Comprehending Corporate Disclosures: What lies beneath?

Sugato Chakravarty[1]
Krannert School of Management
Purdue University
West Lafayette, IN

Prasad Hegde
Faculty of Business, Economics and Law
Auckland University of Technology
Auckland, NZ

Abstract

This paper explores how the readability of corporate financial disclosures, specifically annual 10-K filings, can impact investor reaction. Prior research has shown that language attributes, such as the use of concrete language, influence message delivery and enhance confidence in the company's communications. In this study, we use a novel proxy for document readability that is free of measurement errors that normally plague other approaches to quantification of specific language attributes in written corporate communication: the XBRL characters. We then examine how firm-specific factors, such as size, investments, profitability, and institutional holdings might moderate the relationship between readability and investor reaction. The findings suggest that 10-K readability is positively related to investor reaction, and small firms, firms that invest lower, firms with lower institutional ownership and firms with lower operating profit, exhibit a greater positive response. The paper contributes to the literature by highlighting the impact of document readability on market efficiency and informing firms on how to improve their disclosure practices. Also, we use a simple, and freely available, document readability measure and analyze carefully crafted written statements instead of oral speech, which can introduce inaccuracies and misclassifications. Overall, this study sheds light on the importance of document readability in corporate communication and its implications for researchers, investors, firms, and regulators.

JEL classification: G10; G39

Keywords: concrete language; firm risk; investor response; readability; file size; XBRL

This version: August, 2023

[1] Correspondence: Sugato Chakravarty, Krannert School of Management, Purdue University. Email: Sugato@purdue.edu. We thank the seminar participants at Ball State University and Purdue University for helpful comments. We thank Tom Brush and Rich Makadok for discussions on the topic. We also thank Steven Cahan for helpful comments.

1. Introduction

A growing body of research, under the umbrella of impression management (IM), has examined the extent of impression guidance prevalent in firms' written communication with investors through their choice of context specific information, usage of phrases to promote specific imagery (Bolino, Kacmar, Turnley & Gilstrap, 2008; Pan, McNamara, Lee, Haleblian & Devers, 2018; Washburn & Bromiley, 2014). The importance of document readability is also reflected in the emphasis placed by the Securities and Exchange Commission (SEC) on providing clear and readable financial disclosures. For example, the Plain English Rule 421(d) requires that issuers adhere to plain English principles in the design of firm prospectuses and also provides both linguistic and formatting suggestions for preparing plain English disclosures that are more readable by the appropriate stakeholders and regulators.

Research in psychology suggests that processing fluency, or an individual's subjective feeling about how easy it is to process information from a more readable financial disclosure, will increase investors' beliefs that they can rely on the information in the disclosure (Shah & Oppenheimer, 2007). A fluent message can lead to more favorable evaluations of the messenger (Oppenheimer, 2006). Similarly, research in strategy has examined the extent of impression guidance prevalent in firms' written communication with investors through their choice of context specific information and usage of phrases to promote specific imagery (Pan et al., 2018). Prior research in the communications literature shows that language attributes are an effective tool for managing audiences' impressions since they influence the quality of message delivery (Bradac, Bowers, & Courtright, 1979; Larrimore, Jiang, Larrimore, Markowitz, & Gorski, 2011; Toma & D'Angelo, 2015). The use of concrete language attributes facilitates investors' information processing and enhances confidence in the company's written communications and legal filings (Toma & D'Angelo, 2015; Larrimore et al., 2011); includes descriptive words that provide specific information and entail context specific, and detailed descriptions of situations (Dechêne, Stahl , Hansen & Wänke, 2010); and is characterized by the use of verbs, numbers, and past focused words in comparison to non-concrete (or, abstract) language characterized by adjectives, nonspecific quantifiers, and the use of future-focused words (Elliott, Rennekamp, & White, 2015; Snefjella & Kuperman, 2015). Research shows that the level of language concreteness in top managers' communication is positively related to investor reaction (Pan et al., 2018).

However, while studies on concrete language attributes focus on the specific words, and language, used in written communication, such as the use of descriptive words and context-specific descriptions, there is another distinct way to gauge the "comprehensibility" of a document: through analyzing its overall "readability". The different elements that can contribute to document readability include features like word choice (choosing of a simple versus a complex word – like using "perturbed" instead of "worried"); relative

use of jargon words; usage of unusual words, sentence length, sentence structure; and the average number of syllables per word (You & Zhang, 2009; Bonsall, Leone, Miller & Rennekamp, 2017). Arguably, such factors help determine whether a piece of written communication is likely to be understood widely or lead to confusion in the minds of many. Therefore, while the aim of examining concrete language attributes is to understand how they might influence message delivery, and enhance confidence in the communication, document readability, on the other hand, focuses on the ease, or difficulty, of understanding a piece of text, based on factors like sentence length, word complexity, and average syllables per word. The goal of readability is to analyze how well a document can be understood by its intended audience which, in turn, can impact market efficiency and investor decision-making. Importantly, while both approaches offer valuable insights into the effectiveness of corporate communication, it is worth emphasizing that they still focus on different aspects of language use and document design (Bonsall et al., 2017).

Consequently, in the current paper, we argue that measuring document readability can provide an alternate, and viable, instrument to capture concrete language attributes in important corporate communication beyond what can be captured through studying concrete language attributes (Pan et al., 2018). The classic readability measure has been the Fog index (Gunning, 1952), determined through a formula driven largely by sentence length and word complexity; and the SMOG Index (Mc Laughlin, 1969) which captures how many years of formal education an average person needs to understand a given document. In the past, the SEC has even considered using such indexes to help identify poorly written financial documents even though they happen to be ill-suited in the domain of business writing (Loughran & McDonald, LM, 2014). LM further recommend that in those situations, where readability includes the ability to assimilate valuation relevant information (such as in corporate10-K filings and earnings-related conference calls), researchers should use the file size associated with the document as a proxy for document readability. This is based on the idea that larger documents, in terms of word count or data size, are likely to be more complex, and therefore potentially less readable. Also, a major advantage of using file size is that it is an easily available, and ready to use, metric, not requiring any significant processing time or cost (LM, 2014).

