger in the presence of organizational complexity.

In addition, I expect nonfinancial disclosure innovations to be more informative in the presence of internal control weaknesses. Thus, the next hypothesis parallels that of Chapter II in structure, but differs in specificity through incorporation of nonfinancial disclosure innovations. The ninth hypothesis is stated in the alternative form below.

H9: Ceteris paribus, the negative relation between nonfinancial disclosure innovations and information asymmetry is stronger in the presence of internal control weaknesses.

I also build on the Chapter II analysis through incorporation of disclosure readability as an alternative information environment. Theoretical work has demonstrated that market reactions are slower in the face of information that is difficult to interpret and extract (Grossman and Stiglitz, 1980). Asay, Elliott and Rennekamp (2017) examine whether investors react differently to information in disclosures that are less readable. The research does so through use of an experiment where investors are presented with disclosures that vary in readability in their presentation of mixed news regarding overall firm performance and finds that investors who read disclosures with diminished readability tend to incorporate more outside information (analyst reports, news media, etc.) in their valuation judgments. When investors are faced with opaque disclosures in the form of low readability, they tend to rely more heavily on outside information in their evaluation of the firm. This finding has strategic implications, because the empirical accounting literature often indicates that firms disclose in an attempt to obfuscate their true (potentially poor) performance. One would be tempted to think that firms issue less readable disclosures with the aim of obscuring poor performance. However, given the findings of this experimental research, one might question the efficacy of this strategic obfuscation tactic; investors would instead be expected to merely shift their emphasis to outside information in their evaluation of the firm. In the context of the research at hand, I build on the research by Asay (2017) to incorporate the alternative ESG disclosure channel. I expect that investors will tend to incorporate more information that has been disseminated through the soft ESG disclosure channel when presented with a financial disclosure of diminished readability. I thereby expect that as financial disclosure readability decreases, the informativeness of ESG disclosure innovations in a capital market context will increase. I state this prediction formally and in the alternative form, as the tenth hypothesis below.

H10: Ceteris paribus, the negative relation between nonfinancial disclosure innovations and information asymmetry is stronger in the presence of a less readable financial report (i.e. higher financial disclosure linguistic complexity).

3.2. Sample Selection and Data Collection

3.2.1. Sample

As the main objective of Chapter III is to examine the incremental informativeness of nonfinancial disclosure content relative to that of the financial disclosure, the sample is comprised only of firms that issue standalone sustainability disclosures.[26] As the research builds on the second chapter, I draw off of the Chapter II data collection process to include information asymmetry proxies, and a variety of relevant control variables.

3.2.2. Main Variables

3.2.3. Financial and Nonfinancial Disclosure Innovations

I hand-collect 1,457 ESG disclosures as PDFs from the firm websites and the GRI website over the 2014-2019 time frame. I obtain the disclosure publication year from the GRI website due to inter-firm ESG disclosure name inconsistencies. In order to analyze the content of the PDF documents, I convert the ESG disclosure PDFs to text files. I download the 10-K Annual reports from the SEC company filings that are publicly available on EDGAR. I bulk download the raw text filings via Python by converting html to text files. I match the sustainability reports to the existing data by firm name. After matching the sustainability reports with their respective financial disclosure, Compustat, IBES, CRSP, and Thomson Reuters data, the resulting sample contains 540 observations.

[26] For consistency, I choose to omit the few firms that have elected to issue an integrated report. Such firms report their ESG information alongside that of their financial information in the 10-K Annual Report.

To calculate the degree to which nonfinancial and financial disclosures vary from one year to the next, I use a machine learning approach. After creating a dictionary from the words in all of the documents, I convert each document into a bag of words (words and counts for each document). Then, I employ word embeddings as a manner to represent the text. In Python, I use the Gensim library to develop Word2Vec embedding from all of the sentences in each document. Word2Vec is an algorithm with the capacity to learn word embedding from a text corpus. The word embedding approach yields a vector representation of words that capture intricacies regarding their meaning. Mikolov et al. (2013) recently introduced the Word2Vec software, and the approach has begun to gain traction for word embeddings, although the approach is not yet prevalent in the accounting literature. The learning models behind the software are described in Mikolov et al. (2013a; 2013b)[27]. To create the similarity score, I train a Word2Vec model on the disclosure data.

Based on the entire corpus of text, the algorithm trains a set of fixed-length dense and continuous-valued vectors. In the embedded space, each word represents a point. Points are learned and then shuffled around based on the words that surround the target word. Therefore, words are defined based on the words that surround them. The premise behind Word2Vec is to train a simple neural network with a hidden layer to perform a certain task, but the neural network will not be used for the task that it has been trained on. Instead, I use the learned weights of the hidden layer.

I load and organize the text into sentences and provide them to the constructor of a new Word2Vec instance. I tokenize each sentence (i.e. divide each sentence into words), remove stop words, non-dictionary words, and lemmatize the body of text as a means of pre-processing the disclosures. The process of lemmatization allows for identification of more complex forms and

[27] Word2vec is based on the skip-gram model.

for the regrouping of the forms that correspond to the same root with different inflexions. Through this process, I change verbs to infinitives, plural words to singular, and group forms that correspond to the same root.

I utilize the word embeddings model to create a vector that represents each disclosure document. Then, I compare this vector to that of the prior year through use of the cosine similarity. The cosine similarity is determined by the angle between the vectors and not their lengths (magnitudes). The cosine similarity is robust to length and repetition.

Disclosure innovations represent the extent to which a firm changes the textual content of the disclosure from one year to the next. I focus on the textual narratives rather than the quantitative information in assessing the extent to which each disclosure is boilerplate from one year to the next. The construct captures the extent to which the disclosure is boilerplate from one year to the next. The technique utilizes dictionary words and omits numbers (aside from longhand numbers that are also dictionary words). By focusing on the textual content of the disclosure, I also assume that the quantitative content of the disclosures is captured through quantitative control variables or that the textual content captures the essence of the quantitative content.

In order to ensure validity of the construct, I conduct face validity assessment for some of the disclosures to ensure robustness of the measure. When the algorithm reports low levels of innovation, the disclosures are largely boilerplate from one year to the next. To the extent that the level of innovation increases, I observe more changes to the textual content of the disclosure. At higher levels of innovation, the firm has expended more effort in altering the disclosure content.

3.2.4. Financial Disclosure Readability

As in Chapter II, I employ several measures of readability based on the prior literature. Several common readability measures are employed across accounting, finance, and computational linguistic disciplines. I use the Natural Language Tool Kit (NLTK) in Python to calculate each of the following readability measures: (1) Flesch Reading Ease formula, (2) Flesch-Kincaid Grade Level, (3) SMOG Index (4) Automated Readability Index, (5) The Coleman-Liau Index, (6) Linsear Write Formula, (7) Dale-Chall Readability Score, and (8) Fog Index. As I expect the readability measures to be highly correlated, I estimate a consensus readability score that incorporates measures (1) through (8) to ensure a robust readability score.

Much of the existing accounting literature relies on the Fog Index to measure financial reporting readability and complexity. Gunning (1952) was the first to begin counting poly-syllabic words to obtain a measure of semantic difficulty. Gunning (1952) introduced the Fog index, and the measure employs two components: (1) average sentence length and (2) percentage of complex words. Complex words are defined as the words that have three or more syllables. The two components are multiplied together. The sum is then multiplied by a scalar to gauge the reading grade level.

$$Fog\ Index = 0.4(average\ number\ of\ words\ per\ sentence + percentage\ of\ complex\ words)$$

Higher Fog Index values indicate less readable text. Use of the Fog Index in accounting applications is not without critique. Loughran and McDonald (2014a) raise the concern that business texts contain a large percentage of words that are classified as complex under this metric simply due to their syllable count. Words such as depreciation and liability are often well understood by investors and analysts yet are denoted as complex under this measure.

Despite the prevalence of the Fog index in the accounting literature, alternative readability measures also exist. These additional readability measures are also worthy of discussion as I include several readability measures in the subsequent analyses to ensure robustness. Lexicon count is one of the most basic readability metrics. The measure simply requires calculation of the number of words present in the text. Studies also employ a similar textual analysis measure that simply counts the number of sentences in any given body of text. Document length, words, and file size represent less commonly used measures of disclosure readability. As previously mentioned, Loughran and McDonald (2014a) criticize the use of the Fog Index as a readability measure in the accounting context. Instead, the authors suggest use of total file size of the 10-K. The total file size is defined as the number of megabytes used in the entire 10-K filing recorded on the EDGAR filing system. The aforementioned study by You and Zhang (2009) employs document length as a measure of disclosure quantity. The authors define document length as the number of words in the filing.

