The Effect of 2004 8-K Expansion on Information Asymmetry Among Investors

by

Eric Zhihong He

Business Administration
Duke University

Date:_____

Approved:

_____
Katherine Schipper, Co-supervisor

_____
Rahul Vashishtha, Co-supervisor

_____
Xu Jiang

_____
Emma Rasiel

Dissertation submitted in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy in Business Administration
in the Graduate School of
Duke University

2022

ABSTRACT

The Effect of 2004 8-K Expansion on Information Asymmetry Among Investors

by

Eric Zhihong He

Business Administration
Duke University

Date:_____
Approved:

_____
Katherine Schipper, Co-supervisor

_____
Rahul Vashishtha, Co-supervisor

_____
Xu Jiang

_____
Emma Rasiel

An abstract of a dissertation submitted in partial
fulfillment of the requirements for the degree of
Doctor of Philosophy in Business Administration
in the Graduate School of
Duke University

2022

Abstract

Disclosure complexity may increase information asymmetry between sophisticated and retail investors, due to unsophisticated investors' more limited capacity to process complex disclosures. I explore whether a 2004 Securities and Exchange Commission (SEC) rule that changed the timing of material contract disclosures helps mitigate this problem. The change requires firms to file material contracts on Form 8-K within a few days after signing instead of including those contracts in the next periodic filing, thereby allowing investors more time to process periodic filings. Using data from periodic financial statements and material contracts filed between 2001 and 2007, I find that the 2004 requirement significantly lowers information asymmetry among investors immediately following periodic financial statement disclosures. The effect is more pronounced for firms that file more complex contracts, attract more retail investors, and experience greater change in the timeliness of their material contract filings. The effect is also larger for firms whose financial statements are disclosed during periods when investors are more time constrained. This evidence should be of interest to the SEC and other regulators who strive to level the information playing field among investors and to maintain fair and efficient markets.

Dedication

I dedicate this work to my wife, Amelia, who with incredible compassion has accompanied me in this journey and always believed in my ability to succeed; to my dad, Xiaoan, for his unconditional love, support, and encouragement through all the difficult times in my life; and to the memory of my mom, Wei.

Contents

List of Tables

List of Figures

Acknowledgements

First, I would like to thank my co-supervisors, Katherine and Rahul, for the opportunity to learn from and work with them. Throughout my entire doctoral studies, they have provided invaluable guidance and support. Not only have they taught me how to think critically and conduct rigorous academic research, but they have also shown me what it means to be a caring mentor to others. I am forever grateful for their inspiration, advice, and encouragement.

Second, I would like to express my sincere gratitude to my committee members, Xu and Emma, for their constructive feedback on my research. They have shown me how to use theories to guide empirical work and how to clearly communicate my research to others. I would also like to thank Bill, Mohan, and other accounting faculty members at Duke University for generously providing comments and suggestions on various pieces of my work.

Finally, I thank Matt, Shuyan, Vincent, Yufei, and other past and current Ph.D. student colleagues in the accounting program. I thank Martin, whom I have enjoyed collaborating with. I am also truly grateful to Patricia, Preeti, Daniel, and Joon, for introducing me to accounting research and encouraging me to start this doctoral journey when I was working as a Research Assistant at the Public Company Accounting Oversight Board.

1. Introduction

It has long been established that information processing is costly (Grossman and Stiglitz 1980; Blankespoor et al. 2019). Because investors have different information processing capabilities (Indjejikian 1991; Blankespoor et al. 2019), disclosures with high processing costs can lead to an increase in information asymmetry between sophisticated and unsophisticated investors. Recognizing this issue, regulators have long been concerned about how to reduce information processing costs and thereby level the information playing field between sophisticated and unsophisticated investors.[1] One way to achieve this goal is to raise investors' awareness of disclosures and improve access to required filings such as financial reports. For example, studies have shown that the Securities and Exchange Commission (SEC)'s introduction of EDGAR[2] as an electronic central repository for regulatory filings leveled the playing field for small investors, increased the volume of small but not large trades, and improved informativeness of individual investors' trades (Asthana and Balsam 2001; Asthana et al. 2004; Gao and Huang 2020). Another way to reduce information

[1] The Securities and Exchange Commission (SEC) states on its webpage that "all investors, whether large institutions or private individuals, should have access to certain basic facts about an investment prior to buying it." In 2014, Keith Higgins, then Director of Corporation Finance at the SEC, stated: "Disclosure effectiveness is not just about what information is disclosed to investors, but how it is disclosed as well.… There is a growing concern about disclosure overload. As the number of pages in annual reports has steadily increased, it may become more difficult for investors to find the most salient information."

[2] EDGAR refers to the SEC's Electronic Data Gathering, Analysis, and Retrieval system.

processing costs is by improving readability and promoting easier-to-use formats in required filings. For example, in 1998, the Commission adopted "plain English" rules for all required filing. In another effort, the SEC requires financial statements to be provided in interactive data format using XBRL[3] starting 2009 (SEC 2009). In this paper, I study whether moving certain complex and logically severable parts of periodic filings to a different and more timely disclosure form reduces information asymmetry among investors by allowing investors more time to process the periodic disclosures.

Specifically, I study the effect of a 2004 SEC rule change on information asymmetry between sophisticated and retail investors in the days following periodic financial statement filings. Before the mandate, firms were required to file material contracts as exhibits to 10-K/Q reports; subsequent to the mandate, they disclose those contracts on 8-K forms filed within four business days of the contract signing date.

I choose the material contracts setting for several reasons. First, the 2004 mandate affects only the timing of disclosures while leaving the overall content unchanged. Under both the old and new regimes, investors have access to both the material contracts and the financial statements as of the periodic financial statement disclosure date. After the rule change, contracts are disclosed earlier, providing investors more time to process the contracts ahead of financial statement filings on Form

[3] XBRL is the eXtensible Business Reporting Language format.

10-K/Q and reducing the complexity of the 10-K/Q reports. The separation of material contracts disclosures from periodic financial reports gives investors more time to focus on the periodic reports when they are disclosed, but does not affect the total quantity of information provided. This feature of the rule change allows me to disentangle the effect of disclosure timing from the amount of information disclosed. Second, material contracts can be complex as they describe the obligations, rewards, and penalties of the contracting parties in alternative future states of the world (Eggleston, Posner, and Zeckhauser 2000). Therefore, moving the disclosure of these contracts to a separate filing preceding the financial reports should reduce investors' cognitive load in processing the periodic reports. Finally, prior literature suggests that material contracts are filed frequently and investors pay attention to them (Lerman and Livnat 2010). Taken together, these features of this setting make it particularly appropriate for my study.

Because it is more costly for investors to process more complex contracts even if the language in the contracts is simple, my analysis requires a specialized material contract complexity measure. Traditional linguistic complexity proxies from the accounting literature tend to capture the complexity of language and sentences. In addition to using these measures, I create two novel content-complexity measures tailored to material contracts: (1) the number of payoff-relevant contingencies; and (2) whether the contract is a renewal, in which case investors should be more familiar with

the arrangement. I use textual analysis to collect this information from each sample contract. I also collect traditional linguistic-complexity measures: file size, word count, sentence count, the Fog Index, the Flesch Reading Ease score, the SMOG grade, and the Dale-Chall readability score.[4] I aggregate these contract-level data to the firm-quarter level and use Principal Component Analysis (PCA) to reduce the dimensionality of the complexity measures. While my content complexity measures are correlated with some traditional linguistic complexity measures, the PCA suggests the two types of measures capture different underlying (latent) factors.

Blankespoor, DeHaan, Wertz, and Zhu (2019) describe a conceptual framework for analyzing three types of information processing costs: awareness costs, acquisition costs, and integration costs.[5] On the one hand, when material contracts are disclosed earlier in a given quarter, investors can process them earlier, allowing more time to integrate the information contained in the 10-K/Qs and the information in the material contracts. In other words, investors' integration costs for 10-K/Q reports is reduced under the new regime. Because institutional investors can integrate information more efficiently than retail investors (Miller 2010), retail investors may benefit more from the

[4] I discuss these complexity measures in more detail in Section 5.

[5] Under Blankespoor et al. (2019)'s framework, to process a disclosure the investor needs to (1) be aware of the existence of the disclosure (awareness costs); (2) have access to and extract the information (acquisition costs); and (3) analyze the disclosed information and understand its implications for firm value (integration costs).

mandate because they experience a larger decrease in integration costs, relative to institutional investors. If so, then after the 2004 rule change, information asymmetry between institutional and retail investors immediately following periodic financial report disclosures may be lower. On the other hand, taking contracts out of the financial statements and disclosing them separately on 8-Ks might increase the retail investors' awareness costs if 8-Ks are less salient than 10-K/Qs to these investors. The mandate may also increase information acquisition costs because investors will need to look for multiple filings to access both the material contracts and the financial statements. To the extent that information in material contracts helps investors better understand other financial statement items, separate disclosure of material contracts will then negatively affect retail investors' ability to understand the periodic financial statements. Hence the overall effect of the 2004 rule change on information asymmetry among investors following 10-K/Q disclosures remains an empirical question.

I use a difference-in-differences research design to test my prediction that the 2004 SEC rule change decreased information asymmetry between retail and institutional investors immediately after 10-K/Q disclosures. Although the rule change is applicable to all firms, those that do not rely on external contracting or have always filed material contracts in a timely manner on 8-Ks are less affected by the rule. My control group includes these firms and the treatment group includes all other firms. Following prior literature (e.g., Leuz and Verrecchia 2000), my measure of information asymmetry is

based on the bid-ask spread. Higher information asymmetry means more adverse selection, which implies a larger bid-ask spread (Glosten and Milgrom 1985). I measure information asymmetry as the average bid-ask spread over the three days following 10-K/Q filing dates. Following Bens et al. (2016), I extend the measurement window to 30 days and calculate the natural logarithm of the average bid-ask spread over the calendar month following the 10-K/Q filing dates.

Consistent with the prediction that the SEC mandate reduces information processing cost for retail investors relative to institutional investors, I find that after the 2004 rule change, and relative to the control group, the treatment firms experience about a 5% decrease in the bid-ask spread immediately after 10-K/Q filings. A key assumption in a difference-in-differences study is that absent the treatment shock, the difference between the control and treatment groups is constant over time. In support of this parallel trend assumption, I find no evidence that information asymmetry changes differently for treatment and control firms before the mandate. I further find that in the cross section, firms that file contracts with greater content-complexity before 2004 experience a relatively larger reduction in bid-ask spread. I do not find statistically significant differential effects for firms that file more versus less linguistically complex material contracts.

For the information asymmetry effects to manifest, one critical assumption is that retail investors actually use the information in the material contracts. If not, then

moving material contracts from Form 10-K/Q to Form 8-Kand disclosing earlier would not help these investors understand other information in periodic financial statements. If only institutional investors are able to leverage their superior resources to reduce their information processing cost from this rule change, then information asymmetry between institutional and retail investors may even increase (Blankespoor, Miller, and White 2014). I assume firms that attract more retail investors are more likely to have unsophisticated investors processing materials contracts information. Consistent with this idea, I find that the reduction in information asymmetry after the SEC rule change is greater for firms that draw more retail investor interest.

The effect on information asymmetry should also be larger for firms that significantly increase the timeliness of their material contract filings, because these firms' investors can focus more on the financial statements as they are more likely to have fully processed the contracts before the disclosure of 10-K/Q reports. Consistent with this idea, I find the reduction in the bid-ask spread is significantly larger in the days following 10-K/Q disclosures for firms with greater change in their timeliness of material contract filings.

Finally, I predict that the effect of the 2004 rule change should be more pronounced for financial statements filed during days when investors are busy or distracted by other information releases and therefore more constrained in information processing capabilities. Consistent with this prediction, I find information asymmetry

between sophisticated and retail investors declines more for periodic financial statements filed on Fridays and on days when the Bureau of Labor Statistics releases the Employment Situation report (DellaVigna and Pollett 2009; Damodaran 1989; Graham et al. 2003; Melessa 2012). The effect is also larger for form 10-Ks which are longer and more detailed as compared to 10-Qs.

Collectively, my findings provide evidence of a benefit of the 2004 SEC rule change to file material contracts on Form 8-K within four business days after signing instead of waiting to include these contracts in the next periodic filing, namely, a reduction in information asymmetry among investors following periodic filings. More broadly, my study shows that it is not just the quantity of information disclosed but also the timing of the disclosures that affect information asymmetry among equity investors. This evidence should be of interest to securities regulators as they continue to look for ways to level the information playing field among investors. Prior studies have also examined how managerial discretion over disclosure timing can affect information dissemination. For example, managers can opportunistically time disclosures to hide bad news (e.g., Patell and Wolfson 1982; Penman 1987; deHann, Shevlin, and Thornock 2015) or reduce information overload by spreading out disclosures of good news (Chapman et al. 2019). In these papers, managers strategically choose the timing of disclosures. In contrast, I study a setting in which firms are required to disclose material contracts earlier.

My work also contributes to the accounting literature by introducing two content complexity measures. Traditionally, accounting research has used linguistic complexity measures such as the Fog Index to measure disclosure complexity (e.g., Li 2008). While prior research has provided evidence that content complexity can affect linguistic complexity (Bushee et al. 2018; Lang and Stice-Lawrence 2015), this research has not provided a direct measure of content complexity. I create two novel content-complexity measures applicable to material contracts: the number of payoff-relevant contingencies and whether a contract is a renewal. I use Principal Component Analysis to formally show that linguistic complexity and content complexity are related and distinct concepts.

The rest of the paper is structured as follows. Section 2 discusses the regulatory background and related research and develops the main hypothesis. Section 3 describes the research design and data. Section 4 presents results from my main test. Section 5 discusses my contract complexity measures and presents results from cross-sectional tests. Section 6 concludes.