However, it should be noted that prior to 2012, the SEC allowed financial documents to be filed in ASCII and HTML formats which created relatively small 10-K file sizes. But, beginning in 2012, the SEC mandated that all financial documents be submitted in eXtensible Business Reporting Language (XBRL). XBRL is a freely available and global standard for exchanging business information. It is an open standard, XML-based language for digital business reporting, managed by a global not-for-profit consortium, XBRL International. This language is primarily used to communicate information between businesses and on the internet. The rationale behind using the frequency of XBRL characters in a 10-K document as a proxy for

readability comes from the nature of XBRL itself. XBRL tags provide a standardized method of identifying specific pieces of financial data within a document. The more tags a document has, the more complex and detailed the information it contains may be. Therefore, the frequency of XBRL characters tagged in a financial document could potentially be used as a measure of the complexity of the financial data in the document. A higher frequency of XBRL characters would indicate a more complex, less readable document because it contains more detailed and specific financial data. Conversely, a lower frequency of XBRL characters might suggest a simpler, more readable document. However, this switchover from one kind of language to another has also caused 10-K file sizes to magnify many times in size (going from an average of 1.5 MB in the pre-2012 period to 17.7 MB in the post-2012 or, HBRL, period. Due to this discontinuous jump in file size, and the extremely large file sizes, scholars have questioned the validity of using 10-K file size as a document readability measure in the HBRL period (Hoitash & Hoitash, 2018). In fact, recent research shows that file size as a measure of readability loses its statistical significance in the XBRL period (Cahan, Chang, Siqueira & Tam, 2022).

We, therefore, use the frequency of XBRL characters occurring in a document to investigate the relationship between 10-K readability and market reaction. Additionally, we seek to understand whether prominent firm specific risk factors like firm size, firm investments, institutional ownership and operating profitability, can act as moderators in enhancing, or attenuating, the potential impact of 10-K readability on investor reaction. Based on a theoretical foundation of impression management, and obfuscation theory, we predict the following: Higher level of 10-K readability is positively related to investor reaction. Furthermore, we predict that firm size will moderate the relationship between 10-K readability and investor reaction with greater readability triggering a positive investor reaction only for small firms. We, further, justify that a firm's investments in building its infrastructure, will be an additional moderator of the relationship between readability and investor reaction with readability generating a positive response for firms that make lower investments in themselves. We expect that the proportion of institutional investors in a firm's stock will act as another significant moderator with readability generating a greater positive response for firms with lower institutional ownership of its shares. Finally, we hypothesize that a firm's operating profitability will significantly moderate our main relationship with document readability causing greater positive response for firms with lower profitability. We test our predictions on a large panel data of listed stocks over 11 years (2012 through 2022) and find strong support.

Our paper makes a significant contribution to the strategy literature on a number of dimensions. It highlights the idea that the readability of prominent financial disclosures, like firms' annual 10-K filings, affect market efficiency as investors rely on these disclosures to make informed decisions (Healy & Palepu, 2001) which can then provide insights into whether better readability leads to better investor decision-

making and more efficient markets. Understanding the moderating role of firm-specific factors in the relationship between readability and investor response can inform firms on how to improve their disclosure practices (Li, 2008). For regulators, our study suggests that policies aimed at improving the quality of financial disclosures should take into account the readability of these documents. The SEC's Plain English Rule is a step in the right direction, but further efforts should be made to ensure the highest possible readability of financial disclosures. Regulators should also take into consideration the impact of firm-specific factors on the relationship between document readability and investor reaction when designing policies aimed at improving the quality of financial disclosures. For investors, our study suggests that the readability of financial disclosures can provide valuable information for investment decision-making. Investors are likely to be attentive about the readability of 10-K filings when evaluating potential investment opportunities, as a more readable document is likely to provide more reliable information. Furthermore, our study suggests that the impact of document readability on investor reaction is moderated by observable firm-specific factors such as firm size and the proportion of institutional investors and investors need to factor them in when evaluating the reliability of a firm's financial disclosures.

Our research also goes significantly beyond what has been analyzed by Pan et al. (2018) through a content analysis of earnings call transcripts. While quarterly corporate earnings calls are indeed a significant source of information, another equally, if not more, significant channel for disseminating a company's fiscal position, including its future outlook and sources of expected risk, is through its annual 10-K filings. Also, public companies are required to disclose their economic activities through Form 10-K, making it a significant source of information to investors. According to the Gilmartin Group, the average word count on an earnings call script is about 2,500 words, or about 20 minutes' worth of speaking time. The majority of the companies range from 2,000 to 3,000 words (15 – 25 minutes). In comparison, the average word count of a 10-K is about 49,000. Given this huge discrepancy in word count alone, it appears that an investigation of the extent of investor reaction as a function of its 10-K readability represents a significant step forward in adding to our body of knowledge on strategic corporate communication. Another important contribution of our paper is to employ a document readability measure that both easily obtained and does not suffer from measurement errors that is associated with other textual procedures (Loughran & McDonald, 2014; Jennings, Kim, Lee & Taylor, 2023). Furthermore, we analyze carefully crafted, and extensively vetted, written corporate statements in place of relatively informal oral delivery used in quarterly earnings calls. As McCroskey, Wrench & Richmond (2003) articulate, there are important differences between oral speech versus written reports, which includes oral delivery having smaller variety of words, more colloquial and nonstandard words, more qualifying statements, fewer quantifying terms, or precise numerical words, more extreme and superlative words, etc., all of which can introduce inaccuracies and misclassifications during transcription into written form which is then analyzed by researchers. For

example, the American Hospital Association reports that presence of multiple environmental factors results in significant transcription errors and wrong medications being dispensed as a result (AAO, 2015). In contrast, documents like the annual 10-Ks are carefully crafted, laboriously compiled and vetted, where the exact words employed in the document are chosen carefully.

2. Theory and Hypotheses

2.1 Impression management and the theory of obfuscation

Publicly traded corporations report their state of affairs, various risk factors, and future outlook using diverse channels including written (and narrative heavy) filings like the annual 10-K document (Courtis, 2004; Westphal & Graebner, 2010; Westphal & Zajac, 1998). Such narrative communication is impaired unless there is clarity in writing since distorted and confusing writing hinders a reader's fluent comprehension of intended messages, especially if it uses unnecessary and unfamiliar polysyllabic words and long sentences. Polysyllabic words do serve a useful role when they enable a compression of words and the expression of complex or ambiguous concepts; therefore, any impairment is likely to impede a reader's ability to accurately interpret these messages and, thereby, hinder belief formation and revision (Courtis, 2004; Loughran & McDonald, 2016). The term "obfuscation" is used to describe a narrative that is (sometimes deliberately) vague and unclear and obscures the intended message, or confuses, distracts or perplexes readers. When employed strategically, such opportunistic managerial behavior, assumed to exist at the root of impression management, has given rise to the so-called obfuscation hypothesis, based on the assumption that management is not neutral in its presentation of corporate narratives (Sydserff & Weetman 1999: 460). The resulting predisposition, which arises from strong economic incentives and the opportunity to provide a biased view of corporate performance, leads to management obfuscating failures and emphasizing successes (Adelberg 1979: 187). Obfuscation can also occur through the use of esoteric or obscurantist vocabulary and/or gobbledegook, extraneous and non-relevant information, long sentences with complex grammatical structures and/or high variability in reading ease and convoluted and/or spurious argumentation. Furthermore, linguistic complexity can conflate the information component underlying a narrative and an obfuscation component which can be intentionally used by corporate leadership to reduce the informative component of the disclosure (Bushee, Gow & Taylor, 2018). Importantly, obfuscation is viewed as a form of impression management or perception engineering (Harris, 2000; Courtis, 2002).