A wide breadth of readability measurements exist in the accounting discipline. Several formulas that resemble the Fog Index in their interpretation, but vary in their formula include the following: (1) Flesch Reading Ease formula, (2) Flesch-Kincaid Grade Level, (3) SMOG Index (4) Automated Readability Index, (5) The Coleman-Liau Index, (6) Linsear Write Formula, (7) Dale-Chall Readability Score (8) a consensus readability score that incorporates the measures (1) through (7) in addition to the Fog Index. The Flesch Reading Ease and Flesch-Kincaid Grade Level formulas use the same components as Fog (the percentage complex words and average number of words per sentence). The Flesch Reading Ease formula and Flesch-Kincaid Grade Level differ in that the measures employ an explicit count of complex word syllables rather than the binary classification used for the Fog index. Both the Fog Index and the Flesch-Kincaid

Grade Level produce numeric estimates of grade level. For instance, a Flesch-Kincaid Grade Level of 9.3 indicates that a ninth grader would be able to read the document with ease. The Flesch Reading Ease formula differs in that it returns the Flesch Reading Ease Score based on a scale of 0-100 where score ranges represent levels of difficulty. For instance, a Flesch Reading Ease Score in the range of 90-100 indicates a "Very Easy" text whereas a score of 0-29 is deemed to be "Very Confusing"[28].

The SMOG index measure is also comparable in the sense that the interpretation is that of grade levels[29]. The following SMOG formula yields the grade level at which the text can be interpreted with ease.

$$Smog = 1.0430 \sqrt{number\ of\ polysyllables\ x \frac{30}{number\ of\ sentences}} + 3.1291$$

The Automated Readability Index (ARI), Linsear Write equations, and Dale-Chall Readability Score also yield grade levels. The Dale-Chall Readability Score differs in that the Dale-Chall Readability Score requires a lookup table of the most commonly used 3000 English words.

The SEC provides clear guidance in their recommendation that managers employ certain Plain English attributes. Managers can do so by avoiding writing construct such as passive voice, weak or hidden verbs, legal and financial jargon, numerous defined terms, abstract words, unnecessary details, lengthy sentences, and unreadable design and layout in financial disclosure (SEC, 1998b; Bonsall, Leone, Miller, and Rennekamp, 2017).

In addition to the aforementioned readability measures obtained from Python textstat[30], I also incorporate the Bog Index as a proxy for financial disclosure complexity. Bonsall, Leone,

[28] A more thorough explanation of the Flesch Reading Ease Score can be found in the Appendix section where I enumerate each score range accompanied by the associated interpretation.
[29] A thorough explanation of the original SMOG calculation process as detailed by McLaughlin (1969) is included in the Appendix.

Miller, and Rennekamp (2017) introduce the Bog Index, a new multi-faceted measure of readability based on Plain English attributes[31]. With their new Bog Index measure, the authors intend to capture plain English attributes of the disclosures. Then, the researchers proceed to define their own readability measurement.

In the empirical analysis at hand, I employ the Bog Index as an additional measure of financial disclosure linguistic complexity. The Bog Index is constructed via a software program called StyleWriter. The program identifies attributes that are contained within the SEC Plain English Handbook. For instance, the software captures qualitative attributes such as sentence length, passive voice, weak verbs, overused words, complex words, and jargon. The software does not immediately conclude that every multi-syllabic word is complex, a criticism of prior measurements. Instead, the word complexity is measured based on a lengthy word list. The Bog Index provides a summary of the writing attributes that tend to bog readers down. The index is computed as the sum of Sentence Bog, Word Bog and Pep as follows:

Bog Index=Sentence Bog + Word Bog – Pep

A higher Bog Index indicates diminished readability. Sentence Bog includes readability issues stemming from sentence length. Longer sentences imply lower readability. StyleWriter identifies average sentence length across the entire document, and the average sentence length is squared and scaled by 35 words per sentence. Word Bog is comprised of plain English style problems as well as word difficulty. The plain English style problems and word difficulty are summed, multiplied by 250 and then divided by the number of words. Word difficulty is

---

[30] Textstat is a Python package used to calculate statistics from text to determine readability, complexity, and grade level of a particular corpus.

[31] In conjunction with the proposal of this new measure, the authors ensure the validity of the measurement by running a series of experiments, and also subjecting the measure to empirical verification as compared to the alternate readability measures contained in the literature to date.

computed based on a proprietary list of over 200,000 words. Based on this proprietary list, certain abstract words receive higher scores relative to less abstract words.

The last component of the Bog Index is Pep. Pep includes writing attributes that aid in understanding of texts. It is a sum of names and interesting words that tend to make writing more interesting. Pep is calculated to be the sum of these components that aid in interpretation multiplied by 25 and scaled by the number of words in the document plus sentence variety (standard deviation of sentence length multiplied by ten and scaled by the average sentence length).

Chapter III variables of interest are detailed in TABLE 18 Summary Statistics, and TABLE 19 presents a correlation matrix.

TABLE 18

Chapter III Summary Statistics

	Mean	25th Percentile	Median	75th Percentile	Standard Deviation
ESGINNOVATION	0.0928	0.0265	0.0502	0.1019	0.1244
FININNOVATION	0.0178	0.0029	0.0047	0.0081	0.0468
FINREADABILITY	-36.2144	-33.0000	-23.0000	-20.0000	33.2118
BIDASKSPREAD	0.0134	0.0101	0.0110	0.0139	0.0051
RETVAR	49.1448	6.6642	19.3297	60.7337	73.5928
DISPERSION	0.4953	0.2635	0.4103	0.6463	0.3308
LOGASSET	10.1576	9.4324	10.1300	10.9648	1.3975
BM	4.5258	2.0985	2.9731	4.6254	4.0395
LEVERAGE	0.5074	0.3691	0.4916	0.6411	0.1877
ICW	0.0092	0.0000	0.0000	0.0000	0.0958
STRENGTHS	1.0671	0.7500	1.1111	1.4167	0.5686
CONCERNS	0.8047	0.0000	0.5000	1.5000	0.8792
OWNERSHIP	0.0014	0.0006	0.0010	0.0019	0.0011
FOLLOWING	4.2717	3.0000	4.0000	5.0000	2.0143
ROI	0.0930	0.0493	0.0797	0.1247	0.0538
ORGCOMPLEXITY	0.7256	0.6362	0.7646	0.8198	0.1213
N	541				

This table presents the summary statistics for Chapter III variables of interest.

TABLE 19

Correlation Matrix

	ESGINNOVATION	FININNOVATION	FINREADABILITY	BIDASKSPREAD	RETVAR	DISPERSION	LOGASSET	BM	LEVERAGE	STRENGTHS	CONCERNS	OWNERSHIP	FOLLOWING	ORGCOMPLEXITY	ROI
ESGINNOVATION	1.00														
FININNOVATION	-0.07	1.00													
FINREADABILITY	-0.10**	0.02	1.00												
BIDASKSPREAD	-0.17***	-0.04	0.14***	1.00											
RETVAR	-0.15***	-0.07*	0.15***	0.51***	1.00										
DISPERSION	-0.06	0.03	0.07*	-0.13***	0.17***	1.00									
LOGASSET	0.41***	0.03	-0.01	-0.16***	0.09**	0.33***	1.00								
BM	-0.10**	-0.01	0.20***	0.19***	0.26***	-0.09**	-0.17***	1.00							
LEVERAGE	0.05	-0.06	0.02	-0.06	0.06	0.04	0.19***	0.46***	1.00						
STRENGTHS	0.09**	0.07	0.11***	0.00	0.09**	0.10**	0.36***	0.23***	0.20***	1.00					
CONCERNS	0.42***	-0.06	-0.22***	-0.13***	0.03	0.14***	0.55***	-0.13***	0.25***	0.26***	1.00				
OWNERSHIP	-0.26***	-0.12***	-0.05	0.11**	-0.21***	-0.23***	-0.55***	-0.20***	-0.16***	-0.60***	-0.32***	1.00			
FOLLOWING	0.04	0.08**	0.11***	0.02	0.09**	-0.10**	0.10**	0.11***	-0.01	0.20***	-0.04	-0.24***	1.00		
ORGCOMPLEXITY	0.31***	0.06	0.14***	-0.13***	-0.06	-0.05	0.24***	0.08**	0.05	0.36***	0.26***	-0.16***	0.01	1.00	
ROI	-0.24***	0.08*	0.19***	0.13***	0.11***	-0.19***	-0.31***	0.54***	-0.09**	0.25***	-0.36***	-0.27***	0.14***	-0.02	1.00
N	541														

This table presents Pearson correlation coefficients for the variables of interest in Chapter III.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3.3. Empirical Methods

3.3.1. Empirical Model: H7

I test the seventh hypothesis that, ceteris paribus, information asymmetry is decreasing in nonfinancial disclosure innovations through estimation of the regression below. Equation **(**xi**)** details the empirical model used to test the hypothesis. Note that, in the context of Chapter III, it is no longer necessary to utilize 3SLS as an empirical method as nonfinancial disclosure is no longer the dependent variable of interest. Simultaneity bias is no longer cause for concern. The key parameter of interest in the regression model is the parameter pertaining to nonfinancial disclosure innovations (NFINNOVATION). I also expect financial disclosure innovations to be correlated with the independent variable of interest (NFINNOVATION) and the information asymmetry proxies. Thus, I control for financial disclosure innovations as a means to mitigate omitted variable bias. The model is completed through inclusion of a set of controls.