2. Background and Hypothesis Development

2.1 Institutional Background on Material Contracts Filings

Material contracts are defined by the SEC as material definitive agreements not made in the ordinary course of the registrant's business.[6] Before 2004, firms were required to file material contracts as exhibits in a future 10-K/10-Q report or registration statement per Regulation S-K 17 CFR § 229.601(a)(4): "[i]f a material contract or plan of acquisition, reorganization, arrangement, liquidation or succession is executed or becomes effective during the reporting period reflected by a Form 10-Q or Form 10-K, it shall be filed as an exhibit to the Form 10-Q or Form 10-K filed for the corresponding period." Firms could also voluntarily file contracts immediately after signing via Form 8-K and either reference the 8-K filings or file the contracts again in their 10-K/Q or registration statement. In my sample, from 2001 to 2003, 10.6% of material contracts

[6] Further, the instructions for Form 8-K state that "an agreement is deemed to be not made in the ordinary course of a registrant's business even if the agreement is such as ordinarily accompanies the kind of business conducted by the registrant if it involves the subject matter identified in Item 601(b)(10)(ii) (A) - (D) of Regulation S-K". These subject matters include: (A) Any contract to which directors, officers, promoters, voting trustees, security holders named in the registration statement or report, or underwriters are parties other than contracts involving only the purchase or sale of current assets having a determinable market price, at such market price; (B) Any contract upon which the registrant's business is substantially dependent, as in the case of continuing contracts to sell the major part of registrant's products or services or to purchase the major part of registrant's requirements of goods, services or raw materials or any franchise or license or other agreement to use a patent, formula, trade secret, process or trade name upon which registrant's business depends to a material extent; (C) Any contract calling for the acquisition or sale of any property, plant or equipment for a consideration exceeding 15 percent of such fixed assets of the registrant on a consolidated basis; and (D) Any material lease under which a part of the property described in the registration statement or report is held by the registrant.

were filed via Form 8-K and the rest in periodic filings (Form 10-K/Q). The median calendar days between contract signing and filing is 89 (Table 2).

As part of the Sarbanes-Oxley Act (SOX), Congress requires all public companies to provide information about important corporate events on a real-time basis. Specifically, Section 409 requires that "[e]ach issuer reporting under section 13(a) or 15(d) shall disclose to the public on a rapid and current basis such additional information concerning material changes in the financial condition or operations of the issuer, in plain English, which may include trend and qualitative information and graphic presentations, as the Commission determines, by rule, is necessary or useful for the protection of investors and in the public interest."

In response to SOX, in 2004 the SEC expanded the number of events reportable on Form 8-K.[7] [8] Along with other changes, the 2004 SEC rule change requires firms to announce material contracts via 8-Ks within four business days of signing. The company can file the body of the contract as an exhibit to the announcement 8-K or in

[7] Specifically, it increased the number of reportable events to 22, expanded disclosure requirements on two existing items (departure of directors and amendments to bylaws), and transferred items from the periodic reports. In addition, for most eligible activities, the 2004 8-K expansion requires firms to file 8-K reports within four business days of the occurrence of the triggering events.

[8] The other two items transferred from the periodic reports include unregistered sale of securities and modifications to the rights of security. In contrast to the material contracts, these two items are rarely filed. Noh et al. (2019) report that the number of filings for these two items are less than 7% and 3% of the number of material contracts filings, respectively. Furthermore, material contracts are arguably much more complex than transactions related to these two items. If information is easy to analyze, then changes in disclosure timeliness should have minimal effects on information processing because the cost of analyzing such information is always low.

the next periodic report or registration statement. The SEC originally (in 2002) proposed requiring firms to file material contracts as exhibits to 8-Ks within two business days of contract signing. Firms expressed concerns that they would not always be able to prepare and submit requests for confidential treatment of sensitive contractual terms within two days of contract signing. In response, the SEC altered its proposal to require firms to announce material contracts on 8-Ks within four business days.[9] Nevertheless, the SEC states in the 2004 final rule that they "encourage companies to file the exhibit with the Form 8-K when feasible, particularly when no confidential treatment is requested". Furthermore, the limited safe harbor that grants this option does not protect firms against all disclosure liabilities under the securities laws (Li 2013; Sena 2004). Therefore, firms have good reasons to file the body of the contracts on a current basis in Form 8-K, rather than just announcing them.

According to the SEC, these rule changes are intended to provide investors with "better and faster disclosure of important corporate events", that is, value-relevant information. Prior research shows that material contracts contain value-relevant information. Lerman and Livnat (2010) find that material contract disclosures affect

[9] I assume the 2004 SEC mandate does not change firms' incentives to redact proprietary information from material contracts. If, however, firms are more likely to redact value-relevant information after the rule change, then the rule change can lead to a reduction in information asymmetry because there is simply less information to process in the material contracts.

abnormal trading volume and return volatility around contract filing dates. The 2004 SEC rule change does not change the materiality threshold for 8-K filings, so the content of material contracts filings did not change, only the timing of disclosure.[10] Therefore, this setting is particularly suitable to study how timing of disclosure affects information dissemination in the capital market.

2.2 Prior Studies

My study is related to three areas of prior research. First, it builds on the literature on material contract disclosures which has shown that these contracts are important for corporate actions and contain value-relevant information. In an event study, Lerman and Livnat (2010) demonstrate that both abnormal trading volume and abnormal return volatility are significantly greater than zero in the 3-day period around the filing dates of 8-Ks containing material contracts. Using a large sample of long-term

[10] Although the materiality threshold did not change, contract disclosures could change if the 2004 SEC rule change has real effects on firms' way of writing contracts. In this case, my estimate of the treatment effects would be the combined result of reduced information processing costs and changes in how firms write their contracts. In the post-period, if investors can pay more attention to separately disclosed contracts via 8-K, then it is possible that managers may respond by making contracts more complex. However, contracts are designed to minimize litigation risk while rigorously describing the legal obligations and payoffs of each contracting party. Therefore, firms have incentives to always try to write their contracts in as much detail as possible, regardless of how attentive investors are. Contract complexity may also change if the rule change affects firms' investment decisions. McMullin, Miller, and Twedt (2019) find that the 2004 SEC rule increased the speed of price discovery within the quarter culminating at the earnings announcement. To the extent that more efficient stock prices more accurately reflect all available information, managers may be able to learn more from stock prices and make better-informed decisions. However, it is not clear how investment decisions based on information released at most one quarter earlier translate to uncertainty about future states of the world and complexity of contracts because contract durations can be well over one quarter.

supply contracts, Costello (2013) documents that firms design these contracts to mitigate incentive conflicts in inter-firm relationships. Noh, So, and Webber (2019) find that firms substitute away from management guidance following the 2004 rule change because material contracts contain forward-looking information about firm performance. Relatedly, Li (2013) shows that firms that voluntarily accelerate material contract filings via 8-Ks between 2001 and 2003 are associated with a lower level of information asymmetry among investors in the long run.[11] Altogether, these studies suggest material contracts contain useful information for valuing firms.

Second, my paper is related to how information overload affects investors. Too much information could increase disagreements among investors over how to interpret available information. Psychology research suggests that individuals are likely to use idiosyncratic ways to simplify decision making when the amount of information increases (Einhorn 1971). In addition, prior work finds that information load decreases decision quality even for cognitively sophisticated individuals and financial professionals. For example, Casey (1980) shows that loan officers with the heaviest information load use significantly more time to process the data, but they are no better

[11] While Li (2013) focuses on the long-run information asymmetry effects during the year of voluntarily accelerating filings of material contracts, my paper studies the effects of the 2004 requirement on information asymmetry in the days shortly following periodic financial statement disclosures.

in predicting bankruptcy. Similarly, Malhotra (1982) observes that consumers make worse product choices as information load increases.

Information overload can also affect the information content of share prices. Bloomfield (2002) argues that when information is too complex to interpret, it is not fully incorporated in stock prices. Theory explains that investors with limited attention can be rationally inattentive to complex information due to the large information processing capacity required to process such information, resulting in under-reaction (Hirshleifer and Teoh 2003). For example, Peng and Xiong (2006) show that information overload leads to category-learning behavior, that is, resource-constrained investors rely less on firm-specific information and more on market- and industry-wide information. Similarly, DellaVigna and Pollet (2009) theorize that investors neglect information due to limited processing capabilities. DellaVigna and Pollet (2007) find that higher attention cost induces investors to simplify complex decisions by neglecting long-term information. Consistent with these theoretical implications, Hirshleifer, Lim, and Teoh (2009) demonstrate that stock price responses to earnings news are delayed on days with more earnings announcements. Cohen and Lou (2012) find that given investors' limited information processing capacity, information complexity can lead to a significant delay in the incorporation of positive industry-level information into asset prices.

Lastly, my paper is related to how timing of disclosure affects information dissemination. Patell and Wolfson (1982) document that earnings news released after

trading hours and/or on Fridays is on average worse because managers realize investors pay less attention to news during these times. This finding has been confirmed in different settings by, for example, Penman (1987), DellaVigna and Pollett (2009), deHann et al. (2015), and Segal and Segal (2016). Chapman, Reiter, White and Williams (2019) show managers try to reduce information overload by spreading out disclosures. This spreading out is associated with increased liquidity, reduced price volatility and higher analyst forecast accuracy. In contrast to settings in which managers strategically choose the timing of disclosure, I study the impact of a mandated change in disclosure timing on investor information processing.

2.3 Hypothesis Development

A growing literature in accounting and finance shows that information overload negatively affects information dissemination in the capital markets. The SEC has voiced concerns about this problem and has advocated for disclosures that are most relevant to stakeholders (e.g., Higgins 2014). As previously described, information overload can cause financial statement users to adopt simplifying decision strategies and ignore certain information signals (e.g., Hirshleifer and Teoh 2003; You and Zhang 2009). The rational inattention literature explains that simplified decision strategies can be optimal given users' limited information processing capacity (e.g., Sims 2010; Maćkowiak et al.

2018). Information overload can also increase information asymmetry among investors with differing information capacities (Grossman and Stiglitz 1980; Ball 1992).

Material contracts, which can be complex as they describe the contracting parties' legal rights and obligations under different future states of the world, contribute to information overload for financial statement users when the contracts are disclosed in periodic filings (10-K/Q reports). Because of clustering in fiscal quarter end dates, this information overload problem is exacerbated during the earnings season when many firms disclose their financial statements around the same time. During this time, there is a high rate of information arrival, to the extent that one analyst described the earnings announcement season as a time "when you are drinking [information] from a fire hose".[12] Investors will then have to process large amounts of information as fast as possible if they do not want to be left at an information disadvantage.

As previously described, Blankespoor et al. (2019) establish a conceptual framework for analyzing three types of information processing costs: awareness costs, acquisition costs, and integration costs. Before the 2004 SEC rule change, investors have to process information from both material contracts and the financial statements during days immediately following 10-K/Q disclosures, resulting in high integration costs. After the rule change, material contracts are disclosed on form 8-K within four business

[12] https://www.lifeonthebuyside.com/earnings-season/

days of the contract signing date. Because contracts are disclosed earlier, in a more staggered fashion, all investors have more time to integrate the information in the 10-K/Qs and the material contracts. In other words, the 2004 rule change reduced the information integration cost of 10-K/Qs. This change may benefit retail investors more than sophisticated investors, as retail investors are more constrained in information processing capacity (Barron, Byard, and Enis 2004; Miller 2010). This reasoning suggests information asymmetry between sophisticated and retail investors immediately after periodic financial report filing dates is lower after the 2004 SEC mandate. This is my main hypothesis:

H1*: The 2004 SEC mandate lowers information asymmetry between sophisticated and unsophisticated investors in the short window following periodic financial statement disclosures.*

However, the 2004 SEC mandate could increase information asymmetry among investors. If information in material contracts helps investors better understand other items in the financial statements, then after the 2004 SEC mandate investors must look for multiple filings to get both material contracts information and financial statement information.[13] Under Blankespoor et al. (2019)'s framework, the search for multiple

[13] For example, material supply contracts might represent a significant portion of a filer's revenues or purchases (Costello 2013). Moreover, contracts may include information about future revenues or cost of goods sold (Noh et al. 2019), which can be correlated with other forward-looking information in periodic financial statements.

filings would increase investors' information acquisition cost. Psychology research has demonstrated limitations in humans' short-term working memory (e.g., Miller 1956; Cowan 2017). Further, these limitations can negatively affect information processing abilities (Sweller 1988).[14] To the extent 8-Ks are less salient than 10-K/Qs to investors, the mandate can increase their awareness cost as well. Compared to sophisticated investors, retail investors are more likely to suffer from these increased acquisition and awareness costs. Hence the effect of the 2004 rule change on information asymmetry following 10-K/Q disclosures remains an empirical question.

[14] Consistent with these ideas, prior literature shows that when related information is provided in the same conceptual categories, individuals are able to acquire and comprehend the information better (e.g., Bruner, Goodnow, and Austin 1956; Kozminsky 1977). Similar findings have been documented in the capital market context: Bloomfield, Hodge, Hopkins, and Rennekamp (2015) argue that grouping related information in the same section of a financial disclosure can reduce the cognitive load necessary for integrating the related information. They present experimental evidence that such coordinated presentation helps improve credit analysts' ability to identify and acquire information.

3. Research Design and Sample

3.1 Research Design

To evaluate the impact of the 2004 SEC mandate on information asymmetry among investors following periodic financial statements disclosures, I use a difference-in-differences research design. Specifically, I compare the changes in post-periodic-filing information asymmetry after the 2004 SEC mandate for treatment and control firms using the following regression model with firm index *i* and quarter index *t*:

$$DV_{it} = \beta_0 + \beta_1 Post_t * Treat_i + Controls_{it} + \delta_i + \sigma_t + \epsilon_{it}. \qquad (1)$$

In equation (1), *Post* is an indicator variable that equals one for all firm-quarters with periodic financial statement (10-K/Q) filing dates on or after the effective date of the SEC mandate, and zero otherwise. *Treat* is an indicator variable that equals one for all treatment firms. The 2004 SEC mandate is applicable to all public firms, but two categories of firms are less affected by the rule change: (1) firms that do not rely much on external contracting, and (2) firms that had voluntarily filed material contracts via timely 8-Ks before the mandate became effective. My control group includes these two groups of firms, and the treatment group includes all other firms. I define category (1) control firms as those that did not file any material contracts in the pre-mandate period (2001 to 2003). I define category (2) control firms in two ways. First, I include firms that filed at least one material contract via 8-K in the pre period, and the average filing lag of these pre-period 8-K material contracts is below the pre-period sample median; the

corresponding treatment group is *Treat1*. Second, I relax the timeliness condition and include firms in category (2) if they filed at least one material contract via 8-K in the pre-period; the corresponding treatment group is *Treat2*. In robustness tests, I exclude all category (2) firms from my control group and include them in the treatment group, and generate two more treatment variables, *Treat3* and *Treat4*. *Treat3,* which captures the extent to which a firm was affected by the 2004 SEC mandate, is the average number of material contract filings in the pre-period. *Treat4* is an indicator that equals one when a firm filed at least one material contract in the pre-period, and zero otherwise.[15]

Conceptually, the dependent variable *DV* in equation (1) is information asymmetry. Following prior literature, I measure information asymmetry using a proxy (*Spread*) based on the bid-ask spread (e.g., Glosten and Milgrom 1985; Leuz and Verrecchia 2000; Bens, Cheng, and Neamtiu 2016). As Glosten and Milgrom (1985) explain, a market maker faces a possible adverse selection problem when an investor agrees to trade at the bid or ask price because the investor may be trading on information the market maker does not know. By setting a spread, the market maker can offset losses suffered in trades with informed investors with gains in liquidity

[15] Voluntary filers are not always consistent in their disclosure practices; i.e., they sometimes file contracts via form 8-K and sometimes via form 10-K/Q. As a result, the analyses in my paper can underestimate the magnitude of the effect of the 2004 SEC rule. For example, a firm that files one contract via 8-K and many contracts via 10-K/Qs in the pre-period will be included in the control group, although the rule change should affect information asymmetry among investors following 10-K/Q disclosures.

trades. Higher information asymmetry means more adverse selection, which implies a larger bid-ask spread. To measure short-term information asymmetry, I use the natural logarithm of the average 3-day bid-ask spread after periodic report filing dates. In additional tests, I show my results are robust to using PIN as an alternative measure of information asymmetry. Following Bens et al. (2016), I also extend the measurement window to 30 days and calculate the natural logarithm of the average bid-ask spread over the calendar month following the 10-K/Q filing dates.