The theory of impression management (Goffman, 1959) has shown that organizations strategically seek to influence stakeholder reactions (Leary & Kowalski, 1990; Merkl-Davis, Brennan, Mcleay, 2013; Graffin, Carpenter & Boivie, 2011). Impression management research usually assumes that "*sections of the [annual] reports are allegedly managed so as to present management in as favorable light as possible*" (Stanton et al. 2004: p. 57) and organization leaders actively manage their firm's informational environment

in ways they hope will favorably affect the impressions of targeted stakeholders (Puffer & Weintrop,1991; Zajac & Westphal, 1995; Bolino, Kacmar, Turnley & Gilstrap, 2008). One specific avenue of impression management research involves examining how and when top managers use specific content to explain or justify firms' activities and qualities to avoid harm to their firms' image (Elsbach, 2003; Fiss & Zajac, 2006; Salancik & Meindl, 1984) which converges into strategic obfuscation. Examples include internal-external attributions (Staw, McKechnie, & Puffer, 1983), obfuscation (Li, 2008), and discretionary disclosure strategies (Davidson, Jiraporn, Kim, & Nemec, 2004; Merkl-Davies & Brennan, 2007). An important conclusion from this line of research is that such impression management tactics tend to assuage stakeholders' concerns and serve as an effective tool for managing audiences' impressions because they influence the quality of message delivery (Bradac, Bowers, & Courtright, 1979; Larrimore, Jiang, Larrimore, Markowitz, & Gorski, 2011; Toma & D'Angelo, 2015). However, most of the research effort in IM has focused on the informational content of the messages and paid less attention to the hidden language attributes inherent in corporate communication with shareholders and regulators (Pan et al., 2018) whose use enhances the recipient's confidence in the veracity of the communication including any written filings (Toma & D'Angelo, 2015; Larrimore et al., 2011). Typically, in this body of work, concrete language refers to descriptive words that provide specific information and entail context specific, and detailed descriptions of situations (Dechêne, Stahl, Hansen & Wänke, 2010; Langacker, 1987) and encompass almost every aspect of language including phonology), rules and structure of sentences (syntax), and words (lexicon) (Hosman,2002). For example, verbs are more concrete than adjectives (Assilaméhou, Lepastourel, & Testé, 2013); quantitative numbers like digits are more specific than nonspecific quantifiers such as "many" or "few" (Jerez-Fernandez, Angulo, & Oppenheimer, 2014); past-focused language is considered concrete, while future-focused language cannot be factually vetted (D'Argembeau &Van der Linden, 2004; Snefjella & Kuperman, 2015).

2.2 Document readability as a proxy for obfuscation

We explore an alternative way to examine concreate language attributes through analyzing the overall readability of a financial document which characterizes how easy, or difficult, it is to comprehend and truly understand a piece of text. Different elements can contribute to document readability, like word choice (choosing of a simple versus a complex word – like using "perturbed" instead of "worried"), relative use of jargon words and unusual words, sentence length, sentence structure and the average syllables per word (Bonsall, Leone, Miller & Rennekamp, 2017). Such factors help determine whether a piece of written communication is likely to be understood widely or lead to confusion in the minds of readers, whoever they may be. A classic readability measure is the Fog index (Gunning, 1952) that is determined through a formula driven largely by sentence length and word complexity and the SMOG Index (McLaughlin, 1969)

which captures how many years of formal education an average person needs to understand a given document. Li (2008), for example, uses the Fog index and reports that the annual reports of firms with lower earnings are harder to read (i.e., show higher Fog index scores) and firms with annual reports that are easy to read (greater readability) show persistently positive earnings. In the past, the SEC has even considered using such indexes to help identify poorly written financial documents within the domain of business writing (Loughran & McDonald, 2014). LM further recommend that in those situations where readability includes the ability to assimilate valuation relevant information (such as in corporate10-K filings and earnings-related conference calls), researchers should use the file size associated with the document which is an easily available and ready to use metric and does not require any non-trivial processing time and cost. However, recent study by Cahan et al. (2022) shows that file size as a measure of readability loses its significance after the mandatory adoption of eXtensible Business Reporting Language (XBRL) from 2012 (the XBRL period) by SEC.[2] The authors argue that the size and composition of 10-K filings have fundamentally changed since the XBRL period. As a result, the average file size has increased significantly making it unsuitable to measure readability in the XBRL period. Therefore, we use XBRL as the readability metric in examining the relationship between corporate clarity in communicating with the investment public and the market's response. If we assume that greater document readability implies clarity in communication between a firm and its stakeholders, we expect that arguably one of its most significant corporate communications with its stakeholders, its annual 10-K filing, should display the following characteristic around its publication:

H1: ***Higher levels of 10-K readability is positively correlated with investor reaction.***

3. The impact of firm specific risk factors on investor response to 10-K readability

Purchasing stocks of small firms are considered to be risky investments since there is a significant likelihood for financial loss (Fama & French, 1993; Jegadeesh & Wu, 2013). For example, from 1997 through 2012, the Russell 2000 index (index of small company stocks) returned 8.6% on an annualized basis compared to 4.8% for the S&P 500 (index of relatively large companies). Yet, over the same period, the Russell 2000 had significantly higher volatility (measure of risk). This is partially due to the relative lack of available information on such companies, fewer, if any, coverage by stock analysts, and having low trading volume (Fisher, 1966; Scholes & Williams, 1977; Conrad, Kaul & Gultekin, 1991; Bamber, 1987;

[2] While 2009 is the first year of mandatory XBRL adoption, 2010 is the first full adoption year. However, the total number of XBRL adopters are relatively small till 2012.

Collins, Li & Xie, 2009; Hirshleifer & Teoh, 2003). Research has shown that both firm size and its book-to-market ratio are proxies for pervasive stock specific risk (Chen, Chang, Yu, & Mayes, 2005; Jegadeesh & Wu, 2013). Also, in a series of influential papers, Fama & French (1992; 1996; 2015; 2016) show that risk factors like firm size and firms' book-to-market ratio significantly explain stock returns, with smaller firms posing higher risk, and firms with lower book-to-market ratio posing higher risk than firms with higher book-to-market ratio. Given the higher inherent risk of small companies, stakeholders of such companies will pay particular importance to the readability of their 10-Ks. This implies that small firm 10-Ks with higher readability is associated with positive investor reaction. In contrast, we expect the relationship between readability and investor reaction to weaken for larger firms because larger firms tend to have longer 10-K documents (Li, 2008) and disclose 'more' redundant information through footnotes and tend to repeat disclosures across various sections because of their operational complexity (Cazier & Pfeiffer, 2016). Given that SEC permits a significant amount of disclosure to be incorporated through references to the notes to the financial statements, larger firms are likely to have higher discretionary disclosures because of varied business operations. For example, General Electrical (GE) company's recent 10-K document is 361 page long, similarly Berkshire Hathaway Inc has a 10-K report that spans over 133 pages. These companies fall in the large capitalization category. Furthermore, Chakrabarty, Seetharaman, Swanson and Wang (2018) highlight that larger firms such as GE or JP Morgan Chase and Co. typically have longer 10-K documents causing information overload to ordinary investors. On the other hand, small capitalization firms such as Delek US Holdings Inc. has an annual report of 82 pages. This evidence suggests that firm size is a significant determinant in understanding the 10-K report size. Thus, we posit that investors discount the negative effects of poor readability or redundant information for larger firms (relative to small firms) that can be attributable to firm complexity. Formally stated:

H2A: Firm size decreases the effect of 10-K readability on investor reaction.