H7: Ceteris paribus, information asymmetry is decreasing in nonfinancial disclosure innovations.

$$\text{Information asymmetry} = \beta_0 + \beta_1 \text{NFINNOVATION} + \theta \text{Controls} \quad \text{(xi)}$$

3.3.2. Empirical Model: H8

H8: Ceteris paribus, the negative relation between nonfinancial disclosure innovations and information asymmetry is stronger in the presence of organizational complexity.

$$\begin{aligned}\text{Information asymmetry} &= \beta_0 + \beta_1 \text{NFINNOVATION} + \beta_2 \text{NFINNOVATION} * \text{ORGCOMPLEXITY} \\ &+ \beta_3 \text{ORGCOMPLEXITY} + \theta \text{Controls}\end{aligned} \quad \text{(xii)}$$

3.3.3. Empirical Model: H9

H9: Ceteris paribus, the negative relation between nonfinancial disclosure innovations and information asymmetry is stronger in the presence of internal control weaknesses.

$$\text{Information asymmetry} = \beta_0 + \beta_1 \text{NFDISC} + \beta_2 \text{NFINNOVATION} * \text{ICW} + \beta_3 \text{ICW} + \theta \text{Controls} \quad \text{(xiii)}$$

3.3.4. Empirical Model: H10

H10: Ceteris paribus, the negative relation between nonfinancial disclosure innovations and information asymmetry is stronger in the presence of financial disclosure linguistic complexity.

$$\text{Information asymmetry} = \beta_0 + \beta_1 \text{NFINNOVATION} + \beta_2 \text{NFINNOVATION} * \text{FINREADABILITY} + \beta_3 \text{FINREADABILITY} + \theta \text{Controls} \quad \text{(xiv)}$$

3.4. Empirical Results

3.4.1. H7 Results: Nonfinancial Disclosure Innovations

The results of the H7 model examining the relation between nonfinancial disclosure innovations and information asymmetry, as specified in equation **(xi)**, appear in TABLE 20. I observe the coefficient of NFINNOVATION to be significantly negative at the 1% level for all of the information asymmetry proxies. A negative relation between NFINNOVATION and the information asymmetry proxies provides evidence suggesting that changes to the content of the nonfinancial disclosures correspond to a reduction in information asymmetry. I also observe a significantly negative relation between financial disclosure innovations (FININNOVATION) and return variance (RETVAR). The relation is significant at the 1% level as well.

TABLE 20

Nonfinancial Disclosure Innovation Baseline Model

Independent Variable	Dependent Variables		
	BIDASKSPREAD	RETVAR	DISPERSION
NFINNOVATION	-0.00487***	-165.9***	-0.695***
	(0.00156)	(20.49)	(0.107)
FININNOVATION	-0.00410	-182.0***	-0.0189
	(0.00322)	(37.72)	(0.311)
SIZE	0.00000838	7.290***	0.0669***
	(0.000247)	(2.722)	(0.0167)
GROWTH	0.000421***	7.398***	0.000901***
	(0.0000754)	(1.293)	(0.000156)
STRENGTHS	0.000698	-14.46**	-0.00633
	(0.000656)	(7.330)	(0.0260)
CONCERNS	0.0000352	4.090	-0.0339*
	(0.000315)	(4.641)	(0.0196)
LEVERAGE	-0.00591***	-74.16***	-0.121
	(0.00112)	(17.83)	(0.0771)
OWNERSHIP	0.612*	-17156.2***	-86.04***
	(0.359)	(3167.2)	(19.95)
FOLLOWING	0.0000471	1.235	-0.0250***
	(0.000101)	(1.602)	(0.00797)
ROI	-0.00740	-237.5***	-1.718***
	(0.00597)	(85.18)	(0.354)
Intercept	0.0138***	50.78	0.354*
	(0.00300)	(32.20)	(0.207)
Observations	541	541	541
R^2	0.102	0.188	0.234

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3.4.2. H8 Results: Organizational Complexity Results

The results of the H8 model examining the relation between nonfinancial disclosure innovations and information asymmetry in the context of organizational complexity, as specified in equation **(xii)**, appear in TABLE 21. In the context of organizational complexity, I expect that nonfinancial disclosure innovations will prove to be more informative in a capital market context. That is, I expect the interaction between NFINNOVATION and ORGCOMPLEXITY to significantly negative. The interaction term is observed to be significantly negative for one of the three proxies and insignificantly negative for two of the three proxies.

TABLE 21

Nonfinancial Disclosure Innovations, Organizational Complexity, and Information Asymmetry

Independent Variable	Dependent Variables		
	BIDASKSPREAD	RETVAR	DISPERSION
NFINNOVATION*	-0.00273	-138.6***	-0.472
ORGCOMPLEXITY	(0.00484)	(48.32)	(0.335)
NFINNOVATION	-0.00107	-55.45	-0.370
	(0.00627)	(56.86)	(0.404)
FININNOVATION	-0.00267	-177.2***	0.00842
	(0.00331)	(37.62)	(0.312)
ORGCOMPLEXITY	-0.00687***	-20.85	-0.179
	(0.00217)	(26.82)	(0.147)
SIZE	0.0000913	7.842***	0.0741***
	(0.000236)	(2.744)	(0.0169)
GROWTH	0.000462***	7.543***	0.00888*
	(0.0000762)	(1.280)	(0.00473)
STRENGTHS	0.00132**	-12.04	0.00272
	(0.000668)	(7.447)	(0.0297)
CONCERNS	0.000137	4.385	-0.0246
	(0.000308)	(4.656)	(0.0198)
LEVERAGE	-0.00657***	-76.04***	-0.183*
	(0.00114)	(17.29)	(0.102)
OWNERSHIP	0.805**	16321.2***	-81.11***
	(0.364)	(3275.0)	(20.26)
FOLLOWING	0.0000252	1.198	-0.0262***
	(0.000100)	(1.613)	(0.00794)
ROI	-0.00792	-237.1***	-1.908***
	(0.00606)	(84.93)	(0.412)
Intercept	0.0170***	55.93	0.409*
	(0.00346)	(35.16)	(0.241)
Observations	541	541	541
R^2	0.123	0.189	0.232

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3.4.3. H9 Results: Internal Control Weaknesses Results

In the context of internal control weaknesses, I also expect the content of the nonfinancial disclosures to be more informative. The results of the H9 model examining the relation between nonfinancial disclosure innovations and information asymmetry in the context of internal control weaknesses, as specified in equation **(xiii)**, appear in TABLE 22. I observe a significantly negative relation between NFINNOVATION and all three information asymmetry proxies, reinforcing that the nonfinancial disclosure innovations may be incrementally informative. Further, the relation between the interaction of NFINNOVATION and ICW and the information asymmetry proxies is found to be significantly negative for two of the three proxies. The interaction term is significantly negative at the 5% level for the RETVAR proxy and significant at the 1% level for the BIDASKSPREAD proxy. Results are observed to be somewhat mixed, as the interaction term is observed to be significantly positive for the DISPERSION proxy.