I include the following firm-quarter level control variables following Bens et al. (2016): return on assets, leverage, book-to-market ratio, size, average stock price, average trading volume, analyst following, and previous return volatility. I follow Steffen (2020) and include the average value of the CBOE Volatility Index to control for expectations of future stock market volatility. To control for other time-invariant firm characteristics that may bias my results, I include firm fixed effects in all specifications. The quarter fixed effects in my regressions control for market-wide fluctuations during the sample period. My difference-in-difference regressions do not include the standard *Treat* and *Post* indicators because I include firm and time fixed effects. The coefficient of interest, β_1, carries a difference-in-difference interpretation. If the 2004 SEC mandate reduces information asymmetry following 10-K/Q disclosures more for the treatment firms relative to the control firms, $\beta_1$will be negative.

3.2 Sample Selection

Table 1 details the sample selection. I start with the universe of firms with WRDS SEC Analytics Suite filings 10-K, 10-KT, 10-Q, 10-QT, 10-KSB, or 10-QSB. Following Lerman and Livnat (2010) and Noh et al. (2019), I use a three-year window after the 2004 SEC mandate, paired with a three-year window before the mandate. I define the three-year window before (after) the August 23, 2004 effective date as the pre-period (post-period). Following Noh et al. (2019), I drop 10-K/Qs filed in quarters surrounding the effective date of the mandate (i.e., calendar year 2004) from my sample. I collect filing dates of periodic financial statements between 2001 and 2007 from the WRDS SEC Analytics Suite and merge these filing dates with Compustat and CRSP data. My final sample includes 4,817 firms and 73,883 firm-quarters.

Table 1: Sample Selection

1. Firms with WRDS SEC Analytics Suite filings 10-K, 10-KT, 10-Q, 10-QT, 10-KSB, or 10-QSB with the following characteristics: (1) filed between January 1, 2001 and December 31, 2007; (2) non-missing GVKEY, fiscal period end date (RDATE), and filing date (FDATE); (3) FDATE after RDATE; and (4) no duplicates on GVKEY-RDATE	10,341
2. Firms without CRSP PERMNO	(3,363)
3. Firms without available regression variables	(2,051)
4. Singleton firms (firms that only have one quarter of data)	(110)
Total number of firms in the sample	4,817

To analyze material contract filings, I use a Python script to download Exhibit 10s from all 8-K forms and 10-K/Q forms in my sample period. Using Exhibit 10s to identify material contracts is consistent with prior literature (WRDS 2011; Li 2013; Noh et al. 2019). Exhibit 10 tags are automatically added to the top of every contract by the SEC EDGAR system whenever a firm files one or more material contracts. After downloading Exhibit 10s, I use a second script to extract contract-level information such as signing date, contract type and indicators of contract complexity. I discuss this contract-level information in greater detail in Section 5.1.

3.3 Sample Descriptive Statistics

Table 2, Panel A presents descriptive statistics for the variables used in estimating equation (1). Detailed description of these variables is in Appendix A. I winsorize all continuous variables in the table at 1% and 99%. These descriptive statistics are similar to those in prior studies (e.g., Steffen 2020; Bens et al. 2016).

Table 2, Panel B describes the number and timeliness of material contracts filings via 8-K and 10-K/Q. From the pre-period to the post-period, the number of material contracts filed via 8-Ks increased from 3,453 to 20,324, and the number of contracts filed via 10-K/Q declined from 29,602 to 23,263, suggesting a meaningful change in how firms

file material contracts.[16] The 2004 SEC rule change also increased the timeliness of material contract filings: median calendar days between contract signing and filing dropped from 89 to 38, allowing investors substantially more time to process the material contracts after the rule change.[17]

Table 2: Summary Statistics

Panel A: Regression Variables

	(1) *N*	(2) *Mean*	(3) *Std. Dev.*	(4) *5%*	(5) *25%*	(6) *Median*	(7) *75%*	(8) *95%*
Spread	73,883	-5.308	1.469	-7.651	-6.459	-5.299	-4.201	-2.877
ROA	73,883	-0.016	0.077	-0.167	-0.019	0.007	0.020	0.047
Leverage	73,883	0.459	0.259	0.098	0.254	0.440	0.618	0.897
BTM	73,883	0.605	0.583	0.062	0.257	0.455	0.766	1.735
Size	73,883	5.585	1.940	2.551	4.168	5.474	6.878	9.075
Price	73,883	17.757	17.819	1.095	4.336	11.750	25.513	54.833
Turnover	73,883	-5.500	1.290	-7.908	-6.286	-5.352	-4.590	-3.596
Follow	73,883	1.284	1.019	0.000	0.000	1.386	2.079	2.944
Avgvix	73,883	19.771	6.571	11.087	13.593	18.993	25.030	31.040
Paststdret	73,883	0.030	0.029	0.003	0.011	0.022	0.038	0.089

[16] As described in Section 2.1., the 2004 SEC rule change allows a firm to summarize the material contract in an announcement 8-K within four business days of signing and disclose the entire body of the material contract in the corresponding 10-K/Q. In these cases, contracts are categorized as filed via 10-K/Q.

[17] I use textual analysis to determine a contract's signing date. Specifically, I use a waterfall approach to search for dates in the following places: 1) title; 2) HTML description; 3) first 500 words of the document; and 4) the last 500 words in the document.

Panel B: Material Contract Filings

		(1) Pre	(2) Post	(3) Post - Pre
via 8-K	*N*	3,453	20,324	16,871
	File_Delay (Median)	10	5	-5***
via 10-K/Q	*N*	29,062	23,263	(5,799)
	File_Delay (Median)	98	84	-14***
All	*N*	32,515	43,587	11,072
	File_Delay (Median)	89	38	-51***

Panel A of this table presents the summary statistics for the regression variables used to estimate equation (1). All continuous variables are winsorized at 1% and 99%. Panel B describes the number and timeliness of material contracts filed via 8-K and 10-K/Q. *File_Delay* is calculated as the number of days between contract signing and filing. Detailed variable definitions are provided in Appendix A.

4. Main Results

Table 3 presents the main results on the effects of the 2004 SEC rule change on information asymmetry among investors. In Panel A, I begin by examining the effect of the rule change on information asymmetry by estimating equation (1) using the natural logarithm of the average daily bid-ask spread in the 3-day window following the 10-Q or 10-K filing date as the dependent variable. Column (1) shows the estimates from the specification using *Treat1*; i.e., the treatment group is firms that filed at least one material contract via 8-K or 10-K/Q in the pre period but either never used 8-K to file contracts or the average filing lag of any pre-period 8-K material contracts exceeds the sample median; all other firms are the control group. The coefficient on *Post* × *Treat*, which carries the difference-in-difference interpretation, is negative and significant at the 1% level (coefficient estimate = -0.0585). In terms of economic magnitude, this result suggests that the 2004 SEC rule change is associated with a 5.85% reduction in information asymmetry among investors in the period immediately following 10-K/Q filings.

Table 3: Effects of the SEC 2004 Requirement on Information Asymmetry

Panel A

	(1)	(2)	(3)	(4)
Treatment Variable	*Treat1*	*Treat2*	*Treat3*	*Treat4*
Window	3 days	3 days	3 days	3 days
Post × Treat	**-0.0585*** **	**-0.0463** **	**-0.0454*** **	**-0.1132*** **
	(-2.7153)	**(-2.3627)**	**(-4.1770)**	**(-4.0623)**
ROA	-0.4000***	-0.4014***	-0.3911***	-0.3936***
	(-6.9886)	(-7.0030)	(-6.8124)	(-6.8946)
Leverage	0.5167***	0.5174***	0.5147***	0.5172***
	(15.3626)	(15.3878)	(15.2560)	(15.3588)
BTM	0.2974***	0.2975***	0.2967***	0.2977***
	(27.0408)	(27.0216)	(26.9848)	(27.0947)
Size	-0.2826***	-0.2827***	-0.2832***	-0.2828***
	(-21.0073)	(-20.9554)	(-21.0818)	(-21.0782)
Price	-0.0140***	-0.0140***	-0.0140***	-0.0140***
	(-21.7130)	(-21.7179)	(-21.7573)	(-21.7557)
Turnover	-0.1702***	-0.1703***	-0.1705***	-0.1706***
	(-44.8394)	(-44.8711)	(-44.8904)	(-44.9572)
Follow	-0.1829***	-0.1832***	-0.1816***	-0.1816***
	(-18.6393)	(-18.6767)	(-18.5813)	(-18.5202)
Avgvix	0.0141***	0.0141***	0.0142***	0.0143***
	(11.7266)	(11.6973)	(11.8557)	(11.8427)
Paststdret	1.7910***	1.7920***	1.7862***	1.7842***
	(17.8185)	(17.8152)	(17.7560)	(17.7630)
Observations	73,883	73,883	73,883	73,883
Adjusted R-squared	0.8074	0.8074	0.8075	0.8075
Firm FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes

Panel B

	(1)	(2)	(3)	(4)
Treatment Variable	*Treat1*	*Treat2*	*Treat3*	*Treat4*
Window	30 days	30 days	30 days	30 days
Post × Treat	**-0.0737*** **	**-0.0556*** **	**-0.0509*** **	**-0.1387*** **
	(-3.9046)	**(-3.2507)**	**(-5.3656)**	**(-5.7113)**
ROA	-0.4229***	-0.4247***	-0.4145***	-0.4155***
	(-9.0318)	(-9.0506)	(-8.8208)	(-8.8863)
Leverage	0.4646***	0.4653***	0.4630***	0.4654***
	(15.7412)	(15.7685)	(15.6323)	(15.7634)
BTM	0.2416***	0.2416***	0.2411***	0.2420***
	(25.0183)	(24.9767)	(24.9561)	(25.1016)
Size	-0.2545***	-0.2544***	-0.2549***	-0.2547***
	(-22.3831)	(-22.3460)	(-22.3807)	(-22.4530)
Price	-0.0127***	-0.0127***	-0.0128***	-0.0127***
	(-22.0517)	(-22.0856)	(-22.1138)	(-22.0931)
Turnover	-0.2575***	-0.2574***	-0.2577***	-0.2581***
	(-58.1798)	(-58.1754)	(-58.2939)	(-58.4463)
Follow	-0.1696***	-0.1700***	-0.1682***	-0.1680***
	(-21.6597)	(-21.7125)	(-21.5550)	(-21.4626)
Avgvix	0.0167***	0.0166***	0.0166***	0.0167***
	(16.4068)	(16.3695)	(16.4045)	(16.5180)
Paststdret	3.3421***	3.3475***	3.2898***	3.3172***
	(21.2686)	(21.2431)	(20.9881)	(21.1807)
Observations	73,883	73,883	73,883	73,883
Adjusted R-squared	0.9087	0.9086	0.9088	0.9088
Firm FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes

This table reports OLS regression results on information asymmetry. The dependent variable is the natural logarithm of the average daily bid-ask spread during the measurement window following the 10-Q or 10-K filing date. In Panel A, the measurement window is 3 days while in Panel B it is 30 days. *Price*, *Turnover*, and *Avgvix* are calculated over the measurement window following the 10-Q/K filing date. *Paststdret* is calculated over the measurement window ending one day before the 10-Q/K filing date. From Column (1) to (4) of each panel, I use different ways to define the treatment and control groups. All variables are defined in Appendix A. I include quarter and firm fixed effects in all specifications.

Standard errors are clustered at the firm level. T-statistics are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Column (2) uses *Treat2* as the treatment indicator. This treatment group includes firms that filed at least one material contract via 10-K/Q but never used 8-K in the pre-period, and the control group includes all other firms. The coefficient on *Post* × *Treat* is again negative and significant at the 1% level (coefficient = -0.0463). For robustness, I report in Columns (3) and (4) the results from specifications using *Treat3,* based on the total number of contracts filed in the pre-period, and *Treat4*, an indicator that equals one for firms that filed at least one material contract via 8-K or 10-K/Q in the pre-period, as alternative treatment variables. Inferences are similar to the those from the first two columns. In Panel B, I follow Bens et al. (2016) and use a 30-day window following 10-K/Q disclosures. Results are qualitatively similar to those in Panel A. Altogether, the results in Table 3 are consistent with the hypothesis that the 2004 SEC rule change reduces information asymmetry in the short window following periodic financial filings and thereby helps level the information playing field between sophisticated and unsophisticated investors.

Underlying a difference-in-differences research design is the parallel trends assumption, which states that absent any treatment shock, the difference in the variable of interest between the control and treatment groups is constant over time. To test this assumption, I first plot the time-series of my dependent variable separately for treatment

and control firms and visually check for parallel trends before 2004. Figure 1, Panel A presents the means of my dependent variable *Spread*, the natural logarithm of the average daily bid-ask spread in the 3-day window following the 10-Q or 10-K filing dates by filing quarter when I use *Treat1* as the treatment group. There is no visual evidence that the parallel trends assumption is violated in the pre-period. In Panel B, I use *Treat2* as the treatment group and the findings are similar. In unreported figures, I find similar trends when I plot the medians of *Spread*. The inferences also do not change when I focus on the 30-day window following 10-K/Q filings.