Additionally, a bedrock theory in strategy, the resource-based view of the firm (Cyert & March 1963), posits that firms achieve competitive advantage from value creation through firm-specific investments held by key stakeholders: employees, suppliers, and customers. Value creation can also be accomplished through firms' growth is related to the quality of its internal corporate governance (Andries, Balutel, Ihnatov, & Ursu, 2020); it can also come about through cost cutting and safe investments at the expense of risk-taking ventures (Lovallo, Koller, Uhlaner & Kahneman, 2020). By the same token, when a firm reduces the level of investments in itself year-to-year, it could signal darker days in terms of lost business and other growth opportunities. Therefore, when a firm increases investments in itself, investors are concerned about the change. Investors would be particularly interested in identifying the precise areas of risk reduction and growth in order to be able to accurately price the true value of the stock. Therefore,

investors would want to closely scrutinize the firm's annual 10-K filing which would imply an increased importance of document readability as a function of the change in investments. Formally:

H2B: A higher investment in itself decreases the effect of readability on investor reaction.

Firms have a mix of shareholders between individual or retail investors and institutional investors who hold large blocks of shares in given companies. Institutions are also considered to be smart investors (Chakravarty, 2001; Badrinath, Kale, & Noe, 1995; Sias & Titman, 2006; Chen, 2006; Ke & Petroni, 2003) since they invest millions of dollars in shares and therefore are likely invest money to uncover valuable information on the companies, they invest in. In contrast, individual investors who may hold on average a few hundred shares of a given company do not have any information on the firm's prospects and are generally liquidity providers in the market, earning negative returns (Barrot, Kaniel & Sraer, 2016). Badrinath et al. (1995) find that the past returns on stocks held by informed institutional investors is positively correlated with the contemporaneous returns on stocks held by uninformed retail investors which implies that retail investors' investments follow the prior institutional investments in the same stocks. Research has shown that institutions have a tendency to buy stocks that are overvalued (low BM ratios) and sell stocks that are undervalued (Edelen, Ince, Kadlec, 2016). On the other hand, a higher proportion of institutional ownership in a firm also conveys a form of certification of the company or the institutional seal of approval since the intuition is that if institutions are risking investing millions of dollars in shares of the company it cannot be a bad investment. Conversely, when the company's shares are held in higher proportion by retail investors than institutions, this certification effect will be greatly diluted, and investors correspondingly will be more interested in closely understanding the underlying risks of such companies. This argument is consistent with the previous research that documents that less readable disclosures attract higher demand for analyst research and lower investor trading (Miller, 2010; Lehavy, Li & Merkley, 2011) Thus, based on these arguments we posit that a highly readable 10-K filing from such firms (with higher proportion of individual investor ownership) will therefore have a positive investor reaction. Or in other words, firms with higher proportion of institutional investors will have a negative investor reaction. Formally:

H2C: A higher proportion of institutional investors in a company's shares relative to retail investors decreases the effect of readability on investor reaction.

Companies may strategically manipulate the readability of their annual 10-K reports when they have poor performance to report. Particularly, in these circumstances, they might obfuscate their 10-Ks to create information asymmetry. By making the report more difficult to read, they can potentially discourage investors, analysts, and the public from fully understanding the extent of their financial difficulties or risks.

Similarly, if a company is not performing well, making the 10-K harder to read may prevent investors from quickly understanding the severity of the situation, thus avoiding panic selling of the company's stock. Additionally, companies may also wish to buy time to implement turnaround strategies by making their 10-K reports less readable when they're not doing well. This could potentially reduce the immediate negative impact on their stock price while they're working on improving their situation. By the same token, when they have relative profitability to report, they might want to make their 10-Ks more readable to ensure that their positive financial results are well-understood and widely recognized. However, investors may be less interested in closely perusing a 10-K when the company is making profits and 10-K readability will be less of an issue. There are several reasons for this. For instance, when a company is consistently profitable, investors may perceive it as a lower risk. This can lead to complacency, with investors assuming that the company's management is competent and that their investment is secure, reducing the perceived need to scrutinize the 10-K report (Shefrin, 2002). Investors may exhibit a positive bias towards profitable companies, believing that good past performance indicates good future performance (De Bondt & Thaler, 1985), which could then lead them to overlook potential issues highlighted in the 10-K (Barberis & Thaler, 2003). Also, since the profitability of a company can be an easy-to-understand indicator for some investors, they might feel less compelled to delve into the complexities of the 10-K report. Finally, there is potential for investors to ignore 10-K filings (You & Zhang, 2009) and this is especially true when a company is profitable. In such circumstances, investors might rely more heavily on third-party analyst reports (Irvine, 2004), which typically provide a summarized interpretation of the financial state of the company, thus saving the investors the effort of going through the entire 10-K.

Past research on corporate disclosures (through filings like 10-Ks) has focused mainly on the amount of disclosure made by firms (Healy & Palepu, 2001) – while ignoring the lexical properties of these filings (Li, 2008). Specifically, using the Fog index as a proxy for document readability, Li (2008) finds that the annual reports of firms with lower earnings per share are harder to read (higher Fog index values) and are longer, i.e., more XBRL characters. By the same token, firms with annual reports that are easy to read (lower Fog index) show higher and more persistent positive earnings (or higher operating profit). Schrand & Walther (2000) argue that disclosure readability and lexical features may be used strategically by managers, depending on the what they are trying to trumpet or obfuscate. Courtis (1998) reports that firms manipulate disclosure transparency by reducing clarity when they are performing poorly (Rutherford, 2003). We extrapolate from the above discussion the notion that decreases in company profitability, accompanied by the potential of decrease in 10-K readability, are likely to increase the importance of 10-K readability as investors wish to better understand the company’s positioning relative to specific risks or opportunities. This leads us to our next hypothesis:

H2D: We should see a significant impact of readability on investor reaction when a firm's operating profitability decreases relative to the previous year.