TABLE 22

Nonfinancial Disclosure Innovations Interaction with Internal Control Weaknesses

Independent Variable	Dependent Variables		
	BIDASKSPREAD	RETVAR	DISPERSION
NFINNOVATION	-0.00318*	-157.5***	-0.657***
	(0.00174)	(21.52)	(0.114)
ICW	0.0187***	32.36	-0.354***
	(0.00297)	(21.53)	(0.0769)
NFINNOVATION*ICW	-0.350***	-1524.2**	6.086***
	(0.0964)	(680.4)	(1.608)
FININNOVATION	-0.00271	-177.2***	0.0259
	(0.00328)	(37.71)	(0.314)
ORGCOMPLEXITY	-0.00704***	-30.87	-0.220
	(0.00214)	(25.47)	(0.141)
SIZE	0.000126	7.608***	0.0726***
	(0.000230)	(2.749)	(0.0167)
GROWTH	0.000455***	7.595***	0.00917*
	(0.0000760)	(1.283)	(0.00474)
STRENGTHS	0.00104	-12.11	0.0106
	(0.000677)	(7.554)	(0.0301)
CONCERNS	0.000158	4.514	-0.0242
	(0.000306)	(4.614)	(0.0196)
LEVERAGE	-0.00643***	-77.72***	-0.188*
	(0.00115)	(17.54)	(0.102)
OWNERSHIP	0.626*	-16403.1***	-76.43***
	(0.363)	(3245.1)	(20.51)
FOLLOWING	0.0000262	1.135	-0.0264***
	(0.0000999)	(1.604)	(0.00792)
ROI	-0.00675	-234.3***	-1.933***
	(0.00600)	(85.45)	(0.413)
Intercept	0.0171***	66.42**	0.440*
	(0.00329)	(33.19)	(0.227)
Observations	541	541	541
R^2	0.148	0.192	0.233

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3.4.4. H10 Results: Financial Disclosure Linguistic Complexity Results

The results of the H10 model examining the relation between nonfinancial disclosure innovations and information asymmetry in the context of financial disclosure linguistic complexity, as specified in equation (viii), appear in TABLE 23. I employ a consensus readability score that incorporates eight prevalent readability metrics in the literature. The variable has been transformed such that an increase in this readability score indicates that the document is less readable. Thus, I anticipate that increases in the readability score (i.e. increased linguistic complexity of the financial disclosure) will enhance the ability of nonfinancial disclosure innovations to reduce information asymmetry. I observe the interaction term to be significantly negative for one of the three information asymmetry proxies, DISPERSION. This empirical finding provides some evidence that, as financial statement complexity increases, nonfinancial disclosure innovations are more informative in a capital market context.

TABLE 23

Nonfinancial Disclosure Innovations Interaction with Consensus Readability

Independent Variable	Dependent Variables		
	BIDASKSPREAD	RETVAR	DISPERSION
NFINNOVATION	-0.00343	-178.4***	-0.927***
	(0.00240)	(27.01)	(0.122)
FIN-READABILITY	0.00000975	0.254***	0.00156***
	(0.0000101)	(0.0736)	(0.000453)
NFINNOVATION*	0.0000324	-0.344	-0.00520**
FIN-READABILITY	(0.0000353)	(0.274)	(0.00228)
FININNOVATION	-0.00408	-179.0***	0.00370
	(0.00321)	(38.85)	(0.319)
SIZE	-0.0000805	6.544**	0.0705***
	(0.000255)	(2.790)	(0.0166)
GROWTH	0.000406***	7.028***	0.00549
	(0.0000765)	(1.272)	(0.00464)
STRENGTHS	0.000610	-15.80**	-0.0202
	(0.000673)	(7.391)	(0.0262)
CONCERNS	0.000172	6.289	-0.0169
	(0.000335)	(4.691)	(0.0198)
LEVERAGE	-0.00603***	-70.24***	-0.121
	(0.00122)	(18.42)	(0.103)
OWNERSHIP	0.557	-17266.7***	-83.33***
	(0.353)	(3211.3)	(19.53)
FOLLOWING	0.0000275	1.119	-0.0253***
	(0.000103)	(1.628)	(0.00825)
ROI	-0.00804	-235.6***	-1.836***
	(0.00606)	(85.73)	(0.414)
Intercept	0.0153***	67.22**	0.368*
	(0.00318)	(34.16)	(0.207)
Observations	541	541	541
R^2	0.110	0.196	0.239

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

4. Summary

In this chapter, I build on the empirical conclusions from Chapter II by empirically illustrating that the content of nonfinancial disclosures is informative in a capital market context. I illustrate that, while disclosure innovations are disseminated through both financial and nonfinancial disclosure channels, nonfinancial disclosure innovations are incremental informative in their reduction of information asymmetry relative to financial disclosure innovations. The analyses in Chapter III demonstrate that investors heed changes to the textual content of the nonfinancial information in addition to changes in the textual content of financial information.

The analysis advances the literature by empirically demonstrating the extent to which the content of voluntary nonfinancial disclosures enhance the information environment relative to financial disclosure information. I conduct the analysis in the context of alternative information environments, namely in the presence of organizational complexity, internal control weaknesses, and financial statement complexity. I observe that nonfinancial disclosure innovations are more informative in the presence of more organizational complexity, more internal control weaknesses, and less readable financial disclosures. My results build on the textual analytics literature and further contribute to the burgeoning Natural Language Processing (NLP) subset of the accounting literature by illustrating that the textual content of nonfinancial disclosures is incrementally informative.

Limitations

The empirical analyses conducted in this dissertation are not without shortcomings, and it is important that the reader is aware of some potential limitations of the existing research. I seek to address these challenges in the paragraphs below, but I also encourage my committee members to bring any logical, procedural or empirical concerns to my attention for incorporation in future revisions.

Throughout the duration of the dissertation, I focus on investors as the primary beneficiaries of enhanced financial and nonfinancial disclosure. In this respect, the theoretical framework builds on information economics theory. Nonetheless, one cannot rule out the possibility that firms formulate their disclosure strategy with a wider audience in mind, thereby encompassing other stakeholders as well. Both financial and nonfinancial disclosures are tools that managers can use to inform investors and stakeholders. I acknowledge that throughout the duration of this study, I predominantly focus on ESG disclosure as a practice that informs market participants. However, it is also important to consider that alternative audiences for the information likely exist outside the realm of investors.

I follow Starr (2005, p. 360) in my conservatism with the use of control variables. Starr argues that the inclusion of control variables in empirical models should be based on sound theoretical reasoning and only after "fairly extensive preliminary data analysis reveals….the form of the relationship." Therefore, I include only control variables supported by voluntary disclosure theory with valid theoretical justifications.

There may also be limitations surrounding the Word2Vec model implementation and the corresponding ESG and financial innovation variables that result from this process. One apparent shortcoming is that the variables capture the extent to which a disclosure changes over

time (i.e. whether the disclosure is boilerplate form one year to the next). So, while the variable effectively captures the degree to which disclosure content varies from one year to the next, the variable does not strictly address whether more information has been imparted. If, perhaps, information is redacted from one year to the next, this method treats this withholding of information as an innovation as well. Thus, the variable very effectively captures the extent to which the disclosure changes over time (i.e. is not boilerplate). However, the variable does not solely address the quantity of new information that has been disclosed. This issue can be overcome through incorporation of additional similarity scores in future research. The innovation variables generated through use of Word2Vec capture the extent to which a disclosure is boilerplate over time. Future research will seek to incorporate additional variables that capture incremental information dissemination.

Conclusion

In Chapter I, I introduce hypotheses that examine the interplay of financial disclosure innovations and nonfinancial disclosure innovations in different competitive environments with the intent to ultimately examine the extent to which competition from potential rivals influences firms' voluntary disclosure decisions. I do not observe evidence supporting that firms utilize ESG disclosure innovations as a strategic means to deter entry of potential rivals. I observe a significantly negative relation between nonfinancial disclosure innovations and financial disclosure innovations. I also observe that, in the presence of proprietary costs, the tradeoff is attenuated or eliminated. Overall, Chapter I sheds light on some potential determinants of nonfinancial disclosure innovations.

In Chapter II, I examine whether the disclosure of nonfinancial information, as proxied by the publication of stand-alone sustainability reports, is associated with a reduction in information

asymmetry. After implementing propensity score matching procedures and controlling for potentially confounding variables, I find that the issuance of nonfinancial information, as proxied by sustainability reports, is significantly negatively associated with information asymmetry. Additional analyses divulge that the relation is stronger amongst firms with internal control weaknesses (as a means to gauge financial opacity) and organizational complexity. This finding lends credence to the assertion that the nonfinancial disclosures act to supplement financial disclosures in the reduction of investor uncertainty.