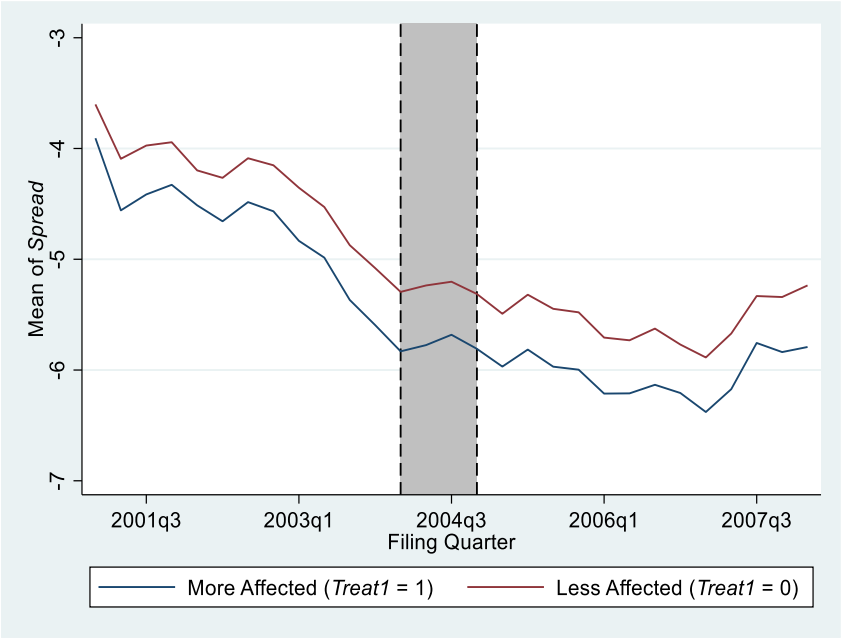
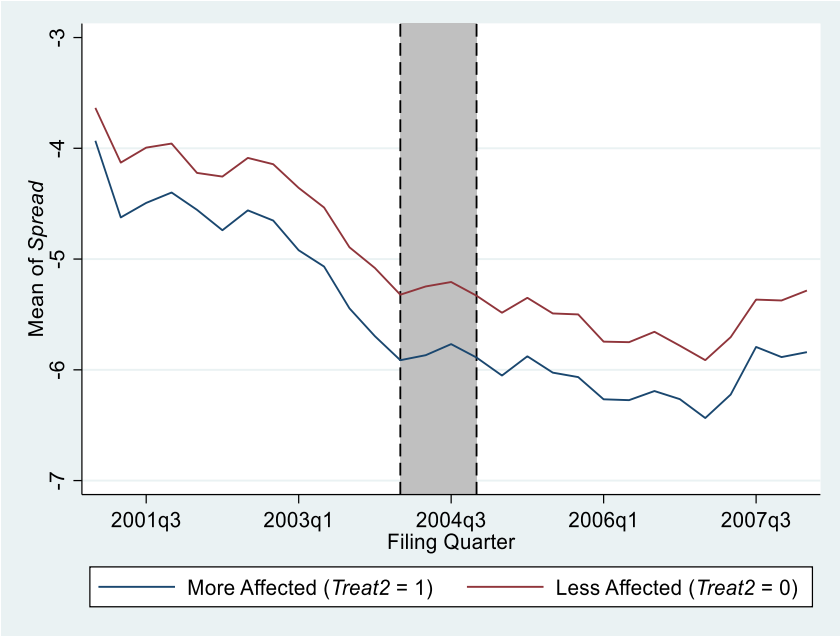


Figure 1: Time Trends of *Spread*

Panel A



Panel B

This figure plots the time series means of *Spread* by firms' treatment status. In Panel A, the treatment variable is *Treat1* and in Panel B it is *Treat2*. *Spread* is measured in the 3 days following periodic financial statement disclosures. All variables are defined in Appendix A. The 2004 SEC rule change was first proposed on June 17, 2002. It was passed on March 16, 2004 and became effective on August 23, 2004. The grey shaded area represents the year of 2004 which is excluded from my regression analyses.

Second, I augment equation (1) with two interaction terms *Before (-1)* × *Treat* and *Before (-2)* × *Treat*, where *Before (-1)* and *Before (-2)* are quarter indicator variables that equal one for the calendar quarter period one quarter (two quarters) before January 2004. Table 4 presents these results. The coefficients for both *Before (-1)* × *Treat* and *Before (-2)* × *Treat* are statistically insignificant at conventional levels with the exception of *Before (-1)* × *Treat* in one specification in Column (4). These results provide support for

the validity of the parallel trends assumption.[18] The coefficients on *Post* × *Treat* continue to be negative and significant as in Table 3.

Table 4: Testing the Parallel Trends Assumption

	(1)	(2)	(3)	(4)
Treatment Variable	*Treat1*	*Treat2*	*Treat1*	*Treat2*
Window	3 days	3 days	30 days	30 days
Before(-2) × Treat	**0.0195**	**0.0352**	**-0.0083**	**-0.0070**
	(0.6546)	**(1.2645)**	**(-0.4437)**	**(-0.4121)**
Before(-1) × Treat	**0.0069**	**-0.0175**	**-0.0234**	**-0.0522*** **
	(0.2359)	**(-0.6429)**	**(-1.2103)**	**(-2.9273)**
Post × Treat	**-0.0764*** **	**-0.0660*** **	**-0.0993*** **	**-0.0849*** **
	(-3.2108)	**(-3.0527)**	**(-4.7773)**	**(-4.5024)**
Observations	73,883	73,883	73,883	73,883
Adjusted R-squared	0.8075	0.8075	0.9088	0.9089
Controls	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes

This table reports OLS regression results on information asymmetry. The dependent variable is the natural logarithm of the average daily bid-ask spread during the measurement window following the 10-Q or 10-K filing date. In Columns (1) and (2), the measurement window is 3 days while in Columns (3) and (4) it is 30 days. *Before(-1)* (*Before(-2)*) is an indicator variable that equals one for one quarter (two quarters) before 2004. All variables are defined in Appendix A. I include quarter and firm fixed effects in all specifications. Standard errors are clustered at the firm level. T-statistics are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

[18] The significant coefficient on *Before (-1) × Treat* in Table 4, Column 4 is consistent with an early response to the rule change. Firms that start to announce contracts on 8-Ks before the effective date of the rule change without disclosing the entire body of the contracts are not included in the control group because they did not file contracts in the pre-period. To the extent contract announcements preempt some information in the body of the contracts, the coefficient on the interaction term in Table 4, Column 4 can be negative.

5. Additional Analyses

5.1 Cross-sectional Variation in Contract Complexity

Material contracts vary in complexity and therefore also in the cognitive burden they impose on investors. Because investors have limited information processing capacity, it is more difficult for them to process complex contracts. However, it is likely that even retail investors can process simple contracts, even when they are disclosed with the financial statements. Therefore, earlier separate disclosure of simple contracts on 8-Ks should not affect information asymmetry among investors. That is, the effect of the 2004 rule change on information asymmetry should be stronger for firms that file more complex contracts. I next discuss contract complexity and how I measure this construct.

5.1.1 Nature of Contract Complexity

The complexity of a material contract, or the cognitive load required to understand a contract, could arise from any of three sources alone or in combination. First, the legal literature suggests that contract complexity increases with the number of payoff-relevant contingencies. As Eggleston, Posner, and Zeckhauser (2000) point out, "evaluating the likelihood of different contingencies arising…requires mental effort". The reasoning is as follows. A contract will typically specify the contracting parties' obligations and entitlements under each future state of the world. Under each state, each contracting party is likely to face different incentives and constraints which can

lead to different actions that in turn affect the eventual payoff, the difference between the value of the entitlements and the obligations, under each scenario. Therefore, to evaluate a contract, investors need to understand different payout scenarios and assign probabilities to the occurrence of each scenario. The expected payoff will be a function of the number of contingencies and their respective probabilities. Thus, a complex contract has many state contingencies in the sense that its terms specify different payoffs to a large number of future states of the world.

The second source of complexity pertains to whether the contract is a renewal or a new arrangement. Conceptually, contract renewals impose lower cognitive burden because the arrangements are familiar and the contracts are similar to existing contracts. New contracts, in contrast, present unfamiliar terms and contingencies that must be processed. Therefore, it is easier to process renewals. The first and second sources of contract complexity relate to the content of the material contracts, so I label them content complexity to distinguish them from linguistic complexity, discussed next.

The third source of contract complexity is linguistic complexity. A contract is linguistically complex if its language imposes a significant cognitive load on its readers. Prior research suggests linguistic complexity affects readers' understanding of the information disclosed. For example, Li (2008) and Bushee et al. (2018) suggest managers have incentives to obfuscate bad news, that is, to make bad news disclosures more difficult to read, to increase investors' information processing costs and delay the market

reaction. Analysts make less accurate forecasts and disagree with each other more when a disclosure is more linguistically complex (Lehavy, Li, and Merkley 2011). Moreover, regulators are also concerned that investors may not be able to understand complex language. In 1998, the SEC issued plain English disclosure guidelines to encourage the use of plain English in corporate filings. Arthur Levitt, then-Chairman of the SEC, urged firms to "speak to investors in words they can understand" in his foreword to *A plain English Handbook*. Collectively, the evidence suggests that a difficult-to-read contract imposes higher processing costs.

5.1.2 Measures of Complexity in the Accounting Literature

Early textual analysis studies in accounting used the Fog Index to measure linguistic complexity or readability (e.g., Li 2008). Developed by Robert Gunning in the 1940's, the Fog Index is originally from the computational linguistics literature. According to Gunning, the index is designed to help professional and non-professional writers improve the clarity of their writing in popular press and memos, letters, and government and industry reports (Gunning 1969). It proposes that more syllables per word and more words per sentence make a document harder to read.[19] Since its introduction, the measure has become one of the most popular readability measures

[19] The Fog Index is calculated as 0.4 × (average sentence length + percentage of complex words), where complex words are defined as words with three or more syllables.

primarily thanks to its ease of calculation (Loughran and McDonald 2014). It has also been widely used in finance and accounting research (e.g., Li 2008; Biddle, Hillary, and Verdi 2009; Miller 2010; Lehavy et al. 2011; Loughran and McDonald 2014; Guay, Samuels, and Taylor 2016).[20] Interestingly, despite its wide usage, there is little evidence documenting the advantage of the Fog Index over other syllable-based linguistic measures such as the Flesch Reading Ease score and the SMOG grade, or non-syllable-based measures such as the Dale-Chall readability score.[21]

However, the applicability of the Fog Index to business documents, including for example financial reports and material contracts, is questionable. Jones and Shoemaker (1994) point out that the measure is developed using non-technical material and argue it is a "dubious instrument" for adequately assessing the readability of narratives that are adult oriented and specialist in nature. More recently, Loughran and McDonald (2014) demonstrate that the Fog Index and any average syllable-count based measures can be a poor measure of readability for business documents for two reasons. First, none of the most frequently occurring "complex" words in financial statements by the standard of these measures, such as "financial, company, operations, management, employees, and

---

[20] See Loughran and McDonald (2016) for a review.

[21] The Flesch Reading Ease score is calculated as 206.835 - 1.015 × average sentence length - 84.6 × average syllable count in a word. The SMOG grade is calculated as 1.043 × sqrt(30 × number of polysyllables/number of sentences) + 3.1291. The Dale-Chall readability score is calculated as 15.79 × percentage of difficult words + 0.0496 × average sentence length, where difficult words are defined as those not included in the Dale-Chall Word List.

customers", would cause investors difficulty. A second problem is that sentence length can be hard to determine, because abbreviations, section headings and long lists can cause trouble for parsing algorithms. The authors further show that the association between the Fog Index and analyst forecast dispersion during 1995 to 2006 as documented in Lehavy et al. (2011) does not persist when the sample period is expanded to 1994 to 2011.

In the context of material contracts, these same issues arise as well. Because the Fog Index and similar measures are not designed for contracts, it is not surprising that the most common "complex" words by the standard of these readability measures are not expected to cause investors trouble understanding the contract. In addition, the structure of contracts (i.e., section headings, lists) can make parsing algorithms less accurate in determining sentence length. Therefore, while I include these measures, I do not make a prediction as to how linguistic complexity as measured by these proxies would be associated with the information asymmetry effects that I document in my setting.

Other commonly used measures for readability include number of words in a document (e.g., Li 2008) and size of the submission file (e.g., Loughran and McDonald 2014), based on the assumption that lengthy documents are more difficult to read. Document length and file size can also be interpreted as measures of content complexity if longer texts are required to deliver more information (Bushee et al. 2018; Lang and

Stice-Lawrence 2015). Bushee et al. (2018) attempt to separate information content from linguistic complexity in the context of earnings conference calls, but their approach is not applicable in the material contracts setting.[22] I include document length, file size, and number of sentences in my contract-level complexity measures and note that these measures could be correlated with both linguistic complexity and content complexity.

5.1.3 My Measures of Content Complexity

As previously explained, linguistic complexity is not the only type of complexity that can impose a substantial cognitive burden on investors in processing documents. Importantly, because contracts are designed to minimize litigation risk while rigorously describing the legal obligations and payoffs of each party to the contract, ease of comprehension is simply not as important as, for example, in articles in the popular press or internal corporate memorandums. While some linguistic complexity measures have been found to be associated with underlying information content, I could not identify direct measures of content complexity in previous research. Thus, I create two direct measures of content complexity directly applicable to material contracts.

The first aspect of content complexity is the number of payoff-relevant contingencies, as described in Section 5.1.1. Following this intuition, I use the number of

[22] Specifically, Bushee et al. (2018) use the linguistic features of questions from the analysts as a benchmark for the information component in managers' responses. There is no question-and-response component of material contracts.

conditional statements in the contract as a proxy for payoff-relevant contingencies. Specifically, I count the number of words considered synonyms for "if". I grab this list of synonyms from thesaurus.com: *assuming that, granted that, on the assumption that, supposing that, wherever, conceding that, in case that, on the occasion that, whenever, with the condition that, contingent upon, in case, with the proviso, given, in the case that, on these terms, in the event, subject to, on condition, supposing, in case*. In addition to the raw count, for each contract I scale the number of conditional statements by the number of words to calculate the percentage of conditional words.

Appendix B displays the first section of a material contract between U.S. Home Corporation and St. Charles Community, LLC. My algorithm identifies four conditional statements in the section. Each statement describes either conditions allowing parties to terminate the contract or detailed requirements on a key instrument (the "Lot Contract") which can affect the probability of termination. To assess the likelihood the contract will be terminated by either party, investors need to understand the probability of each scenario. Therefore, to evaluate the implications of this contract for the contracting parties, investors will need to expend significant effort to collect information about the arrangement. According to the Fog Index, complex words in this section include "intention", "undertake", "liability", "company", and "terminate", none of which is linguistically challenging for investors.