4. Methods

4.1 Sample

For our analyses, we utilize the firms listed on NYSE, AMEX, and NASDAQ with share codes 10 and 11 on CRSP (i.e., only common stocks) from the years 2012 through 2022. The unit of analyses is firm-year observations. We obtain the readability measures from SeekEdgar database and from Prof. Tim Lougran's website. SeekEdgar LLC provides the annual readability measures for all firms that file their 10-K with SEC. We obtain the Bog index data from Prof. Brian Miller's website The explanatory and control variables are derived from the Compustat, CRSP, and Thomson Reuters 13f databases. In addition, consistent with prior studies (Jegadeesh & Wu, 2013; Loughran & McDonald, 2011) we employ the following rules to construct the final sample for analysis.

- We consider only the initial 10-K filing in a given year by a given firm. In other words, we exclude subsequent amendments to the initial 10-K filed in the same year since most information is reported in the initial filing.
- Since SEC EDGAR identifies the firms through Central Index Key (CIK), we use the WRDS CRSP-COMPUSTAT database to match the CIKs in EDGAR and WRDS. In other words, we exclude all firms for which we cannot match CIKs between the two databases.
- Our analyses use various accounting and financial variables such as market capitalization and stock returns. We exclude all firms for which such data are not available.
- We require that the sample firms have a closing price of at least $5 on the filing date. Eliminating low-priced firms is standard in finance literature to reduce the bid-ask bounce's impact on prices. We also exclude firms below $100 million market capitalization to eliminate micro-cap stocks.

The above filters, and the requirement of matching the firms with CRSP/COMPUSTAT, reduce the original sample to 13,256 firm-year observations. We should note that, by comparison, Cahan et al. (2022) report a firm-year sample of 17,341, however they do not exclude the micro-cap stocks resulting in slightly more observations.

4.2 Dependent variable: Cumulative abnormal returns

Consistent with the prior research in strategy, we capture investor reaction by adopting cumulative abnormal returns (CAR) adjusted for CRSP value-weighted returns (Pan et al., 2017). A common approach in observing investor reaction is through examining the abnormal returns during the announcement window

(i.e., abnormal returns during t-*1* to t+*1*) of the 10-K filing. The announcement returns implicitly assumes that the stock price quickly incorporates the information from the 10-K documents. However, previous studies suggest that investors tend to underreact to new information (e.g., Bernard & Thomas, 1990)), especially complex information (e.g., You & Zhang, 2009). Specifically, You & Zhang (2009) show that investors and stock analysts generally ignore certain important information around the announcement dates, that could also be exacerbated by the complexity and lack of immediate comprehension (readability) of the information. The authors argue that investors often view the filing of 10-K as a formality and largely ignore it in the immediate aftermath of the earnings announcement. Consistent with this insight, the authors do not find any significant investor reactions immediately around the filing date. They attribute this phenomenon to investor under-reaction and show that 10-K filings contain useful information about firms' future performance. As a result, the authors suggest that investor reaction is appropriately captured through 3-, 6- and 12-month cumulative abnormal returns. In another recent study, Cohen, Malloy & Nguyen (2020) suggest that investors are inattentive to these simple changes across the universe of public firms. They document a significant future return that accrues until about 18-months from the reporting date. Similarly, Loughran & McDonald (2017) show that the average publicly traded firm's 10-K is downloaded only about 28 times immediately following the filing, indicating a lesser investor reaction. Womack (1996) shows that the market reaction to analysts' recommendations is incomplete and slow. Womack finds significant positive (negative) stock returns following new buy (sell) recommendations over 6-months after the recommendations are issued. The implication of this body of work is that, in order to accurately capture investor response to the contents of a 10-K, we need to consider an extended time window after the actual filing. Therefore, we capture investor reaction associated with a specific company's 10-K by computing the CAR over a 3-month window (*t*+3) beginning with the 10-K filing month (month *t*).

4.3 Main explanatory variable: Document readability

We define the readability of 10-K documents, READ, as the natural logarithm of the number of XBRL characters in each 10-K document following the recommendation of Hoitash and Hoitash (2018; 2022), Hoitash, Hoitash and Morris (2021) and Cahan et al. (2022). Particularly, Cahan et al. (2022) argue that the size and composition of 10-K filings have fundamentally changed since the XBRL period. As a result, the average file size has increased significantly making it an unsuitable to measure readability in the XBRL period. Earlier works on readability have predominantly use Gunning Fog index as a measure of readability. According to the Gunning Fog index, a document is harder to read (i.e., less readable) when there are more syllables per word or, more words per sentence. Using the Fog index, Li (2008) finds that firms with lower reported earnings tend to have annual reports that are harder to read (i.e., high Fog index values). Following Li (2008), other researchers have used the Fog index as a measure of document

readability in finance and accounting (see for example, Biddle, Hilary, & Verdi, 2009; Guay, Samuels, & Taylor, 2016; Miller, 2010). However, Loughran & McDonald (2014) also argue that, since the Fog index naturally considers complex words as those containing more than two syllables, it is problematic in the context of financial documents and transcriptions as many financial terms with two or more syllables are not considered as complex words to investment professionals and policymakers. This makes Fog index ineffective in capturing the readability of 10-Ks. They further recommend that 10-K file size is a good measure of document readability since larger 10-K files are consistently more difficult to read. It should, however, be noted that Loughran and McDonald (2014), in arguing for file size as a readability proxy, examined the data over 1994-2011. But, over this time period, firms could file their 10-Ks in either ASCII or HTML formats resulting in smaller file sizes on average. But, as Cahan et al., (2021) observe, that average 10-K file sizes have grown exponentially in the post 2012 period when the SEC started requiring firms to file their financial statements in XBRL format instead. For example, the average 10-K file size has increased significantly between the pre- and post- 2012 periods (Cahan et al, 2021). A major reason for this increase in file size, according to the SEC (2017), is that the 10-K files now include both an XBRL component submitted by the filer and an XBRL component created by the SEC from the filer-submitted XBRL component. Importantly, given this significant discontinuity in average 10-K file sizes in the two time periods, it is not clear if file size is still a reasonable proxy for document readability post 2011 (Hoitash & Hoitash, 2018; Guo & Yu, 2020).

XBRL has advantages over other data sources because financial reports can be read consistently, and efficiently, extracted by computers, without any human intervention (Hoitash & Hoitash, 2018). Consequently, classification errors and other omissions that occur in any traditional document detail identification routines used in identifying concrete language attributes or overall readability using word/sentence characteristics, are eliminated. XBRL makes business and accounting reporting dynamic and accurate (Guo & Yu, 2020), and it is often referred to as "bar codes of reporting" because the digital reports are easy to access, efficient and standardized (see https://www.xbrl.org/).[3] Since our goal is to capture that relationship between document readability and investor reaction, we use the total number of XRBL characters proxy for our readability measure, READ. Particularly, we define readability as the negative of the log of XRBL characters (i.e., Log XBRL * -1). The multiplication by -1 is done purely for ease of coefficient interpretation.