In Chapter III, I build on the empirical conclusions from Chapter II by considering the dissemination of information across multiple different disclosure channels, and the extent to which the textual attributes of each disclosure channel affect the incremental informativeness of the nonfinancial disclosures relative to the financial disclosures in a capital market context. While the prior chapter provides evidence that the firm-level decision to issue the reports sends a signal to the market relative to the firms that do not disclose ESG information at all, Chapter III further examines whether investors heed changes to the textual content of the nonfinancial information over time. I assess the extent to which the information contained in the disclosed reports is informative from one year to the next through similarity scores intended to measure nonfinancial and financial disclosure innovations. In doing so, I examine whether the nonfinancial disclosure innovations are incrementally informative after controlling for financial disclosure innovations. While I believe that drawing definitive conclusions at this point in the empirical analysis is likely premature, I observe some evidence consistent with this assertion.

I further extend the existing literature by examining nonfinancial disclosure implications in a variety of alternative information environments. Namely, I examine the relation in the contexts

of organizational complexity, internal control weaknesses, and financial statement linguistic complexity.

The results presented in this dissertation reinforce the importance of considering various disclosure attributes rather than assuming that disclosures are homogeneous and equally informative. The research contributes to multiple streams of accounting literature. Namely, the research highlights and extends the established association between voluntary disclosure and information asymmetry by incorporating disclosure of nonfinancial information into the discussion. Results of this study are particularly relevant as stakeholder demand for corporate environmental information has increased over the past decade; the information has become material to investor decision-making (Khan, 2016). The results reinforce the argument that the nonfinancial disclosures serve an informative role in the market. The findings also corroborate the recent empirical work by Dhaliwal (2014) regarding CSR disclosure. In doing so, this research also contributes to a better understanding of disclosure implications by thorough examination of the capital market consequences surrounding ESG disclosure innovations.

The overall link between disclosure and information asymmetry has proven to be elusive in prior literature. My findings suggest that nonfinancial disclosures act to inform the capital market in tandem with financial disclosure; by doing so, the research, at least in part, answers the call for further voluntary disclosure research.

References

Akerlof, G. A. (1970). The Market for "Lemons": Quality Uncertainty and the Market Mechanism. *The Quarterly Journal of Economics*, *84*(3), 488.

Ali, A., Klasa, S., & Yeung, E. (2014). Industry concentration and corporate disclosure policy. *Journal of Accounting and Economics*, *58*(2–3), 240–264.

Allee, K. D., & Deangelis, M. D. (2015). The Structure of Voluntary Disclosure Narratives: Evidence from Tone Dispersion: the structure of voluntary disclosure narratives. *Journal of Accounting Research*, *53*(2), 241–274.

Allee, K. D., Deangelis, M. D., & Moon, J. R. (2018). Disclosure "Scriptability": DISCLOSURE "SCRIPTABILITY." *Journal of Accounting Research*, *56*(2), 363–430.

Al-Tuwaijri, S. A., Christensen, T. E., & Hughes, K. E. (2004). The relations among environmental disclosure, environmental performance, and economic performance: a simultaneous equations approach. *Accounting, Organizations and Society*, *29*(5–6), 447–471.

Antweiler, W., & Frank, M. Z. (2004). *The Market Impact of Corporate News Stories*. 39.

Asay, H. S., Elliott, W. B., & Rennekamp, K. (2017). Disclosure Readability and the Sensitivity of Investors' Valuation Judgments to Outside Information. *The Accounting Review*, *92*(4), 1–25.

Asay, H. S., Libby, R., & Rennekamp, K. (2018). Firm performance, reporting goals, and language choices in narrative disclosures. *Journal of Accounting and Economics*, *65*(2–3), 380–398.

Ashbaugh-Skaife, H., Collins, D. W., & Kinney, W. R. (2007). The discovery and reporting of internal control deficiencies prior to SOX-mandated audits. *Conference Issue on Corporate Governance: Financial Reporting, Internal Control, and Auditing*, *44*(1), 166–192.

Bamber, L. S., Jiang, J. (Xuefeng), & Wang, I. Y. (2010). What's My Style? The Influence of Top Managers on Voluntary Corporate Financial Disclosure. *The Accounting Review*, *85*(4), 1131–1162.

Bao, Y., & Datta, A. (2014). Simultaneously Discovering and Quantifying Risk Types from Textual Risk Disclosures. *Management Science*, *60*(6), 1371–1391.

Barth, M. E., Cahan, S. F., Chen, L., & Venter, E. R. (2017). The economic consequences associated with integrated report quality: Capital market and real effects. *Accounting, Organizations and Society*, *62*, 43–64.

Barth, M. E., & McNichols, M. F. (1994). Estimation and Market Valuation of Environmental Liabilities Relating to Superfund Sites. *Journal of Accounting Research*, *32*, 177.

Beneish, M. D., Billings, M. B., & Hodder, L. D. (2008). Internal control weaknesses and information uncertainty. *The Accounting Review*, *83*(3), 665–703.

Bertomeu, J., Beyer, A., & Dye, R. A. (2011). Capital Structure, Cost of Capital, and Voluntary Disclosures. *The Accounting Review*, *86*(3), 857–886.

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.

Biddle, G. C., Hilary, G., & Verdi, R. S. (2009). How does financial reporting quality relate to investment efficiency? *Journal of Accounting and Economics*, *48*(2–3), 112–131.

Billings, M. B., Jennings, R., & Lev, B. (2015). On guidance and volatility. *Journal of Accounting and Economics*, *60*(2–3), 161–180.

Bird, A., Karolyi, S. A., & Ma, P. (2016). Strategic Disclosure Misclassification. *SSRN Electronic Journal*.

Blei, Ng, and Jordan (2003). *Latent Dirichlet Allocation*. 30.

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

Bonsall, S. B., & Miller, B. P. (2017). The impact of narrative disclosure readability on bond ratings and the cost of debt. *Review of Accounting Studies*, *22*(2), 608–643.

Botosan, C. A., & Plumlee, M. A. (2002). A Re-examination of Disclosure Level and the Expected Cost of Equity Capital. *Journal of Accounting Research*, *40*(1), 21–40.

Brown, N. C., & Elliott, W. B. (2016). What are You Saying? Using Topic to Detect Financial Misreporting. *SSRN Electronic Journal*.

Brown, S., & Hillegeist, S. A. (2003). Disclosure quality and information asymmetry. *Kellogg School of Management, Evanston, IL*.

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.

Chen, S., Miao, B., & Shevlin, T. (2015). A New Measure of Disclosure Quality: The Level of Disaggregation of Accounting Data in Annual Reports: A NEW MEASURE OF DISCLOSURE QUALITY. *Journal of Accounting Research*, *53*(5), 1017–1054.

Chen, Y.-C., Hung, M., & Wang, Y. (2018). The effect of mandatory CSR disclosure on firm profitability and social externalities: Evidence from China. *Journal of Accounting and Economics*, *65*(1), 169–190.

Cheng, B., Ioannou, I., & Serafeim, G. *Corporate Social Responsibility and Access to Finance. 44.*

Cheynel, E. (2013). A theory of voluntary disclosure and cost of capital. *Review of Accounting Studies*, *18*(4), 987–1020.

Cho, C. H., & Patten, D. M. (2007). The role of environmental disclosures as tools of legitimacy: A research note. *Accounting, Organizations and Society*, *32*(7–8), 639–647.

Cho, C. H., Phillips, J. R., Hageman, A. M., & Patten, D. M. (2009). Media richness, user trust, and perceptions of corporate social responsibility: An experimental investigation of visual web site disclosures. *Accounting, Auditing & Accountability Journal*, *22*(6), 933–952.

Cho, C. H., & Patten, D. M. (2013). Green accounting: Reflections from a CSR and environmental disclosure perspective. *Critical Perspectives on Accounting*, *24*(6), 443–447.

Cho, C. H., Roberts, R. W., & Patten, D. M. (2010). The language of US corporate environmental disclosure. *Accounting, Organizations and Society*, *35*(4), 431–443.

Cho, S. Y., Lee, C., & Pfeiffer, R. J. (2013). Corporate social responsibility performance and information asymmetry. *Journal of Accounting and Public Policy*, *32*(1), 71–83.

Christensen, H. B., Floyd, E., Liu, L. Y., & Maffett, M. (2017). The real effects of mandated information on social responsibility in financial reports: Evidence from mine-safety records. *Journal of Accounting and Economics*, *64*(2–3), 284–304.

Clarkson, P. M., Fang, X., Li, Y., & Richardson, G. (2013). The relevance of environmental disclosures: Are such disclosures incrementally informative? *Journal of Accounting and Public Policy*, *32*(5), 410–431.

Clarkson, P. M., Li, Y., & Richardson, G. D. (2004). The Market Valuation of Environmental Capital Expenditures by Pulp and Paper Companies. *The Accounting Review*, *79*(2), 329–353.