The second aspect of content complexity I identify in Section 5.1.1 is whether a contract is a renewal of an existing commercial arrangement or a new contract. Because investors are familiar with the content of a renewal contract, renewals should have lower content complexity than new contracts and impose less cognitive burden. To identify a contract renewal, I create a list of renewal indicators and count the occurrence of these indicators in a contract. The renewal indicators are: *renew*, *renewal*, *extend*, *continue*, and *resume*. As before, I calculate the percentage of renewal words by scaling the raw count by the total number of words in a material contract.

5.1.4 Firm-quarter Level Measures of Contract Complexity

I collect the following contract-level measures of contract complexity: file size, number of words, number of sentences, number of conditional statements, and number of renewal indicators, percentage of conditional statements and renewal indicators, the Fog Index, the Flesch Reading Ease score, the SMOG grade, and the Dale-Chall readability score. Because the unit of observation in my main test is a firm-quarter, I aggregate the contract-level complexity measures of all the contracts disclosed by a firm in a given quarter to create my firm-quarter level measures of contract complexity.

Specifically, for each of the following contract-level measures, I calculate the sum across all contracts disclosed in the quarter: file size and number of words, sentences, conditional statements, and renewal indicators. For each of the remaining contract-level

measures, I calculate the average across all contracts disclosed by the firm in the quarter. I do not calculate the sum for these measures because the sum does not have a clear interpretation. For example, the cognitive burden to process two documents of Fog Index 9 is not the same as the cognitive burden to process one document of Fog Index 18. I also include a simple count of all material contracts disclosed in a given quarter because all else equal, the number of contracts can be directly linked to the amount of information that investors need to analyze.

As explained in Section 5.1.1, linguistic complexity and content complexity are distinct but correlated. For example, more linguistically complex words and lengthy sentences can be necessary to convey complex content. Consistent with this idea, prior research has shown that when managers try to convey inherently complex information rather than to obfuscate bad news, linguistic complexity increases (Bushee et al. 2018). Lang and Stice-Lawrence (2015) find both the Fog Index and the number of words increase following IFRS adoption which increased discussion of technical topics in financial reports.

In Table 5, Panel A, I present the correlation matrix for the complexity measures. Pearson correlation coefficients are shown in the lower triangle, while Spearman rank correlations appear above the diagonal. As expected, most of these measures are positively correlated. Panel B displays sample summary statistics for the firm-quarter level complexity measures. On average, firms file 1.03 contracts per quarter. The

average number of words from all contract filings is 7615.9. The average number of conditional statements is 61.2, and the average number of renewal indicators is 12.2. Compared to prior research that studies annual report readability (e.g., Li 2008), material contracts have much larger median Fog Index (34.5 vs. 19.2). According to the standard interpretation, any document with a Fog Index exceeding 18 is "unreadable" (Li 2008). As with other traditional readability measures created to benchmark popular press articles and grade school textbooks, the Fog Index estimates the number of years of education needed to understand the text on a first reading (Loughran and McDonald 2014). Therefore, it is not surprising to see high Fog Index levels in material contracts, which are designed for purposes other than ease of communication.

Table 5: Contract Complexity Measures

Panel A: Correlation Matrix

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Ncontracts (1)	-	0.5550	0.5528	0.5507	0.5448	0.5066	-0.0605	0.0893	0.0359	0.1082	0.1145	0.0671
		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
*Total_File_*	0.5348	-	0.9988	0.8591	0.9532	0.8790	-0.2868	0.3270	0.4047	0.0133	0.0755	0.1106
Size (2)	0.0000		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0152	0.0000	0.0000
*Total_Word_*	0.5306	0.9967	-	0.8626	0.9576	0.8806	-0.2737	0.3211	0.4015	0.0070	0.0869	0.1126
Count (3)	0.0000	0.0000		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.2050	0.0000	0.0000
Total_Senten	0.5869	0.8218	0.8194	-	0.8315	0.7724	0.1448	-0.1071	0.0355	-0.3836	0.0895	0.1162
ce_Count (4)	0.0000	0.0000	0.0000		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
*Total_If_*	0.5721	0.9173	0.9113	0.8357	-	0.8489	-0.2390	0.3050	0.3694	-0.0052	0.3113	0.1064
Count (5)	0.0000	0.0000	0.0000	0.0000		0.0000	0.0000	0.0000	0.0000	0.3460	0.0000	0.0000
*Total_Renew_*	0.4578	0.8699	0.8631	0.7398	0.8594	-	-0.2142	0.2653	0.3385	-0.0186	0.0830	0.4866
Count (6)	0.0000	0.0000	0.0000	0.0000	0.0000		0.0000	0.0000	0.0000	0.0007	0.0000	0.0000
Flesch (7)	0.0221	-0.0724	-0.0678	0.1126	-0.0676	-0.0673	-	-0.9421	-0.7636	-0.8164	0.0475	0.0117
	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000		0.0000	0.0000	0.0000	0.0000	0.0334
Fog (8)	-0.0201	0.0726	0.0686	-0.1088	0.0709	0.0696	-0.9984	-	0.8020	0.8467	0.0248	0.0151
	0.0002	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000		0.0000	0.0000	0.0000	0.0059
Smog (9)	-0.0156	0.1858	0.1806	-0.0735	0.1671	0.1774	-0.4963	0.4944	-	0.6275	0.0005	0.0515
	0.0045	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000		0.0000	0.9283	0.0000

Dale_Chall (10)	-0.0169 0.0021	0.0439 0.0000	0.0403 0.0000	-0.1317 0.0000	0.0430 0.0000	0.0394 0.0000	-0.9968 0.0000	0.9982 0.0000	0.4794 0.0000	-	-0.0043 0.4380	-0.0328 0.0000
If_Percent (11)	0.0323 0.0000	-0.0437 0.0000	-0.0406 0.0000	-0.0144 0.0089	0.1078 0.0000	-0.0351 0.0000	-0.0322 0.0000	0.0423 0.0000	-0.0213 0.0001	0.0426 0.0000	-	0.0197 0.0003
Renew_Percent (12)	-0.0043 0.4377	-0.0407 0.0000	-0.0406 0.0000	-0.0391 0.0000	-0.0436 0.0000	0.0980 0.0000	0.0254 0.0000	-0.0242 0.0000	-0.0404 0.0000	-0.0214 0.0001	-0.0387 0.0000	-

Panel B: Summary Statistics

	(1) *N*	(2) *Mean*	(3) *Std. Dev.*	(4) *25%*	(5) *Median*	(6) *75%*
Ncontracts	73,883	1.030	2.092	0.000	0.000	1.000
Total_File_Size	73,883	49,630.480	142,805.564	0.000	0.000	29,328.000
Total_Word_Count	73,883	7,615.854	22,095.898	0.000	0.000	4,431.000
Total_Sentence_Count	73,883	74.922	206.045	0.000	0.000	48.000
Flesch	28,287	-120.905	421.203	-81.730	-22.740	3.570
Fog	28,287	72.018	165.048	25.320	34.053	54.021
Smog	28,287	27.489	14.224	20.900	24.300	29.550
Dale_Chall	28,287	13.825	20.336	8.240	9.185	11.487
Total_If_Count	73,883	61.175	167.925	0.000	0.000	34.000
Total_Renew_Count	73,883	12.212	37.525	0.000	0.000	6.000
If_Percent	28,287	0.009	0.005	0.006	0.007	0.010
Renew_Percent	28,287	0.002	0.002	0.001	0.001	0.002

Panel C: Principal Component Analysis

	Factor 1 - Content Complexity 1		Factor 2 - Content Complexity 2		Factor 3 - Linguistic Complexity	
Variable	*Loading*	*Standard Error*	*Loading*	*Standard Error*	*Loading*	*Standard Error*
Ncontracts	0.3001	0.0038	0.0959	0.0122	-0.0373	0.0023
Total_File_Size	**0.4385**	0.0013	-0.0313	0.0034	0.0333	0.0010
Total_Word_Count	**0.4371**	0.0022	-0.0291	0.0035	0.0309	0.0009
Total_Sentence_Count	**0.4107**	0.0013	0.0330	0.0043	-0.0912	0.0034
Total_If_Count	**0.4348**	0.0010	0.0935	0.0034	0.0258	0.0010
Total_Renew_Count	**0.4073**	0.0016	-0.1319	0.0043	0.0359	0.0019
Flesch	0.0101	0.0005	-0.0158	0.0028	**-0.5379**	0.0030
Fog	-0.0090	0.0005	0.0225	0.0029	**0.5377**	0.0030
Smog	0.0479	0.0026	-0.0762	0.0182	0.3473	0.0133
Dale_Chall	-0.0210	0.0007	0.0247	0.0033	**0.5360**	0.0030
If_Percent	0.0074	0.0018	**0.7343**	0.0110	-0.0002	0.0026
Renew_Percent	-0.0157	0.0018	**-0.6447**	0.0161	0.0005	0.0030

This table explores the statistical relationships among firm-quarter level contract complexity measures. All variables are defined in Appendix A. In Panel A, I present the correlation matrix. Pearson's correlation coefficients are shown in the lower triangle, while Spearman's rank correlations appear above the diagonal. p-values are reported below each correlation coefficient. This table presents the summary statistics for the firm-quarter level contract complexity variables. Panel C displays the loadings from the Principal Component Analysis. I retain factors with an eigenvalue greater than one. Standard errors are calculated using the bootstrap method (1,000 samples with replacement).

To better understand the relations between content complexity and linguistic complexity and to reduce the dimensionality of the complexity measures, I conduct a Principal Component Analysis (PCA). Without imposing restrictions on the number of dimensions, the PCA can help identify underlying (latent) complexity constructs and determine which measures are associated with each latent construct. The input complexity measures include the number of contracts, total file size, total word count, total sentence count, total number of conditional statements and renewal indicators,

average Flesch Reading Ease score, SMOG grade, Fog Index, Dale-Chall score, and average percentage of conditional statements and renewal indicators, all measured at the firm-quarter level.

Following prior literature (e.g., Larcker, Richardson, and Tuna 2007; Verdi 2005), I retain factors with eigenvalues greater than one and after the last significant drop in the scree plot. This step results in three factors that retain 77.2% of the total variance in the original data.[23] I rotate the reduced factor solution using an oblique Promax rotation that allows the retained factors to be correlated to enhance interpretability of the loadings. To determine standard errors of the loadings, I use the bootstrap method (1,000 samples with replacement) for the rotated factor solution.

Table 5, Panel C presents the loadings of the principal component analysis. The first factor, which I call content complexity 1, is significantly associated with total file size, total word count, total sentence count, total number of conditional statements, and total count of renewal indicators. The second factor, which I call content complexity 2, is significantly associated with average percentage of conditional words and renewal words. Factor three, which I name linguistic complexity, is significantly associated with average Flesch Reading Ease score, average Fog Index, and average Dale-Chall score.

---

[23] The total variance retained by the three underlying factors, 77.2%, is comparable to PCA results from prior studies in different settings. For example, in Larcker et al. (2007), 61.7% of the total variance is retained by their underlying corporate governance constructs.

There are no significant cross loadings, meaning no input variable is associated with more than one factor. Many prior accounting studies on textual analysis use number of words and Fog index as their measures of financial statement readability (e.g., Li 2008). While these two measures are correlated, as shown in Panel A, the results from the PCA analysis indicate they are associated with different underlying complexity constructs.

Based on these results, I retain input variables with two characteristics: a loading greater than 0.4 in absolute value and test statistics indicating significance at conventional levels, following Larcker et al. (2007) and Verdi (2005). I construct firm-quarter level content complexity and linguistic complexity index scores as the loadings-weighted average of the standardized input variables retained on each factor. I classify a firm as high in content (linguistic) complexity if the firm's content (linguistic) complexity index score in the pre-period, 2001 to 2003, is above sample median.

5.1.5 Results from the Cross-sectional Test Based on Contract Complexity

To test the prediction that the effect of the 2004 rule change on information asymmetry is stronger for firms that file more complex contracts, I replace the *Post* × *Treat* variable in equation (1) with interaction terms *Post* × *Treat* × *High_Complexity* and *Post* × *Treat* × *Low_Complexity*, and re-run the main regression. Table 6 displays the results on this cross-sectional test. In Panel A (Panel B), the measurement window for information asymmetry is 3 days (30 days). I first present results on content complexity

1 in Columns (1) and (2). In Panel A, consistent with my prediction, the interaction term between *Post* × *Treat* and *High_Complexity* is negative and significant at the 1% level while the interaction term between *Post* × *Treat* and *Low_Complexity* is statistically indistinguishable from zero regardless of my treatment group choices. In terms of economic magnitude, in Column (1) the coefficient on *Post* × *Treat* × *High_Complexity* is -0.0892, suggesting that compared to the control group, information asymmetry of treatment firms with more content-complex contracts is about 9% smaller in the 3 days following periodic financial statement disclosures after the 2004 SEC rule change. In contrast, the coefficient on *Post* × *Treat* × *Low_Complexity* is -0.0214 and insignificant at conventional levels, suggesting that when compared to the same control group, information asymmetry of treatment firms with less content-complex contracts is statistically similar. Moreover, a *t* test indicates a significant (at the 1% level) difference between the two coefficient estimates. In Columns (3) and (4), I partition the sample based on content complexity 2 and obtain similar results.

Turning to Columns (5) and (6), I separate the sample based on linguistic complexity. Both coefficients on *Post* × *Treat* × *High_Complexity* and *Post* × *Treat* × *Low_Complexity* are insignificant at conventional levels, indicating treatment firms, whether they are high or low in contract linguistic complexity, do not have differential bid-ask spreads from the control group. Because I calculate contract linguistic complexity only for firms that disclosed at least one contract in the pre-period, the

control group for this test excludes firms that never disclosed material contracts. Results in Panel B, analyzing bid-ask spreads over the 30 days following 10-K/Q disclosures, are similar to those in Panel A. Collectively, these results suggest the information asymmetry effects I document are concentrated in firms that disclosed more content-complex contracts.