4.4 Control variables

[3] For a detailed explanation about XBRL reporting, please follow the article found here: https://www.xbrl.org/the-standard/what/an-introduction-to-xbrl/

We also controlled for additional factors that could influence our dependent variable, including variables related to the content of 10-K documents, firm-level and industry-level variables. We controlled for firm size and firm performance since they can influence investors' reactions to firm communication, because larger firms tend to have more to disclose (Hendricks, Singhal, & Zhang, 2009; Pan et al., 2017). Firm size is captured through the logarithm of total assets (ASSETS) at the end of the fiscal year (Audia & Greve, 2006; Benischke, Martin & Glaser, 2018; Chakravarty & Hegde, 2019; Ferris Jagannathan & Pritchard, 2003; Fich & Shivdasani, 2006). Additionally, we incorporated the return on assets (ROA) to account for firm performance where ROA is computed as the earnings before interest, taxes, depreciation, and amortization divided by the book value of assets (Coles, Daniel & Naveen, 2008; Fich & Shivdasani, 2006; Pan et al., 2018). Additionally, we controlled for other market measures that could potentially affect investor reaction: TURNOVER, defined as the number of shares traded during the period from day -252 to day -6 before filing date (day 0) divided by the number of shares outstanding on filing date (Jegadeesh & Wu, 2013). We include return volatility (VOLATILITY) to capture the overall risk of firms, where volatility is defined as the standard deviation of the daily returns estimated using 60 prior months of data ending as of the end of the month before the filing month (Huang, Teoh & Zhang, 2014; Jegadeesh & Wu, 2013). To control for financial risk, we include, LEVERAGE, defined as the ratio of total debt to book value of assets, since previous studies argue that capital structure (or financial leverage) is an important determinant of performance (Chen & Zhao, 2006; Lang, Ofek & Stulz, 1996; Baker & Wurgler, 2002; Frank & Goyal, 2009; Bhagat & Bolton, 2013 and Fich & Shivdasani, 2006).

Further, we control for firm age (AGE), defined as the number of years since the company's first trading date from CRSP (Loderer & Waelchli, 2010; Khanna & Palepu, 2000; Fich & Shivdasani, 2006), since older firms are generally value firms. Also, previous studies have emphasized that firms typically do not report balance sheet information necessary to compute accruals during their preliminary earnings announcements. For such firms the market first receives information about accruals through their 10-Ks (Chen, Defond, & Park, 2002; Jegadeesh & Wu, 2013). We therefore include the measure, ACCRUALS, defined as a one-year change in current assets excluding cash minus change in current liabilities excluding long-term debt in current liabilities and taxes payables minus depreciation divided by average total assets (Sloan, 1996). Additionally, to account for the effects of communication content on investors' reactions, prior studies in finance and accounting suggest that positive and negative tone of financial disclosures (particularly, 10-Ks) influences investors' evaluations and reactions (Davis, Piger, & Sedor, 2012; Feldman, Govindara, Livnat, & Segal, 2010). Accordingly, we included the fraction of positive (POSTONE) and negative words, (NEGTONE), in our regressions where POSTONE (NEGTONE) is calculated through the ratio of total number of positive (negative) words divided by the total word count in a 10-K document (Loughran & McDonald, 2011; Pan et al., 2018). Finally, to control for variations regarding readability

used across industries, we included the 2-digit SIC level industry dummies over the 12 industries represented in our data (Fama and French, 1997; Benischke, Martin & Glaser, 2018).

4.5 Moderating variables

To test the effect of readability on investor reaction under different levels of observable firm characteristics, we used five commonly observed measures:

(1) Firm size (SIZE): Firm size can significantly influence the amount and types of resources available for financial reporting and the complexity of financial disclosures. Larger firms are more likely to have greater resources and complex operations, which could lead to more detailed and potentially less readable financial disclosures (Lang & Lundholm, 1993). Furthermore, investors may respond differently to the readability of financial statements of large firms compared to small firms due to differences in information asymmetry and the potential for market manipulation (Bamber et al., 1997). Research has shown that SIZE proxies for stock specific risks (Chen, Chang, Yu, & Mayes, 2005; Jegadeesh & Wu, 2013; Fama & French, 1992; 1996; 2015; 2016). We operationalized SIZE as logarithm of market-value at the end of fiscal year. To do so, we use the end of year market value from the CRSP database.

(2) Firm's investments in Itself (INV): Firms that invest heavily in themselves (e.g., research and development, capital expenditures among others) often have higher growth prospects and more uncertain future cash flows. This can make it more challenging for investors to assess the value of these firms (Lev & Sougiannis, 1996). The readability of 10-K filings may play a crucial role in how investors evaluate firms with varying levels of investments, as clear and understandable disclosures can reduce the uncertainty associated with these firms and improve investor decision-making. INV is computed as the change in total assets from the fiscal year ending in year *t-2* to the fiscal year ending in *t-1*, divided by *t-2* total assets, computed annually at the end of each June. We operationalized INV through the Compustat data item – 6 (item name – AT).

(3) The Fraction of Institutional Ownership in the Firm's Stock (IO): Institutional investors often have greater resources and expertise to process and analyze financial information compared to individual investors (Sias, 2004). The fraction of institutional ownership in a firm's stock may influence how readability affects investor response, as institutional investors may be better equipped to understand complex financial disclosures. Moreover, the presence of institutional investors can affect the overall market reaction to financial disclosures, as they are known to be more sophisticated and less prone to behavioral biases (Gompers & Metrick, 2001). Prior research has shown that institutional investors are smart and informed (Chakravarty, 2001; Badrinath, Kale, & Noe, 1995; Sias & Titman, 2006), whereas retail investors are less informed

about the companies, therefore we posit that IO has an effect on the document readability and investor reaction relationship. We use 13-F filings data to obtain the institutional ownership.

(4) Change in Operating Profitability (OP): Profitability is a key factor that investors consider when making investment decisions. For example, Novy-Marx (2013) argues that operating profit is a cleaner accounting measure to capture the firms' true economic profitability. The readability of 10-K filings may influence investor response differently for firms with varying profitability levels. For example, investors may scrutinize financial disclosures more closely for firms with lower profitability to better understand the reasons for the poor performance, and clear disclosures can facilitate this process (Biddle et al., 2009). We operationalized OP in a similar spirit as Novy-Marx (2013) and Fama & French (2015), that uses earnings and free cash flows of the firms. Profitability is defined as revenues minus cost of goods sold, minus selling, general, and administrative expenses, minus interest expense all divided by book equity through the Compustat data items [SALE - (COGS - XSGA - XINT)] divided by BE [BE = sum (AT, -LT, -PSTKL, TXDITC, DCVT]. For our fourth moderator we define OP dummy as 1 when the change in profitability from the fiscal year ending in year *t-1* to the fiscal year ending in *t* is negative and zero otherwise.