Clarkson, P. M., Li, Y., Richardson, G. D., & Vasvari, F. P. (2008). Revisiting the relation between environmental performance and environmental disclosure: An empirical analysis. *Accounting, Organizations and Society*, *33*(4–5), 303–327.

Clarkson, P. M., Li, Y., Richardson, G. D., & Vasvari, F. P. (2011). Does it really pay to be green? Determinants and consequences of proactive environmental strategies. *Journal of Accounting and Public Policy*, *30*(2), 122–144.

Comprix, J., Graham, R. C., & Moore, J. A. (2011). Empirical Evidence on the Impact of Book-Tax Differences on Divergence of Opinion among Investors. *The Journal of the American Taxation Association*, *33*(1), 51–78.

Crowley, R. M. (2016). *Disclosure Through Multiple Disclosure Channels*. 116.

Crowley, R. M., Huang, W., & Lu, H. (2018). *Discretionary Dissemination on Twitter*. 74.

Darrough, M. N., & Stoughton, N. M. (1990). Financial disclosure policy in an entry game. *Journal of Accounting and Economics*, *12*(1–3), 219–243.

Das, S. R., & Chen, M. Y. (2007). Yahoo! for Amazon: Sentiment Extraction from Small Talk on the Web. *Management Science*, *53*(9), 1375–1388.

Dedman, E., & Lennox, C. (2009). Perceived competition, profitability and the withholding of information about sales and the cost of sales. *Journal of Accounting and Economics*, *48*(2–3), 210–230.

Dhaliwal, D., Li, O. Z., Tsang, A., & Yang, Y. G. (2014). Corporate social responsibility disclosure and the cost of equity capital: The roles of stakeholder orientation and financial transparency. *Journal of Accounting and Public Policy*, *33*(4), 328–355.

Dhaliwal, D. S., Li, O. Z., Tsang, A., & Yang, Y. G. (2011). Voluntary Nonfinancial Disclosure and the Cost of Equity Capital: The Initiation of Corporate Social Responsibility Reporting. *The Accounting Review*, *86*(1), 59–100.

Diamond, D. W. (1985). *Optimal Release of Information By Firms*. 25.

Diamond, D. W., & Verrecchia, R. E. (1991). *Disclosure, Liquidity, and the Cost of Capital*. 36.

Dorrell, J. T., & Darsey, N. S. (1991). An Analysis of the Readability and Style of Letters to Stockholders. Journal of Technical Writing and Communication, 21(1), 73–83. https://doi.org/10.2190/KD3W-W2AF-60K6-92H6

Doyle, J., Ge, W., & McVay, S. (2007). Determinants of weaknesses in internal control over financial reporting. *Conference Issue on Corporate Governance: Financial Reporting, Internal Control, and Auditing*, *44*(1), 193–223.

Dye, R. A. (1985). Disclosure of Nonproprietary Information. *Journal of Accounting Research,23*(1), 123-145.

Dyer, T., Lang, M., & Stice-Lawrence, L. (2017). The evolution of 10-K textual disclosure: Evidence from Latent Dirichlet Allocation. *Journal of Accounting and Economics*, *64*(2–3), 221–245.

Easley, D., Hvidkjaer, S., & O'hara, M. (2002). Is information risk a determinant of asset returns? *The Journal of Finance*, *57*(5), 2185–2221.

Easley, D., & O'hara, M. (2004). Information and the Cost of Capital. *The Journal of Finance*, *59*(4), 1553–1583.

Eccles, Iannou and Serafeim (2017). The Impact of Corporate Sustainability on Organizational Processes and Performance. *http://ssrn.com/abstract=1964011*

Eccles, R. G., Ioannou, I., & Serafeim, G. (2012). *The impact of a corporate culture of sustainability on corporate behavior and performance*.

Eccles, R. G., Serafeim, G., & Krzus, M. P. (2011). Market interest in nonfinancial information. *Journal of Applied Corporate Finance*, *23*(4), 113–127.

Elkington, J. (1998). Cannibals with forks : the triple bottom line of 21st century business. Gabriola Island, BC ; Stony Creek, CT: New Society Publishers.

Ellis, J.A., Fee, C.E., & Thomas, S.E. (2012). Proprietary Costs and the Disclosure of Information About Customers. *Journal of Accounting Research*, *50*(3), 685–727.

Fekrat, M. A., Inclan, C., & Petroni, D. (1996). Corporate environmental disclosures: Competitive disclosure hypothesis using 1991 annual report data. *The International Journal of Accounting*, *31*(2), 175–195.

Feng, M., Li, C., & McVay, S. (2009). Internal control and management guidance. *Journal of Accounting and Economics*, *48*(2–3), 190–209.

Feng, Li (2010). The Information Content of Forward-Looking Statements in Corporate Filings—A Naïve Bayesian Machine Learning Approach. *Journal of Accounting Research*, *48*(5), 1049-1102.

Fisher, I. E., Garnsey, M. R., & Hughes, M. E. (2016). Natural Language Processing in Accounting, Auditing and Finance: A Synthesis of the Literature with a Roadmap for Future Research: NLP in Accounting, Auditing and Finance. *Intelligent Systems in Accounting, Finance and Management*, *23*(3), 157–214.

Francis, J., Nanda, D., & Olsson, P. (2008). Voluntary Disclosure, Earnings Quality, and Cost of Capital: Disclosure and Earnings Quality. *Journal of Accounting Research*, *46*(1), 53–99.

Francis, J. R., Khurana, I. K., & Pereira, R. (2005). Disclosure Incentives and Effects on Cost of Capital around the World. *The Accounting Review*, *80*(4), 1125–1162. h

Franco, F., Urcan, O., & Vasvari, F. P. (2016). Corporate Diversification and the Cost of Debt: The Role of Segment Disclosures. *The Accounting Review*, *91*(4), 1139–1165.

Friedman, M. (2007). The Social Responsibility of Business Is to Increase Its Profits. In W. C. Zimmerli, M. Holzinger, & K. Richter (Eds.), *Corporate Ethics and Corporate Governance* (pp. 173–178).

Gao, P. (2010). Disclosure Quality, Cost of Capital, and Investor Welfare. *The Accounting Review*, 30.

Ge, W., & McVay, S. (2005). The Disclosure of Material Weaknesses in Internal Control after the Sarbanes-Oxley Act. *Accounting Horizons*, *19*(3), 137–158.

Glaeser, S. (2018). The effects of proprietary information on corporate disclosure and transparency: Evidence from trade secrets. *Journal of Accounting and Economics*, *66*(1), 163–193.

Graham, J. R., Harvey, C. R., & Rajgopal, S. (2005). The economic implications of corporate financial reporting. *Journal of Accounting and Economics*, *40*(1), 3–73.

Greene, W. H. (2012). Econometric Analysis. Upper Saddle River, N.J: Prentice Hall.

Grossman, S. J. (1981). An Introduction to the Theory of Rational Expectations Under Asymmetric Information. *The Review of Economic Studies*, *48*(4), 541.

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

Guidry, R. P., & Patten, D. M. (2012). Voluntary disclosure theory and financial control variables: An assessment of recent environmental disclosure research. *Accounting Forum*, *36*(2), 81–90.

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

Healy, P. M., & Myers, S. C. (2002). R&D Accounting and the Tradeoff between Relevance and Objectivity. *Journal of Accounting Research, 40*(3), 677-710.

Healy, P. M., & Palepu, K. G. (2001). Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical disclosure literature. *Journal of Accounting and Economics*, *31*(1), 405–440.

Healy, P. M., & Serafeim, G. (2016). An Analysis of Firms' Self-Reported Anticorruption Efforts. *The Accounting Review*, *91*(2), 489–511.

Holthausen, R. W., & Verrecchia, R. E. (1988). The Effect of Sequential Information Releases on the Variance of Price Changes in an Intertemporal Multi-Asset Market. *Journal of Accounting Research*, *26*(1), 82–106.

Hope, O.-K. (2003). Disclosure Practices, Enforcement of Accounting Standards, and Analysts' Forecast Accuracy: An International Study. 39.

Huang, A. H., Lehavy, R., Zang, A. Y., & Zheng, R. (2018). Analyst Information Discovery and Interpretation Roles: A Topic Modeling Approach. *Management Science*, *64*(6), 2833–2855.

Ittner, C. D., & Larcker, D. F. (1998). Are Nonfinancial Measures Leading Indicators of Financial Performance? An Analysis of Customer Satisfaction. *Journal of Accounting Research*, *36*, 1.