Table 6: Comparing High and Low Contract Complexity

Panel A

	(1)	(2)	(3)	(4)	(5)	(6)
Treatment Variable	*Treat1*	*Treat2*	*Treat1*	*Treat2*	*Treat1*	*Treat2*
Complexity Type	*Content 1*	*Content 1*	*Content 2*	*Content 2*	*Linguistic*	*Linguistic*
Window	3 days	3 days	3 days	3 days	3 days	3 days
***Post* × *Treat* ×**	**-0.0892*** **	**-0.0760*** **	**-0.1022*** **	**-0.0890*** **	**-0.0369**	**-0.0336**
High_Complexity (A)	**(-3.7044)**	**(-3.2566)**	**(-4.3060)**	**(-3.8845)**	**(-1.2161)**	**(-1.3049)**
***Post* × *Treat* ×**	**-0.0214**	**-0.0146**	**-0.0052**	**-0.0007**	**0.0200**	**0.0098**
Low_Complexity (B)	**(-0.8812)**	**(-0.6510)**	**(-0.2102)**	**(-0.0311)**	**(0.6570)**	**(0.3842)**
Difference (A) vs. (B)	-0.0678***	-0.0614**	-0.0970***	-0.0883***	-0.0569**	-0.0434*
	(-3.098)	(-2.5749)	(-4.4102)	(-3.7081)	(-2.5769)	(-1.8055)
Observations	73,883	73,883	73,883	73,883	58,332	58,332
Adjusted R-squared	0.8085	0.8085	0.8085	0.8085	0.8004	0.8004
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes

Panel B

	(1)	(2)	(3)	(4)	(5)	(6)
Treatment Variable	*Treat1*	*Treat2*	*Treat1*	*Treat2*	*Treat1*	*Treat2*
Complexity Type	*Content 1*	*Content 1*	*Content 2*	*Content 2*	*Linguistic*	*Linguistic*
Window	30 days	30 days	30 days	30 days	30 days	30 days
***Post* × *Treat* ×**	**-0.1009*** **	**-0.0823*** **	**-0.1189*** **	**-0.1028*** **	**-0.0404**	**-0.0323**
High_Complexity (A)	**(-4.8062)**	**(-4.0597)**	**(-5.7779)**	**(-5.2347)**	**(-1.5200)**	**(-1.4409)**
***Post* × *Treat* ×**	**-0.0412***	**-0.0277**	**-0.0192**	**-0.0060**	**0.0131**	**0.0089**
Low_Complexity (B)	**(-1.9350)**	**(-1.4029)**	**(-0.8807)**	**(-0.2957)**	**(0.4946)**	**(0.3983)**
Difference (A) vs. (B)	-0.0597***	-0.0546***	-0.0997***	-0.0968***	-0.0535***	-0.0412*
	(-3.1193)	(-2.6172)	(-5.1865)	(-4.6508)	(-2.7713)	(-1.9570)
Observations	73,883	73,883	73,883	73,883	58,332	58,332
Adjusted R-squared	0.9085	0.9085	0.9087	0.9086	0.9054	0.9054
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes

This table reports OLS regression results on information asymmetry. The dependent variable is the natural logarithm of the average daily bid-ask spread during the measurement window following the 10-Q or 10-K filing date. In Panel A, the measurement window is 3 days while in Panel B it is 30 days. High (low) complexity is an indicator that equals one for firms with above (below) median average firm-quarter level contract complexity in the pre-period of 2001 to 2003. In Columns (1) and (2), I focus on content complexity 1. In Columns (3) and (4), I partition the treatment firms based on content complexity 2. In Columns (3) and (4) I study linguistic complexity. Content complexity 1, content complexity 2, and linguistic complexity are constructed as the loadings weighted average of the standardized input variables retained on each factor. All variables are defined in Appendix A. I include quarter and firm fixed effects in all specifications. Standard errors are clustered at the firm level. T-statistics are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

5.2 Cross-sectional Variation in Retail Trader Interest

As previously discussed, I predict that the effect of the 2004 SEC rule change should be greater for firms that attract more retail investor interest. I reason that, regardless of information processing capability, investor interest is a pre-condition for processing information disclosed by a firm. For firms with low retail investor interest, the 2004 SEC rule change affects only sophisticated institutional investors, who have better information processing capability. As a result, information asymmetry for these would not decrease and may even increase, to the extent the SEC rule change improves institutional investors' information processing (Blankespoor et al. 2014). In contrast, before the 2004 rule change, the periodic financial statements of firms with substantial retail investor interest were not fully processed by retail investors who have less processing capability. After 2004, when material contracts are separately disclosed earlier and on a staggered basis, retail investors will better process the information in periodic financial statements. Therefore, I predict the information asymmetry effects to be stronger for firms that attract more retail trader interest.

I use the Trade and Quote (TAQ) database to identify retail trades. Following prior literature, I use trade size as a proxy for retail order flow (e.g., Lee and

Radhakrishna 2000; Barber, Odean, and Zhu 2008).[24] Specifically, I define retail trades as those less than $5,000. In robustness tests, I use alternative thresholds and find similar results. For each day, I aggregate the total dollar amount of retail transactions. Next, I define firm-quarter level retail trader interest as the total dollar size of retail trades over days [-70, -11] relative to the earnings announcements for each quarter. Finally, I take the average of this firm-quarter level measure over the pre-period 2001 to 2003. I define firms with high (low) retail investor interest as those with above (below) median average retail trade dollar size and create indicator variables *High_Retail* and *Low_Retail* for firms with high and low retail investor interest, respectively.

To conduct this cross-sectional test, I replace the *Post* × *Treat* variable in equation (1) with interaction terms *Post* × *Treat* × *High_Retail* and *Post* × *Treat* × *Low_Retail*. Table 7, Panel A presents the results. Consistent with my prediction, the interaction term between *Post* × *Treat* and *High_Retail* is negative and significant at the 1% level while the interaction term between *Post* × *Treat* and *Low_Retail* is indistinguishable from zero across all specifications. For example, in Column (1) where I use *Treat1* as the treatment group and $5,000 as the cutoff for retail trades, the coefficient on *Post* × *Treat* ×

[24] Recent developments in the finance literature suggests that using trade size as the measure of retail trades can be problematic as computer algorithms that "slice and dice" large institutional parent orders into sequences of small child orders become more prevalent (e.g., Cready, Kumas, and Subasi 2014; Boehmer, Jones, Zhang, and Zhang 2021). However, almost all the declines in trade size occurred after 2005 (Cready et al. 2014). Because I only use 2001 to 2003 data in determining retail orders and firms that attract more retail interest, this problem is less concerning in my study.

High_Retail is -0.1019, suggesting that compared to the control group, information asymmetry of treatment firms that draw high retail investor interest is about 10.19% smaller in the 3 days following periodic financial statement disclosures after the 2004 SEC rule change. In contrast, the coefficient on *Post* × *Treat* × *Low_Retail* is -0.0019 and insignificant at conventional levels, suggesting that when compared to the same control group, information asymmetry of treatment firms that draw low retail investor interest is statistically similar. A *t* test indicates a significant (at the 1% level) difference between the two coefficients. In Columns (3) and (4), I set the small trade threshold at $10,000 and find similar results. Panel B reports quantitively similar results for the 30 days following periodic disclosures. Collectively, the results in Table 7 support the idea that information asymmetry effects are stronger for firms that attract more retail trader interest.

Table 7: Comparing High and Low Retail Trader Interest

Panel A

	(1)	(2)	(3)	(4)
Treatment Variable	*Treat1*	*Treat2*	*Treat1*	*Treat2*
Small Trade Threshold	*$5,000*	*$5,000*	*$10,000*	*$10,000*
Window	3 days	3 days	3 days	3 days
***Post* × *Treat* × *High_Retail* (A)**	**-0.1019*** **	**-0.0875*** **	**-0.1080*** **	**-0.0938*** **
	(-4.2151)	**(-3.8738)**	**(-4.4826)**	**(-4.1686)**
***Post* × *Treat* × *Low_Retail* (B)**	**-0.0019**	**0.0102**	**0.0069**	**0.0200**
	(-0.0764)	**(0.4366)**	**(0.2815)**	**(0.8489)**
Difference (A) vs. (B)	-0.1000***	-0.0773***	-0.1011***	-0.0738***
	(-4.5453)	(-4.0755)	(-5.2202)	(-4.7508)

Observations	67,519	67,519	67,519	67,519
Adjusted R-squared	0.8089	0.8089	0.8090	0.8089
Controls	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes

Panel B

	(1)	(2)	(3)	(4)
Treatment Variable	*Treat1*	*Treat2*	*Treat1*	*Treat2*
Small Trade Threshold	*$5,000*	*$5,000*	*$10,000*	*$10,000*
Window	30 days	30 days	30 days	30 days
Post × Treat × High_Retail (A)	**-0.1124*** **	**-0.0914*** **	**-0.1170*** **	**-0.0953*** **
	(-5.3565)	**(-4.6832)**	**(-5.6025)**	**(-4.9010)**
Post × Treat × Low_Retail (B)	**-0.0265**	**-0.0068**	**-0.0198**	**-0.0007**
	(-1.2247)	**(-0.3271)**	**(-0.9127)**	**(-0.0354)**
Difference (A) vs. (B)	-0.0859***	-0.0846***	-0.0972***	-0.0946***
	(-4.4688)	(-4.0125)	(-5.0547)	(-4.4800)
Observations	67,519	67,519	67,519	67,519
Adjusted R-squared	0.9087	0.9087	0.9088	0.9087
Controls	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes

This table reports OLS regression results on information asymmetry. The dependent variable is the natural logarithm of the average daily bid-ask spread during the measurement window following the 10-Q or 10-K filing date. In Panel A, the measurement window is 3 days while in Panel B it is 30 days. High (low) Retail is an indicator that equals one for firms with above (below) median average firm-quarter level retail trading activities in the pre-period of 2001 to 2003. In Columns (1) and (2), I define retail trades as those smaller than $5,000 and in Columns (3) and (4) the cutoff is $10,000. All variables are defined in Appendix A. I include quarter and firm fixed effects in all specifications. Standard errors are clustered at the firm level. T-statistics are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

5.3 Cross-sectional Variation in Change in Timeliness of Material Contract Filings

The reduction in post-periodic-filing information asymmetry associated with the 2004 SEC rule change should be more pronounced for firms that substantially increased the timeliness of material contract filings, because investors of these firms gain relatively more time before the periodic financial statement disclosures to process the information in the contracts.[25] For each firm, I define change in timeliness as the median days between material contracts filings and the corresponding periodic financial statement filings in the post-period minus that in the pre-period. I classify a firm as high (low) in timeliness change if its change in timeliness is above (below) the sample median. Then, I replace the *Post* × *Treat* variable in equation (1) with interaction terms *Post* × *Treat* × *High_Timeliness_Change* and *Post* × *Treat* × *Low_Timeliness_Change*. Table 8 presents the results. Consistent with my prediction, I find the reduction in *Spread* is 45.76% (= 0.0221/0.0483) larger in the three days following 10-K/Q disclosures for firms with an above-median change in timeliness of their material contract filings (Column 1, Panel A). The inference does not change when I use *Treat2* as the treatment variable or when I analyze the bid-ask spread over the 30 days following 10-K/Q filings.

[25] Firms that have more complex contracts might be more likely to file contracts on a timelier basis. In this case, it should bias against me finding the larger reduction in information asymmetry.

Table 8: Comparing Change in Timeliness of Material Contract Filings

Panel A

		(1)	(2)
Treatment Variable		*Treat1*	*Treat2*
Window		3 days	3 days
Post × Treat × High_Timeliness_Change	***(A)***	**-0.0704*** **	**-0.0606*** **
		(-3.1268)	**(-2.9010)**
Post × Treat × Low_Timeliness_Change	***(B)***	**-0.0483** **	**-0.0341* **
		(-2.1894)	**(-1.6928)**
Difference (A) vs. (B)		-0.0221**	-0.0265**
		(-1.9799)	(-2.1909)
Observations		73,883	73,883
Adjusted R-squared		0.8074	0.8074
Controls		Yes	Yes
Firm FE		Yes	Yes
Time FE		Yes	Yes

Panel B

		(1)	(2)
Treatment Variable		*Treat1*	*Treat2*
Window		30 days	30 days
Post × Treat × High_Timeliness_Change	***(A)***	**-0.0878*** **	**-0.0731*** **
		(-4.5136)	**(-4.0894)**
Post × Treat × Low_Timeliness_Change	***(B)***	**-0.0615*** **	**-0.0406** **
		(-3.1963)	**(-2.3218)**
Difference (A) vs. (B)		-0.0263***	-0.0325***
		(-3.1496)	(-3.6097)
Observations		73,883	73,883
Adjusted R-squared		0.9087	0.9087

Controls	Yes	Yes
Firm FE	Yes	Yes
Time FE	Yes	Yes

This table reports OLS regression results on information asymmetry. The dependent variable is the natural logarithm of the average daily bid-ask spread during the measurement window following the 10-Q or 10-K filing date. In Panel A, the measurement window is 3 days while in Panel B it is 30 days. *High (Low)_Timeliness_Change* is an indicator that equals one (zero) if *Timeliness_Change* is above (below) sample median. For each firm, *Timeliness_Change* is calculated as the median days between material contracts filings and periodic financial statement filings in the post-period minus that in the pre-period. All variables are defined in Appendix A. I include quarter and firm fixed effects in all specifications. Standard errors are clustered at the firm level. T-statistics are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

5.4 Cross-sectional Variation in Investor Busyness

My final cross-sectional tests are based on investor busyness. Based on the reasoning that busy investors are more constrained in information processing capabilities, the information asymmetry effects associated with the 2004 SEC rule change should be more pronounced for financial statements filed on days when investors are busy, or, alternatively, distracted.

My first measure of investor busyness is whether a periodic financial report is released on a Friday. DellaVigna and Pollett (2009) and Damodaran (1989) theorize that market participants are distracted just before weekends and, as a result, they allocate less attention to financial statements released on a Friday, as compared to Monday through Thursday. Consistent with this idea, DellaVigna and Pollett (2009) find lower earnings response coefficients, higher post earnings announcement drift, and lower trading volume on Fridays. My second measure of investor busyness is whether the periodic financial statement is a 10-K as opposed to a 10-Q. Because the 10-K report is

longer and more detailed than the 10-Q report, investors must process more information when a 10-K form is released. Finally, based on research showing that the well-documented Friday effect can be partially driven by the scheduled monthly release of the Employment Situation (ES) reports from the Bureau of Labor Statistics because investors pay attention to the ES reports (e.g., Ederington and Lee 1996; Graham et al. 2003; Melessa 2012), my last measure of investor busyness is whether a periodic financial statement is disclosed on an ES report day.

Similar to cross-sectional tests in the previous sections, I replace the *Post* × *Treat* variable in equation (1) with interaction terms *Post* × *Treat* × *Busy* and *Post* × *Treat* × *Not_Busy*, and re-run the main regression. Table 9, Panel A presents the results for the 3 days following financial statements disclosure. Across all specifications, the coefficient estimates on both *Post* × *Treat* × *Busy* and *Post* × *Treat* × *Not_Busy* are negative and significant at conventional levels, indicating declines in information asymmetry following the 2004 SEC rule change regardless of whether investors are busy. Consistent with my prediction, the magnitude of the coefficient estimates on *Post* × *Treat* × *Busy* exceed those on *Post* × *Treat* × *Not_Busy* and *t* tests confirm the differences are statistically significant at the 1% level across all four columns. In Panel B where the measurement window is extended to 30 days after financial statement disclosures, the coefficient estimates on the two variables are statistically different only when investor busyness is measured by *Friday* or *ES_Day*. In summary, these tests provide evidence

that *Spread* declines more for periodic financial statements that are disclosed when investors are busier.