5. Results

To evaluate the effects of readability on investor reaction, we used Generalized Estimating Equation (GEE) estimation. In general, GEE estimates are consistent, unbiased, and asymptotically normal and the estimation procedure controls for unobserved firm heterogeneity in longitudinal data (Ballinger, 2004; Liang & Zeger, 1986). GEE, and related GLM, estimations are a popular tool to analyze panel data (Pan, McNamara, Lee, Haleblian & Devers, 2018; Hallen & Pahnke, 2016; Henderson, Miller, & Hambrick, 2006; Quigley & Hambrick, 2012). These estimation procedures are used extensively toward pooled data with firm-year observations (Hardin & Hilbe, 2008; Stillman, 2003; McCullagh & Nelder,1989; Benischke, Martin & Glaser, 2018). Therefore, GEE estimation is appropriate as it allows us to specify any degree of interaction effects. Furthermore, GEE has the advantage of producing robust standard errors that are corrected for over dispersion (Ballinger, 2004; Krause, Filatotchev, & Bruton, 2016; McCullagh, 1984). Overall, GEE is more flexible and better suited for analyzing the data (McCullagh & Nelder,1989). Consistent with previous studies (Pan et al, 2018 and Benischke, Martin & Glaser, 2018) we specified our GEE model with the Gaussian distribution and specified the identity function as the link function which simply means that expected value of the dependent variable is just the linear predictor, similar to a linear

regression model. We do so since we are interested in estimating a linear relationship with the un-transformed dependent variable, CAR.

Table 1 presents the descriptive statistics (Panels A and B) and pairwise correlations (Panel C) for the variables included in our model. Table 2 presents the summary of the theoretical predictions of the effect of the moderators on readability and investor reaction. Table 3 reports the GEE model results, where the Control Model provides the baseline that tests the main effect of document readability on investor reaction. Models 1 - 5 each test the interactions of document readability with firm size, investments, institutional ownership and OP, respectively. Model 6 adds all the hypothesized interactions.

From Panel A of Table 1 we find that the average CAR is 5.3% in our sample and an average 10-K document has about 3.4 million XBRL characters. We also find that the mean (median) market value of firms is $9,831M ($1,526M) suggesting that majority of the firms are mid-and small capitalization according to the market capitalization cut-offs from Financial Industry Regulatory Authority (FINRA).[4] On average, the firms have an ROA of -0.2 and are about 26 years old. We also find that firms have a mean institutional ownership of 70%, suggesting that a significant portion of firms have a majority of institutional shareholders in their shareholder pool which is not surprising given that institutions buy stocks in large blocks relative to retail shareholders. Panel C of Table 1 reports the pairwise correlations across the variables used in our analyses. We find that CAR and READ have a positive correlation, indicating that investor reaction is positive for firms with higher readability. We also find that CAR has a positive correlation with TURNOVER and VOLATILITY, suggesting that higher investor trading is positively associated with investor reaction. On the other hand, ASSETS, INV and IO have a negative correlation with CAR indicating that larger firms and firms with higher institutional ownership have lower investor reaction. Because larger firms (proxied by higher asset size) are less risky compared to small firms, thus earn lower investor return (Banz, 1981, Keim, 1990; Fama & French, 1993; 2015; Horowitz, Loughran & Savin, 2000). Also, firms with high INV have uncertain future cashflows resulting in lower investor returns (citation). Correspondingly, institutions are more likely to actively monitor and oppose management, therefore disagree more relative to retail investors, thereby causing a negative impact on returns (Brickley, Lease & Smith, 1988; Cornett, Marcus, Saunders & Tehranian, 2007; Ferreira & Matos, 2008). While the pairwise correlation values, ranging from 0.3 to -0.6, are below the generally acceptable threshold of |0.7| (Dormann et al. 2013; Toubiana & Maruenda, 2021), additional post regression diagnostics revealed that the VIF scores remain well below the acceptable threshold of 10 (Cohen, Cohen, West & Aiken, 2003). Specifically,

[4] FINRA provides the suggestive range for market capitalization classification under this article: https://www.finra.org/investors/insights/market-cap

we find the mean VIF among the explanatory variables is 1.25, ranging from 1.01 to 1.76. Overall, the pairwise correlations and the VIF scores indicate that multicollinearity is not a concern in our specifications. Table 2 summarizes our hypotheses and their empirical support.

Table 3 presents the main results. We begin with our Control Model in Col. (1), which tests our baseline hypothesis (i.e., H1) related to the effect of readability on investor reaction along with other the control variables. Models 1-6 are presented in Cols. (2) – (6). In Models 1-4, we introduce our moderating variables one at a time. In Model 5 in Col. (6), we include the complete model with all interactions. We note that all models are highly significant, based on the Wald χ^2 test, suggesting that the right-hand side (RHS) variables are collectively significant statistically.

Specifically, H1 predicts that firms that have a highly readable 10-K document should engender positive investor reaction. Specifically, from Table 3 – the control, or baseline, model, we find that the coefficient for READ is positive and significant (b = 6.156, p = 0.000), suggesting that document readability is positively associated with investors' reaction thereby supporting H1. To provide perspective about the economic significance of the coefficient for READ, the difference between a very poorly written 10-K (-2 SD) and a very well written 10-K (+2 SD) would result in an annual abnormal return difference of about 16%.[5] This finding translates to an improvement in market value of $ 1.2 B [i.e., 12.31% of $9.83 B] when firms use a highly readable 10-K document (+2 SD) as opposed to poorly written 10-K document (-2 SD). Given that the firms in our data have an average market value of just under $9.83 (from Panel A of Table 1). Put differently, our finding implies that a 10% increase (decrease) in average XBRL characters (i.e., 10% of 3.4 M or 0.3 M) translates to a decrease (increase) in CAR of 1.13%. This relationship is also confirmed in our Model 6 of Table 3, where we find that the coefficient for readability is positive and significant (b = 4.03, p = 0.000) suggesting that a low level of readability (−2 SD) to a high level of readability (+2 SD) would result in an annual abnormal return change of about 8% and it will generate, on average, a $790 M more favorable market reaction.

Hypothesis 2A predicts that the effect of readability on investor reaction will be negatively moderated by the firm size, with a stronger effect for smaller firms. Shown by Model 1 in Table 3, readability and firm size have a negative interaction effect (b = −1.591, p = 0.000). Therefore, investors react positively to higher/better readability for smaller firms but react more positively for lower readability in large firms. This relationship is also supported in our interaction plot (Figure 1). In other words, the results suggest that we find a strong positive reaction for smaller firms, however such benefits of better readability erode as the firm size increases. Thus, Hypothesis 2A received support. For smaller firms (-1

[5] This is computed by multiplying the coefficient 6.156 with the standard error (SE) of READ from Panel A of Table 1 (i.e.,0.5) for +- 2SE [i.e., Coefficient * (±2*SE), translating into 6.156 x 0.5 x 4 = 12.31%].