Jensen (2001). Value Maximization, Stakeholder Theory, and the Corporate Objective Function. *Journal of Applied Corporate Finance*, *12* (1).

Jensen, C., & Meckling, H. (1976). Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure. *Journal of Financial Economics, 3,* 305-360.

Jones & Shoemaker (1994). Accounting Narratives: A Review of Empirical Studies of Content and Readability. *Journal of Accounting Literature, 13,* 142-184.

Khan, M., Serafeim, G., & Yoon, A. (2016). Corporate Sustainability: First Evidence on Materiality. *The Accounting Review*, *91*(6), 1697–1724.

Kim, J.-B., & Zhang, L. (2016). Accounting Conservatism and Stock Price Crash Risk: Firm-level Evidence. *Contemporary Accounting Research*, *33*(1), 412–441.

Kim, O., & Verrecchia, R. E. (1994). Market liquidity and volume around earnings announcements. *Journal of Accounting and Economics*, *17*(1–2), 41–67.

Kim, Y., Park, M. S., & Wier, B. (2012). Is Earnings Quality Associated with Corporate Social Responsibility? *The Accounting Review*, *87*(3), 761–796.

Kothari, S. P., Li, X., & Short, J. E. (2009). The Effect of Disclosures by Management, Analysts, and Business Press on Cost of Capital, Return Volatility, and Analyst Forecasts: A Study Using Content Analysis. *The Accounting Review*, *84*(5), 1639–1670.

Lang, M., & Stice-Lawrence, L. (2014). Textual Analysis and International Financial Reporting: Large Sample Evidence. *Journal of Accounting and Economics, 60*(2-3), 110-135.

Lang, M., & Lundholm, R. (1993). Cross-Sectional Determinants of Analyst Ratings of Corporate Disclosures. *Journal of Accounting Research*, *31*(2), 246.

Lang, M., & Stice-Lawrence, L. (2015). Textual analysis and international financial reporting: Large sample evidence. *Journal of Accounting and Economics*, *60*(2–3), 110–135.

Lang, M., & Sul, E. (2014). Linking industry concentration to proprietary costs and disclosure: Challenges and opportunities. *Journal of Accounting and Economics*, *58*(2–3), 265–274.

Lannelongue, G., & González-Benito, J. (2012). Opportunism and environmental management systems: Certification as a smokescreen for stakeholders. *Ecological Economics*, *82*, 11–22.

Lawrence, A. (2013a). Individual investors and financial disclosure. *Journal of Accounting and Economics*, *56*(1), 130–147.

Leuz, C., & Verrecchia, R. E. (2005). Firms' Capital Allocation Choices, Information Quality, and the Cost of Capital. *SSRN Electronic Journal*.

Lewellen, J., & Shanken, J. (2002). Learning, Asset-Pricing Tests, and Market Efficiency. *The Journal of Finance*, *57*(3), 1113–1145.

Li, F. (2010). Textual Analysis of Corporate Disclosures: A Survey of the Literature. *Journal of Accounting Literature, 29,* 143-165.

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

Li, X. (2010). The impacts of product market competition on the quantity and quality of voluntary disclosures. *Review of Accounting Studies*, *15*(3), 663–711.

Lin, C., & He, Y. (2009). Joint Sentiment/Topic Model for Sentiment Analysis. *https://dl.acm.org/doi/proceedings/10.1145/1645953*.

Lo, K., Ramos, F., & Rogo, R. (2017). Earnings management and annual report readability. *Journal of Accounting and Economics*, *63*(1), 1–25.

Lohmann, L. (2009). Toward a different debate in environmental accounting: The cases of carbon and cost–benefit. *Accounting, Organizations and Society*, *34*(3), 499–534.

Loughran, T., & McDonald, B. (2011). When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks. *The Journal of Finance*, *66*(1), 35–65.

Loughran, T., & McDonald, B. (2015). The Use of Word Lists in Textual Analysis. *Journal of Behavioral Finance*, *16*(1), 1–11.

Loughran, T., & McDonald, B. (2016). Textual Analysis in Accounting and Finance: A Survey: Textual Analysis in Accounting and Finance. *Journal of Accounting Research*, *54*(4), 1187–1230.

Loughran, T., & McDonald, B. (2014). Measuring Readability in Financial Disclosures. *The Journal of Finance, 69* (4), 1643-1671.

Loughran, T., & McDonald, B. (2011). When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks. *The Journal of Finance, 66*(1), 35-65.

Lys, T., Naughton, J. P., & Wang, C. (2015). Signaling through corporate accountability reporting. *Journal of Accounting and Economics*, *60*(1), 56–72.

Matsumura, E. M., Prakash, R., & Vera-Muñoz, S. C. (2014). Firm-Value Effects of Carbon Emissions and Carbon Disclosures. *The Accounting Review*, *89*(2), 695–724.

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient Estimation of Word Representations in Vector Space. *ArXiv:1301.3781 [Cs]*.

Milgrom, P. R. (1981). Rational Expectations, Information Acquisition, and Competitive Bidding. *Econometrica*, *49*(4), 921.

Milne, M. J. and Patten, D.M. (2002). Securing Organizational Legitimacy: An experimental Decision Case Examining the Impact of Environmental Disclosures. *Accounting, Auditing, and Accountability, 15*(3), 372-405

Ogneva, M., Subramanyam, K. R., & Raghunandan, K. (2007). Internal Control Weakness and Cost of Equity: Evidence from SOX Section 404 Disclosures. *The Accounting Review*, *82*(5), 1255–1297.

Pastor and Veronesi (2003). Stock Prices and IPO Waves. *NBER Working Paper Series. http://www.nber.org/papers/w9858.*

Patten, D. M. (2002). The relation between environmental performance and environmental disclosure: a research note. *Accounting, Organizations and Society*, *27*(8), 763–773.

Plumlee, M., Brown, D., Hayes, R. M., & Marshall, R. S. (2015). Voluntary environmental disclosure quality and firm value: Further evidence. *Journal of Accounting and Public Policy*, *34*(4), 336–361.

Plumlee, M., & Yohn, T. L. (2010). An Analysis of the Underlying Causes Attributed to Restatements. *Accounting Horizons*, *24*(1), 41–64.

Rennekamp, K. (2012). Processing Fluency and Investors' Reactions to Disclosure Readability. *Journal of Accounting Research*, *50*(5), 1319–1354.

Richardson, A. J., & Welker, M. (2001). Social disclosure, financial disclosure and the cost of equity capital. *Accounting, Organizations and Society*, *26*(7–8), 597–616.

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

Rogers, J. L., & Van Buskirk, A. (2009). Shareholder litigation and changes in disclosure behavior. *Journal of Accounting and Economics*, *47*(1–2), 136–156.

Rosenbaum, P. R., & Rubin, D. B. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika, 70*(1), 41-55.

S. P. Kothari, Xu Li, & James E. Short. (2009). The Effect of Disclosures by Management, Analysts, and Business Press on Cost of Capital, Return Volatility, and Analyst Forecasts: A Study Using Content Analysis. *The Accounting Review*, *84*(5), 1639–1670.

Simnett, R., Vanstraelen, A., & Chua, W. F. (2009). Assurance on Sustainability Reports: An International Comparison. *The Accounting Review*, *84*(3), 937–967.

Sutton, J. (1991). Sunk costs and market structure : price competition, advertising, and the evolution of concentration. Cambridge, Mass.: MIT Press.

Tetlock, P. C., Saar-Tsechansky, M., & Macskassy, S. (2008). More than Words: Quantifying Language to Measure Firms' Fundamentals. *The Journal of Finance*, *63*(3), 1437-1467.

Ashbaugh-Skaife, H., Collins, D., Kinney, W. & LaFond, R. (2008). The Effect of SOX Internal Control Deficiencies and Their Remediation on Accrual Quality. *The Accounting Review, 83*(1), 217-250.

Tversky, & Kahneman (1974). Judgment under Uncertainty: Heuristics and Biases. *Science, 185* (4157), 1124-1131.

Verrecchia, R. E. (2001). Essays on Disclosure. *Journal of Accounting and Economics, 32* (1-3), 97-180.

Villiers, C. de, & van Staden, C. J. (2011). Where firms choose to disclose voluntary environmental information. *Journal of Accounting and Public Policy*, *30*(6), 504–525.

Watts, R.L. and Zimmerman, J.L. (1986) Positive Accounting Theory. Prentice Hall, Englewood Cliffs.

You, H., & Zhang, X. (2009). Financial reporting complexity and investor underreaction to 10-K information. *Review of Accounting Studies*, *14*(4), 559–586.