Table 9: Comparing Investor Busyness

Panel A

	(1)	(2)	(3)	(4)	(5)	(6)
Treatment Variable	*Treat1*	*Treat2*	*Treat1*	*Treat2*	*Treat1*	*Treat2*
Measure of Busyness	*Friday*	*Friday*	*10-K*	*10-K*	*ES_Day*	*ES_Day*
Window	3 days	3 days	3 days	3 days	3 days	3 days
***Post* × *Treat* × *Busy* (A)**	**-0.0862*** **	**-0.0805*** **	**-0.0804*** **	**-0.0651*** **	**-0.1237*** **	**-0.1138*** **
	(-3.6854)	**(-3.6654)**	**(-3.5240)**	**(-3.0852)**	**(-4.5190)**	**(-4.2707)**
***Post* × *Treat* × *Not_Busy* (B)**	**-0.0512** **	**-0.0372***	**-0.0517** **	**-0.0405** **	**-0.0533** **	**-0.0406** **
	(-2.3671)	**(-1.8859)**	**(-2.3986)**	**(-2.0622)**	**(-2.4705)**	**(-2.0696)**
Difference (A) vs. (B)	-0.0350***	-0.0433***	-0.0287***	-0.0246***	-0.0704***	-0.0732***
	(-3.1670)	(-3.5972)	(-3.2909)	(-2.6870)	(-3.9357)	(-3.8171)
Observations	73,883	73,883	73,883	73,883	73,883	73,883
Adjusted R-squared	0.8075	0.8074	0.8074	0.8074	0.8075	0.8075
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes

Panel B

	(1)	(2)	(3)	(4)	(5)	(6)
Treatment Variable	*Treat1*	*Treat2*	*Treat1*	*Treat2*	*Treat1*	*Treat2*
Measure of Busyness	*Friday*	*Friday*	*10-K*	*10-K*	*ES_Day*	*ES_Day*
Window	30 days	30 days	30 days	30 days	30 days	30 days
Post × Treat × Busy (A)	**-0.0848*** **	**-0.0720*** **	**-0.0789*** **	**-0.0597*** **	**-0.1106*** **	**-0.0941*** **
	(-4.2404)	**(-3.8774)**	**(-4.0485)**	**(-3.3405)**	**(-4.9233)**	**(-4.3566)**
Post × Treat × Not_Busy (B)	**-0.0708*** **	**-0.0513*** **	**-0.0721*** **	**-0.0544*** **	**-0.0708*** **	**-0.0524*** **
	(-3.7413)	**(-2.9864)**	**(-3.8268)**	**(-3.1864)**	**(-3.7488)**	**(-3.0631)**
Difference (A) vs. (B)	-0.0140*	-0.0207**	-0.0068	-0.0053	-0.0398***	-0.0417***
	(-1.7635)	(-2.4083)	(-1.3601)	(-1.0149)	(-3.1417)	(-3.0545)
Observations	73,883	73,883	73,883	73,883	73,883	73,883
Adjusted R-squared	0.9087	0.9086	0.9087	0.9086	0.9087	0.9086
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes

This table reports OLS regression results on information asymmetry. The dependent variable is the natural logarithm of the average daily bid-ask spread during the measurement window following the 10-Q or 10-K filing date. In Panel A, the measurement window is 3 days while in Panel B it is 30 days. Measure of busyness is *Friday* in Columns (1) and (2) in both panels. *Friday* is an indicator that equals one if the financial statement is disclosed on a Friday. Measure of busyness is *10-K* in Columns (3) and (4). *10-K* is an indicator that equals one if the financial statement is a 10-K form as opposed to a 10-Q form. Finally, in Columns (5) and (6), busyness is measured by *ES_Day*, an indicator that quals one if the financial statement is disclosed on the release day of the monthly Employment Situation from the Bureau of Labor Statistics. All variables are defined in Appendix A. I include quarter and firm fixed effects in all specifications. Standard errors are clustered at the firm level. T-statistics are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

5.5 Other Robustness Tests

I conclude my analyses with several robustness tests for my main results. First, to further address the concern that treatment is not randomly assigned, I use propensity score matching to create more similar control and treatment groups. More specifically, I conduct one-to-one matching without replacement based on return on assets, leverage, book-to-market ratio, size, average stock price, average trading volume, analyst following, and previous return volatility. I then re-run the main regressions based on equation (1) using the matched sample. Table 10, Panel A presents the results. My inferences do not change.

Table 10: Additional Robustness Tests

Panel A

	(1)	(2)	(3)	(4)
Treatment Variable	*Treat1*	*Treat2*	*Treat1*	*Treat2*
Window	3 days	3 days	30 days	30 days
Post* × *Treat	**-0.1795*** **	**-0.1684*** **	**-0.1664*** **	**-0.1533*** **
	(-6.4753)	**(-6.0541)**	**(-6.9427)**	**(-6.3705)**
Observations	46,200	46,200	46,200	46,200
Adjusted R-squared	0.7921	0.7920	0.9030	0.9029
Controls	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes

Panel B

	(1)	(2)	(3)	(4)
Treatment Variable	*Treat1*	*Treat2*	*Treat1*	*Treat2*
Window	3 days	3 days	30 days	30 days
Post* × *Treat	**-0.0184*** **	**-0.0203*** **	**-0.0045***	**-0.0076*** **
	(-5.2928)	**(-6.5702)**	**(-1.6612)**	**(-3.1914)**
Observations	71,931	71,931	71,931	71,931
Adjusted R-squared	0.1231	0.1235	0.2256	0.2257
Controls	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes

This table reports OLS regression results on information asymmetry. In Panel A, the dependent variable is the natural logarithm of the average daily bid-ask spread during the measurement window following the 10-Q or 10-K filing date. I use propensity score matching to create more similar control and treatment groups. Specifically, I conduct one-to-one matching without replacement based on return on assets, leverage, book-to-market ratio, size, average stock price, average trading volume, analyst following, and previous return volatility. Then I re-run the same tests in Table 3. In Panel B, the dependent variable is the probability of informed trading, PIN, from Easley et al. (1996) during the measurement window following the 10-Q or 10-K filing date. All other variables are defined in Appendix A. I include quarter and firm fixed effects in all specifications. Standard errors are clustered at the firm level. T-statistics are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Second, I conduct sensitivity tests using alternative measures of information asymmetry. Easley et al. (1996) develop a measure of the probability of informed trading, PIN, based on the model of Glosten and Milgrom (1985). Following Easley et al., I calculate the 3-day and 30-day PIN following 10-K/Q disclosures as alternative measures of information asymmetry among investors. Then I re-run equation (1) using PIN as the dependent variable. Table 10, Panel B presents the results. Again, the inferences remain unchanged.

6. Conclusion

Previous research documents that disclosure complexity can make it difficult for investors, especially retail investors who are more constrained in information processing capability, to fully process the disclosed information. The result can be higher information asymmetry between processing-capability-constrained retail investors and less-constrained sophisticated investors. In this study, I provide evidence that making periodic financial reports less complex, without reducing the total amount of information provide, helps mitigate this problem. Using a 2004 SEC mandate that moves material contract disclosures from 10-K/Q reports to event-based 8-K reports which must be filed within four business days of contract signing, I find that information asymmetry among investors following the release of 10-K/Q reports decreases after 2004. Furthermore, I find the reduction in information asymmetry is more pronounced for firms with more complex contract filings and more retail trader interest, and firms that experience greater increases in the timeliness of their material contract filings. The effects are larger for 10-K reports as compared to 10-Q reports, and also larger when these reports are released during times when investors are typically paying less attention to financial reports (e.g., on Fridays and on days when Employment Situation reports are released).

My findings are subject to the caveat that firms do not receive random assignment of treatments in my study, because the 2004 SEC rule change is applicable to

all registrants. I try to measure the extent to which firms are differently affected by this mandate by studying prior disclosure patterns, while acknowledging that variation in disclosure patterns could reflect variation in economic conditions. Although I cannot draw the conclusion that the 2004 requirement caused a reduction in information asymmetry, I can conclude that my results suggest a negative association between the bid-ask spread and the mandate, consistent with the idea that the latter reduces the information overload problem in periodic financial statements.

My study provides evidence on how the form and timing of disclosures can affect how information is processed by equity market participants. The evidence in this paper should be of interest to regulators as it adds to the growing body of evidence on the effects of disclosure complexity on disclosure effectiveness.

Appendix A: Variable Definitions

Spread	The natural logarithm of the average daily bid-ask spread during the 3 days following the 10-Q or 10-K filing date. For each firm-day, the bid-ask spread is measured from the CRSP Daily Stock File as ASK-BID scaled by the average of ASK and BID.
ROA	Return on assets measured from the Compustat Quarterly database as NIQ scaled by ATQ.
Leverage	Leverage measured from the Compustat Quarterly database as LTQ scaled by ATQ.
BTM	Book-to-market ratio measured from the Compustat Quarterly database as SEQQ scaled by (PRCCQ*CSHOQ).
Size	Firm size measured from the Compustat Quarterly database as the natural logarithm of ATQ.
Price	The average price during the 3 days following the 10-Q or 10-K filing date. For each firm-day, price is measured from the CRSP Daily Stock File as the absolute value of PRC.
Turnover	The natural logarithm of the average daily trading volume during the 3 days following the 10-Q or 10-K filing date. For each firm-day. The volume is measured from the CRSP Daily Stock File as VOL scaled by (SHROUT*1000).
Follow	The natural logarithm of 1 + the number of unique analysts issuing an EPS forecast during the three calendar months immediately preceding the 10-Q or 10-K filing date. Unique analysts are identified by the ANALYS variable in the IBES Detail U.S. EPS file.
Avgvix	The average value of the CBOE Volatility Index during the 3 days following the 10-Q or 10-K filing date.
Paststdret	The standard deviation of daily returns from the CRSP Daily Stock File during the 3 days ending one day prior to the 10-Q or 10-K filing date.

Post	An indicator variable that equals one (zero) for all observations with a 10-Q or 10-K filing date on or after (before) August 23, 2004.
Treat1	An indicator variable that equals zero for all firms that (1) did not file any material contracts, or (2) were timely 8-K material contract filers between 2001 and 2003. I define timely 8-K material contract filers as firms that filed at least one material contract via 8-K in the pre period and the average filing lag of these pre-period 8-K material contracts are below sample median between 2001 and 2003. This variable is set to one for all other firms.
Treat2	An indicator variable that equals zero for all firms that (1) did not file any material contracts, or (2) were 8-K material contract filers between 2001 and 2003. I define 8-K material contract filers as firms that filed at least one material contract via 8-K in the pre period. This variable is set to one for all other firms.
Treat3	The average number of quarterly material contract filings between 2001 and 2003.
Treat4	An indicator variable that equals zero for all firms that (1) did not file any material contracts between 2001 and 2003. This variable is set to one for all other firms.
Ncontracts	Total number of material contracts filed either with 10-Q/Ks or 8-Ks in a quarter.
Total_File_Size	The sum of the size of material contract documents in a quarter.
Total_Word_Count	The total number of words in material contracts in a quarter.
Total_Sentence_Count	The total number of sentences in material contracts in a quarter.
Total_If_Count	The total number of conditional statements in material contracts in a quarter. I use the following keywords to identify conditional statements: *if, assuming that, granted that, on the assumption that, supposing that, wherever, conceding that, in case that, on the occasion that, whenever, with the condition that, contingent upon, in case, with the proviso, given, in the case that, on these terms, in the event, subject to, on condition, supposing, in case*.

Total_Renew_Count	The total number of renew indicators in the material contracts filed in a quarter. I use the following keywords to identify renew indicators: *renew, renewal, extend, continue, resume.*
Flesch	The average Flesch Reading Ease Score of the material contracts filed in a quarter. For each contract, the score is calculated as 206.835-1.015×average sentence length-84.6×average syllable count in a word.
Fog	The average Gunning Fog Index of the material contracts filed in a quarter. For each contract, the index is calculated as 0.4×average sentence length+40×percentage of complex words, where complex words are defined as words with three or more syllables.
Smog	The average SMOG grade of the material contracts filed in a quarter. For each contract, the grade is calculated as 1.043×sqrt(30×number of polysyllables/number of sentences)+3.1291.
Dale_Chall	The average Dale-Chall readability score of the material contracts filed in a quarter. For each contract, the score is calculated as 15.79×percentage of difficult wrods+0.0496×average sentence length, where difficult words are defined as those not included in the Dale-Chall Word List.
If_Percent	The average percentage of conditional statements of the material contracts filed in a quarter. For each contract, the percentage is calculated as number of conditional statements scaled by total number of words. This measure is set to zero if a firm did not file any contracts in the quarter.
Renew_Percent	The average percentage of renew indicators of the material contracts filed in a quarter. For each contract, the percentage is calculated as number of renew indicators scaled by total number of words. This measure is set to one if a firm did not file any contracts in the quarter.

Appendix B: An Example of a Material Contract

This appendix displays the first section of the Joint Venture Agreement for St. Charles Active Adult Community between U.S. Home Corporation, a Delaware corporation (hereinafter "USH"), and St. Charles Community, LLC, a Delaware limited liability company. Texts in bold represent the conditional statements identified by my Python script.