Appendices

Appendix A: Data Definitions[32]

Variable		Definition
ASSET	=	Book value of the firm's total assets (AT) at the end of fiscal year *t* (in $ millions);
FOLLOWING	=	The average number of analysts following firm *i* during year *t*;
BIDASKSPREAD	=	Average bid-ask spread, calculated using the absolute spread scaled by the average of bid and ask (bid-ask);
CAPIN	=	Net property, plant, and equipment for firm *i* in year *t* scaled by prior year total assets;
LIAB	=	Book value of the firm's total liabilities (LT) at the end of fiscal year *t* (in $ millions)
OPINC	=	Firm's operating income (in $ millions) after depreciation (OIADP) for fiscal year *t*;
CONCERNS	=	Number of damaging ratings (concerns) for the firm identified in KLD;
STRENGTHS	=	Number of proactive ratings (strengths) for the firm identified in KLD;
SIZE	=	Log of the firm's total assets at the end of the fiscal year;
GROWTH	=	Firm's market-to-book ratio measured as market value of equity divided by book value of equity (CEQ)
LEVERAGE	=	Firm's leverage, measured as (DLTT + DLC)/(DLTT +DLC+CEQ);
INSTOWN	=	Percentage of total shares outstanding held by institutional investors, from the Thomson Reuters 13-F database;
INTANG	=	Intangible assets (including goodwill) for firm *i* in year *t* scaled by prior year total assets;
TURNOV	=	The mean daily trading volume scaled by total market value of equity (share price*shares outstanding) for firm *i* over the one-year period from the second quarter of year *t* to the first quarter of year *t+1*;
RETVAR	=	The standard deviation of daily stock returns for firm *i* over the one-year period from the second quarter of year *t* to the first quarter of year *t+1*;
ROI		The firm's return on invested capital, defined as operating profit scaled by total assets
NFDISC	=	An indicator variable equal to 1 if the firm issues a standalone sustainability report to the public, and 0 otherwise.

[32] Although data for all variables listed above has been collected and appropriately merged, several of the variables have not yet been incorporated into the first pass model. Noted omissions will likely be included as controls in future drafts.

INDCONC		A revenue-based Hirfindahl-Hirschman index intended to measure within-firm concentration by industry, computed as the sum of the square of a firms' sales in a particular segment as a percentage of total firm sales.
ORGCOMPLEXITY		A variable intended to measure organizational complexity, computed as -1**INDCONC+1*.
BOG		An index variable intended to capture readability, per Bonsall et al. (2017). The variable represents a new, multi-faceted measure of readability based on attributes contained in the SEC Plain English Handbook and is calculated via the software Stylewriter. The basic equation is as follows: *Bog Index=Sentence Bog + Word Bog – Pep*
FIN-READABILITY		A consensus readability score for financial disclosures incorporating the most prevalent readability measures. I employ the Natural Language Tool Kit (NLTK) in Python to calculate each of the following (1) Flesch Reading Ease formula, (2) Flesch-Kincaid Grade Level, (3) SMOG Index (4) Automated Readability Index, (5) The Coleman-Liau Index, (6) Linsear Write Formula, (7) Dale-Chall Readability Score, and (8) Fog Index.
IND-PPE	=	The weighted average of property, plant, and equipment of all firms in an industry. The firm's market share, calculated as the ratio of its segment sales to industry aggregate sales, is used as its weight. The firm's segment PP&E is allocated according to the ratio of the segment sales to the firm's total sales.
IND-R&D	=	The weighted average of research and development of all firms in an industry. A firm's market share, calculated as the ratio of its segment sales to industry aggregate sales, is used as its weight. If a firm's segment R&D is missing, it is replaced by a firm's total R&D multiplied by the ratio of the segment sales to the firm's total sales.
IND-CAPEX		The weighted average of capital expenditures of all firms in an industry. A firm's market share, calculated as the ratio of its segment sales to industry aggregate sales, is used as its weight. If a firm's segment capital expenditures are missing, they are replaced by the firm's total capital expenditures multiplied with the ratio of the segment sales to the firm's total sales.
IND-MKT	=	Product market size as measured by the natural log of industry aggregate sales.
COMPETITIONPCA		A variable intended to measure competition from potential entrants obtained via principal component analysis on previously defined IND-MKT, IND-CAPEX, IND-R&D, and IND-PPE variables.

NFINNOVATION	=	A similarity score between the ESG disclosure in time t and that of time t-1, calculated via word embedding in Word2Vec.
FININNOVATION	=	A similarity score between the 10-K in time t and that of time t-1, calculated via word embedding in Word2Vec.
DISPERSION	=	From IBES, the standard deviation of analysts' year t forecasts scaled by the mean analysts' year t forecasts for firm *i* over the last three quarters of year *t*.

Appendix B: Readability Measures

Flesch Reading Ease Score

Score Range	Score Interpretation
90-100	Very Easy
80-89	Easy
70-79	Fairly Easy
60-69	Standard
50-59	Fairly Difficult
30-49	Difficult
0-29	Very Confusing

SMOG Grading Process (per McLaughlin, 1969)

1. Count 10 consecutive sentences near the beginning of the text to be assessed, 10 in the middle, and 10 near the end. A sentence includes any string of words ending with a period, question mark, or exclamation point.
2. In the 30 sentences that have been selected, count every word of three or more syllables. If a polysyllabic word is repeated, each repetition is also counted.
3. Estimate the square root of the number of polysyllabic words counted. That is, take the square root of the nearest perfect square. If the count lies roughly between two perfect squares, the lower number is chosen.
4. Add 3 to the approximate square root calculation.
5. The resulting calculation is the SMOG Index

Automated Readability Index (ARI)

$$ARI = 4.71\left(\frac{characters}{words}\right) + 0.5\left(\frac{words}{sentences}\right) - 21.43$$

Appendix C: Summary Statistics and Correlation Matrix with Inclusion of Firm-Level Competition from Potential Entrants Variable

TABLE 24
Chapter I Summary Statistics

Variable	Mean	25th Percentile	Median	75th Percentile	Standard Deviation
FININNOVATION	0.0167	0.0029	0.0046	0.0080	0.0436
NFINNOVATION	0.0730	0.0243	0.0473	0.0893	0.0886
INDCONC	0.3130	0.1999	0.2827	0.4276	0.1225
COMPETITIONPCA2	0.0000	-0.8819	-0.3729	0.8854	1.0000
SIZE	9.9839	9.3831	9.9802	10.7344	1.0003
BM	4.1050	2.0232	3.1785	4.9757	2.7613
LEVERAGE	0.4825	0.3550	0.4606	0.5950	0.1641
CONCERN	0.6044	0.0000	0.2857	1.1250	0.6858
STRENGTHS	0.9986	0.6250	1.0000	1.4000	0.4727
OWNERSHIP	0.0013	0.0007	0.0011	0.0019	0.0008
ROIC	0.0960	0.0530	0.0887	0.1342	0.0493
N	246				

This table presents the summary statistics for the variables used in Chapter I including an alternative firm-level competition from potential entrants variable (COMPETITONPCA2).

TABLE 25
Correlation Matrix

	FININNOVATION	NFINNOVATION	INDCONC	COMPETITIONPCA2	SIZE	BM	LEVERAGE	CONCERNS	STRENGTHS	OWNERSHIP	ROIC
FININNOVATION	1.00										
NFINNOVATION	-0.03	1.00									
INDCONC	-0.10	-0.14**	1.00								
COMPETITIONPCA2	-0.07	-0.10	0.44***	1.00							
SIZE	0.04	0.15**	-0.04	0.13**	1.00						
BM	0.01	-0.02	-0.08	-0.33***	-0.18***	1.00					
LEVERAGE	-0.06	-0.02	0.05	0.07	0.16**	0.33***	1.00				
CONCERNS	-0.02	0.15**	-0.12*	0.33***	0.43***	-0.21***	0.15**	1.00			
STRENGTHS	0.10	0.04	-0.32***	-0.21***	0.27***	0.26***	0.13**	0.12*	1.00		
OWNERSHIP	-0.14**	-0.12*	0.08	0.12*	-0.61***	-0.27***	-0.13**	-0.20***	-0.55***	1.00	
ROIC	0.10	-0.11*	0.02	-0.30***	-0.24***	0.59***	-0.12*	-0.34***	0.27***	-0.34***	1.00
N	246										

This table presents Pearson correlation coefficients for the variables of interest in Chapter I, including a firm-level metric for competition among potential entrants (COMPETITIONPCA2).

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$