1. FORMATION OF ST. CHARLES ACTIVE ADULT COMMUNITY, L.L.C.

It is the intention of the parties to undertake the Project with a limited liability company in which USH and SCC shall be 50% members each. Provided USH does not terminate this Agreement pursuant to its Study Period set forth in Section 3 below, promptly following the end of the Study Period USH and SCC shall form a Maryland limited liability company, the name of which shall be "St. Charles Active Adult Community, L.L.C.", or such other name as is agreed to by USH and SCC (the "LLC"), and further agree that the operation of the LLC shall be governed by an operating agreement in the form attached to this Agreement as Exhibit A and incorporated herein by reference (the "Operating Agreement"). Each party shall contribute to this joint venture the considerations and contributions as are set forth in this Agreement, and each party shall make the capital contributions to the LLC as are set forth in the Operating Agreement. Promptly following the end of the Study Period, USH agrees to file the

Articles of Organization for the LLC with the Maryland State Department of Assessments and Taxation, at the cost of the LLC. For the considerations and on the terms set forth in this Agreement, SCC shall contribute the Property to this Joint Venture and to the LLC. The LLC will subdivide and develop the Project on the Property and will in turn sell finished lots to USH pursuant to a lot option contract, in form and substance approved by both SCC and USH prior to the end of the Study Period (such approval not to be unreasonably withheld, conditioned or delayed) (the "Lot Contract"). **If the parties are unable to agree on the form of the Lot Contract prior to the end of the Study Period, in their reasonable discretion, then either party, by written notice to the other, shall have the right to terminate this Agreement, whereupon neither party shall have any further rights or obligations hereunder. If, prior to any such termination by either party' the parties agree on the form of the Lot Contract, this right to terminate shall be of no further force or effect.** The parties acknowledge and agree that the purchase price for lots to be set forth in the Lot Contract shall be established at a sum equal to 120% of the LLC's cost of delivering finished lots (including, but not limited to, land costs as determined pursuant to Paragraph 7 (d) below, engineering costs incurred by the LLC and all costs of constructing improvements to the Property by the LLC). The Lot Contract will require that USH begin purchasing lots from the LLC on the date no later than the date that the lots are certified by the LLC to be fully subdivided and ready for the issuance of building

permits pursuant to the Lot Contract. Thereafter, USH shall be required to purchase a minimum of sixty (60) lots per year pursuant to the Lot Contract, and the purchase price of the lots pursuant to the Lot Contract shall be subject to an escalator of 6% per annum, commencing on the date of the first lot takedown.

For purposes of the calculation of the purchase price under the Lot Contract, the LLC's costs of delivering finished lots shall also include, but shall not be limited to, all of the costs to: subdivide the Property from the larger parcel owned by SCC, obtain all permits and governmental approvals necessary to subdivide and develop the project, bond the development and construction of the infrastructure for the Project, furnish utilities to the lots in the Project, design and engineer the Project and the Project infrastructure, design and relocate the public right of way known as Demarr Road, design and relocate the private right of way on the Property used by the Demarr family and others, pay all interest carry due to the development loan lender on the Property, pay any applicable association fees or shortfall in the Association budget to the Project neighborhood association, pay the "Assumed Value" of the Property as provided in Section 7(d) herein (plus a 6% per annum escalation from January 1, 2003 until the initial lot purchase take down under the Lot Contract), pay the 3% management fee payable to SCC as provided in the LLC's Operating Agreement, pay a $t,500.00 per lot pump station contribution, pay a $6,000.00 per lot road fee, pay for the development and construction of a community recreation center and related facilities, and pay all and any

other costs and expenses necessary for preparation, delivery and/or conveyance of the Lots under the Lot Contract and for all labor and materials for the work required by the Lot Contract. **USH shall have the right to terminate this transaction if the total of said costs, plus 20%, would cause the purchase price of the Lots on the date of the initial take down to exceed the amounts set forth in Section 5(e) hereof.** Such rights of termination on the part of USH shall be governed by the provisions of Section 5(e) hereof. In addition to the foregoing, the Lot Contract shall require USH to place an earnest money deposit thereunder, upon execution of the Lot Contract, in the amount of $1,000,000.00 in good U.S. funds or by way of letter of credit issued by a recognized financial institution.

SCC shall pay the costs for any off-site utility work to bring utilities to the boundary line of the Property. **The Lot Contract shall provide that the LLC will sell lots to other builders in the event of any default thereunder by USH.**

References

Asthana, S., & Balsam, S. (2001). The effect of EDGAR on the market reaction to 10-K filings. *Journal of Accounting and Public Policy*, *20*(4-5), 349-372.

Asthana, S., Balsam, S., & Sankaraguruswamy, S. (2004). Differential response of small versus large investors to 10-K filings on EDGAR. *The Accounting Review*, *79*(3), 571-589.

Ball, R. (1992). The earnings-price anomaly. *Journal of Accounting and Economics*, *15*(2-3), 319-345.

Barber, B. M., Odean, T., & Zhu, N. (2008). Do retail trades move markets? *The Review of Financial Studies*, *22*(1), 151-186.

Barron, O. E., Byard, D., & Enis, C. R. (2004). Leveling the informational playing field. *Review of Accounting and Finance*.

Beatty, R. P., & Hand, J. R. (1992). The causes and effects of mandated accounting standards: SFAS No. 94 as a test of the level playing field theory. *Journal of Accounting, Auditing & Finance*, *7*(4), 509-530.

Bebchuk, L. A., & Stole, L. A. (1993). Do short-term objectives lead to under-or overinvestment in long-term projects? *The Journal of Finance*, *48*(2), 719-729.

Bens, D. A., Cheng, M., & Neamtiu, M. (2016). The impact of SEC disclosure monitoring on the uncertainty of fair value estimates. *The Accounting Review*, *91*(2), 349-375.

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.

Blankespoor, E., DeHaan, E., Wertz, J., & Zhu, C. (2019). Why do individual investors disregard accounting information? The roles of information awareness and acquisition costs. *Journal of Accounting Research*, *57*(1), 53-84.

Bloomfield, R. J. (2002). The "incomplete revelation hypothesis" and financial reporting. *Accounting Horizons*, *16*(3), 233.

Bloomfield, R., Hodge, F., Hopkins, P., & Rennekamp, K. (2015). Does coordinated presentation help credit analysts identify firm characteristics? *Contemporary Accounting Research*, *32*(2), 507-527.

Boehmer, E., Jones, C. M., Zhang, X., & Zhang, X. (2021). Tracking retail investor activity. *The Journal of Finance*, *76*(5), 2249-2305.

Bruner, Jerome S. & Goodnow, Jacqueline J. & Austin, George A. & Harvard Cognition Project. (1956). *A study of thinking*. New York : Wiley.

Bushee, B. J., Gow, I. D., & Taylor, D. J. (2018). Linguistic complexity in firm disclosures: Obfuscation or information? *Journal of Accounting Research*, *56*(1), 85-121.

Casey Jr, C. J. (1980). Variation in accounting information load: The effect on loan officers' predictions of bankruptcy. *The Accounting Review*, 36-49.

Chapman, K. L., Reiter, N., White, H. D., & Williams, C. D. (2019). Information overload and disclosure smoothing. *Review of Accounting Studies*, *24*(4), 1486-1522.

Costello, A. M. (2013). Mitigating incentive conflicts in inter-firm relationships: Evidence from long-term supply contracts. *Journal of Accounting and Economics*, *56*(1), 19-39.

Cowan, N. (2017). The many faces of working memory and short-term storage. *Psychonomic Bulletin & Review*, *24*(4), 1158-1170.

Cready, W., Kumas, A., & Subasi, M. (2014). Are trade Size-Based inferences about traders reliable? evidence from institutional Earnings-Related trading. *Journal of Accounting Research*, *52*(4), 877-909.

Damodaran, A. (1989). The weekend effect in information releases: A study of earnings and dividend announcements. *The Review of Financial Studies*, *2*(4), 607-623.

DeHaan, E., Shevlin, T. J., & Thornock, J. R. (2015). Market (in) attention and the strategic scheduling and timing of earnings announcements. *Journal of Accounting and Economics*, *60*(1), 36-55.

DellaVigna, S., & Pollet, J. M. (2007). Demographics and industry returns. *American Economic Review*, *97*(5), 1667-1702.

DellaVigna, S., & Pollet, J. M. (2009). Investor inattention and Friday earnings announcements. *The Journal of Finance*, *64*(2), 709-749.

Easley, D., Kiefer, N. M., O'hara, M., & Paperman, J. B. (1996). Liquidity, information, and infrequently traded stocks. *The Journal of Finance*, *51*(4), 1405-1436.

Ederington, L. H., & Lee, J. H. (1996). The creation and resolution of market uncertainty: the impact of information releases on implied volatility. *Journal of Financial and Quantitative Analysis*, *31*(4), 513-539.

Eggleston, K., Posner, E. A., & Zeckhauser, R. (2000). Simplicity and complexity in contracts. Working paper. *University of Chicago Law School, John M. Olin Program in Law and Economics Working Paper No. 93.*

Eggleston, K., Posner, E. A., & Zeckhauser, R. (2000). The design and interpretation of contracts: why complexity matters. *Northwestern University Law Review*, *95*, 91.

Einhorn, H. J. (1971). Use of nonlinear, noncompensatory models as a function of task and amount of information. *Organizational Behavior and Human Performance*, *6*(1), 1-27.

Gao, M., & Huang, J. (2020). Informing the market: The effect of modern information technologies on information production. *The Review of Financial Studies*, *33*(4), 1367-1411.

Glosten, L. R., & Milgrom, P. R. (1985). Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics*, *14*(1), 71-100.

Graham, M., Nikkinen, J., & Sahlström, P. (2003). Relative importance of scheduled macroeconomic news for stock market investors. *Journal of Economics and Finance*, *27*(2), 153-165.

Grossman, S. J., & Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets. *The American Economic Review*, *70*(3), 393-408.

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.

Gunning, R. (1969). The fog index after twenty years. *Journal of Business Communication*, *6*(2), 3-13.

Higgins, K. (2014). Disclosure effectiveness: remarks before the American Bar Association business law section Spring meeting. https://www.sec.gov/news/speech/2014-spch041114kfh.

Hirshleifer, D., Lim, S. S., & Teoh, S. H. (2009). Driven to distraction: extraneous events and underreaction to earnings news. *The Journal of Finance*, *64*(5), 2289-2325.

Hirshleifer, D., & Teoh, S. H. (2003). Limited attention, information disclosure, and financial reporting. *Journal of Accounting and Economics*, *36*(1-3), 337-386.

Indjejikian, R. J. (1991). The impact of costly information interpretation on firm disclosure decisions. *Journal of Accounting Research*, *29*(2), 277-301.

Jones, M. J., & Shoemaker, P. A. (1994). Accounting narratives: A review of empirical studies of content and readability. *Journal of Accounting Literature*, *13*, 142.

Kozminsky, E. (1977). Altering comprehension: The effect of biasing titles on text comprehension. *Memory & Cognition*, *5*(4), 482-490.

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.

Larcker, D. F., Richardson, S. A., & Tuna, I. R. (2007). Corporate governance, accounting outcomes, and organizational performance. *The Accounting Review*, *82*(4), 963-1008.

Lee, C. M., & Radhakrishna, B. (2000). Inferring investor behavior: Evidence from TORQ data. *Journal of Financial Markets*, *3*(2), 83-111.

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

Lerman, A., & Livnat, J. (2010). The new Form 8-K disclosures. *Review of Accounting Studies*, *15*(4), 752-778.

Leuz, C., & Verrecchia, R. E. (2000). The economic consequences of increased disclosure. *Journal of Accounting Research*, 91-124.

Lev, B. (1988). Toward a theory of equitable and efficient accounting policy. *The Accounting Review*, 1-22.

Li, E. X. (2013). Revealing future prospects without forecasts: The case of accelerating material contract filings. *The Accounting Review*, *88*(5), 1769-1804.

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

Loughran, T., & McDonald, B. (2014). Measuring readability in financial disclosures. *the Journal of Finance*, *69*(4), 1643-1671.

Loughran, T., & McDonald, B. (2016). Textual analysis in accounting and finance: A survey. *Journal of Accounting Research*, *54*(4), 1187-1230.

Maćkowiak, B., Matějka, F., & Wiederholt, M. (2018). Dynamic rational inattention: Analytical results. *Journal of Economic Theory*, *176*, 650-692.

Malhotra, N. K. (1982). Information load and consumer decision making. *Journal of Consumer Research*, *8*(4), 419-430.

McMullin, J. L., Miller, B. P., & Twedt, B. J. (2019). Increased mandated disclosure frequency and price formation: Evidence from the 8-K expansion regulation. *Review of Accounting Studies*, *24*(1), 1-33.

Melessa, S. J. (2012). Monthly Employment Reports and the Pricing of Firm-Level Earnings News. *Dissertation, Duke University.*

Merton, R. C. (1987). A Simple Model of Capital Market Equilibrium with Incomplete Information. *The Journal of Finance*, *42*(3), 483-510.

Miller, B. P. (2010). The effects of reporting complexity on small and large investor trading. *The Accounting Review*, *85*(6), 2107-2143.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*(2), 81.

Noh, S., So, E. C., & Weber, J. P. (2019). Voluntary and mandatory disclosures: Do managers view them as substitutes? *Journal of Accounting and Economics*, *68*(1), 101243.

Patell, J. M., & Wolfson, M. A. (1982). Good news, bad news, and the intraday timing of corporate disclosures. The *Accounting Review*, 509-527.

Peng, L., & Xiong, W. (2006). Investor attention, overconfidence and category learning. *Journal of Financial Economics*, *80*(3), 563-602.

Penman, S. H. (1987). The distribution of earnings news over time and seasonalities in aggregate stock returns. *Journal of Financial Economics*, *18*(2), 199-228.

Segal, B., & Segal, D. (2016). Are managers strategic in reporting non-earnings news? Evidence on timing and news bundling. *Review of Accounting Studies*, *21*(4), 1203-1244.

Sena, J. (2004). 8-K changes: Safe harbor, or false sense of security. *Compliance Week*.

Sims, C. A. (2010). Rational inattention and monetary economics. In *Handbook of Monetary Economics* (Vol. 3, pp. 155-181). Elsevier.

Steffen, T. (2020). The information asymmetry effects of expanded disclosures about derivative and hedging activities. *Yale University Working Paper.*

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*(2), 257-285.

Verdi, R. S. (2005). Information environment and the cost of equity capital. *MIT Sloan School of Management Working Paper*.

Wharton Research Data Services. (2011). WRDS-SEC platform: overview of the filings exhibits data, *https://wrds-www.wharton.upenn.edu/documents/749/Overview_of_the_Filings_Exhibits_Data.pdf*.

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

Biography

Eric Zhihong He was born and raised in Suzhou, Jiangsu Province, China. After completing high school, Eric moved to the United States at age 17. He earned a B.A. in Economics and Mathematics from the University of North Carolina at Chapel Hill in 2012, and an M.S. in Agricultural and Resource Economics from the University of California, Berkeley in 2013. Eric worked as a Research Assistant in the Washington, District of Columbia office of the Public Company Accounting Oversight Board before beginning his Ph.D.

During his time as a Ph.D. student at Duke University, Eric published the article "Does differential taxation of short-term relative to long-term capital gains affect long-term investment?" in *Journal of Accounting and Economics* with Martin Jacob, Rahul Vashishtha, and Mohan Venkatachalam. Eric earned his Ph.D. in Business Administration from Duke University in 2022.