SD), the relationship strengthens, and a highly readable 10-K document (+2 SD) results in a CAR increase of 9.84% or about a $967 M improvement in market value. Whereas for larger firms (+1 SD), the relationship weakens, and a highly readable 10-K document (+2 SD) results in only a $483 M improvement in market value, or a 4.92% increase in CAR. This effect of simple slope tests is usually used for the interpretation of the interaction effect of two continuous predictor variables (see Aiken & West,1991; Cohen & Cohen,1983). Thus, firm size as a moderator attenuates the effect of readability on investor reaction.

Hypothesis 2B predicts that the effect of readability on investor reaction will be negatively moderated by the firms' investment. We posited that the market reaction associated with the level of readability will be contingent on firms' level of investments. Model 2 shows that the interaction effect between readability and investments is negative (b = -0.098, p = 0.47). As we expect, investors reacted negatively to higher readability when firms' investment level is high, but more positively to lower readability when the firms' investments is low. As such, for firms with lower investments (- 1 SD), this finding translates to a CAR increase of 1.10% or $90 M increase in market value when the firms' readability is high (+2 SD) as opposed to lower readability (-2 SD). In contrast, for firms with high investments (+1 SD), the relationship weakens, and a higher readability (+2 SD) as opposed to lower readability (-2 SD) results in a 0.71% increase in CAR, or $70 M increase in market value. This relationship is also supported in our interaction plot (Figure 2). The estimates indicate decreased benefits associated with using clarity when the firm has higher investments. Therefore, investments as a moderator accentuates the effect of readability on investor reaction.

Hypothesis 2C predicted that the effect of readability on investor reaction will be negatively moderated by institutional ownership. We posited that the market reaction associated with the level of readability will be contingent on the firm's ownership level. Shown by Model 3 in Table 3, readability and institutional ownership had a negative interaction effect (b = -0.053, p = 0.007). As we expected, investors reacted more positively to higher readability in firms with high (low) retail (institutional) ownership, but investors react more positively for lower readability when firms have retail ownership. For firms with lower institutional ownership (-1 SD), the relationship strengthens, and a highly readable 10-K documents results in a CAR increase of 4.36% or about a $429 M improvement in market value. Whereas for firms with higher institutional ownership (+1 SD), the relationship weakens, and a highly readable 10-K documents results in only a $298 M improvement in market value, or a 3.03% increase in CARs. This relationship is also supported in our interaction plot (Figure 3). Therefore, institutional ownership as a moderator accentuates the effect of readability on investor reaction.

Hypothesis 2D predicted that the effect of readability on investor reaction will be positively moderated by the firms' change in operating profitability. We posited that the market reaction associated

with the level of readability will be contingent on firms' true economic profit (i.e., OP). Model 4 shows that the interaction effect between readability and a decrease in operating profitability (captured through the OP dummy) is positive (b = 0.064, p = 0.001). As we expected, investors reacted more positively to higher readability when firms have a lower profitability or a negative change in OP. As such, for firms with low OP (- 1 SD), this finding translates to a CAR increase of 12.8% or $942 M increase in market value when the firms' readability is high (+2 SD) as opposed to lower readability (-2 SD). In contrast, for firms with a positive change in OP (+1 SD), the relationship weakens, and a higher readability (+2 SD) as opposed to lower readability (-2 SD) results in a 9% or $884 M increase in CAR. The estimates indicate increased benefits associated with using clear communication when the firm has higher profitability. This relationship is also supported in our interaction plot (Figure 4). Therefore, change in profitability from the fiscal year ending in year *t-1* to the fiscal year ending in *t* as a moderator accentuates the effect of readability on investor reaction.

6. Supplementary analyses – Robustness and endogeneity

So far, we argued that XBRL characters appropriately capture the document readability and complexity of narrative disclosure after the mandate of XBRL filing by SEC in 2012. In this section, we explore whether the relationship between readability and investor reaction yields a positive relationship in the pre-XBRL period by using log of gross file size, a readability measure promoted by Loughran and McDonald (2014), that has been widely used as a proxy for readability (Li, Zhao, 2015; Ertugrul, Lei, Qiu & Wan, 2017; Bao, Fung, Su, 2018 among others) as an alternative measure of readability. Because Loughran & McDonald (2014) in their 1994-2011 sample show that gross (and net) file size is able to better capture document readability. Particularly, the authors argue that the other available readability measures are generally inappropriate in measuring the readability of financial documents which include the Gunning Fog index (Gunning, 1952; Brown, Crowley & Elliott, 2020). Recently, Gao and Huang (2020) highlight that modern reporting approaches are transforming the way business documents are reported and such technologies have an impact on the traditional approaches. Furthermore, Cahan et al. (2022) show that while file size is able to significantly capture document readability but only in the pre-XBRL period (1994-2011) because the size and composition of file size has changed after the mandatory adoption of XBRL in 2012. Thus, we determine whether the predicted relationship between readability and investor reaction is robust by using the alternative readability measure (i.e., log of gross file size).

Consequently, we replace the XBRL measure with log of gross file size (READ) and re-estimate our control model. To do so, we first obtain the gross file size from Prof. Tim Loughran's website for each of the 10-K documents in our sample for the sample period 1994-2011. We also note that the average gross file size in the sample is about 1.17 megabytes. We reported the results in Model 1 of Table 4, where we

find the coefficient for log (gross file size) is positive (b = 0.437, p = 0.000), suggesting a positive effect on the investor reaction. In other words, the positive coefficient indicates that a lower file size reduces obfuscation and results in a positive investor reaction, that is consistent with the finding reported in Table 3. Therefore, with this analysis, we found results essentially are consistent with those presented in our primary analysis.

Additionally, to ensure our results are not solely driven by a computerized measure – XBRL characters, we also include a traditional measure of reporting by using HTML characters (Hypertext Markup Language). Because SEC requires the companies to file the 10-K documents in HTML format (adopted in 1999), investors can use the HTML reports. We posit that since HTML documents are mandatory and requires disclosure of all the textual information of the 10-K documents, it proxies the readability of the document. Thus, as an alternative measure of readability for XBRL we use log of HTML characters for each 10-K document and re-estimate the control model in Table 3 for the XBRL period (2012-2022). In Model 2 of Table 4, we find that the coefficient for log (HTML characters) is positive (b = 5.707, p = 0.000), suggesting a positive effect on the investor reaction.

Since, previous studies (Cohen, Malloy & Nguyen, 2020; You & Zhang, 2008) argue that investor reaction to changes in 10-K tonality and earnings announcements are appropriately captured between 6-12 months. Therefore, to ensure that the results are not driven by the time horizon used in the calculation of CAR, in our second test we re-estimated our control model by replacing 3-Month CAR with 6- and 12-month CARs to ensure that the effect of readability on investor reaction is persistent up to 12-months. We reported the results in Models 3 and 4 of Table 4, Panel A. From Model 3 we find the coefficient for readability is positive (b = 9.938, p = 0.000) and in Model 4 also we find the coefficient for readability is positive (b = 13.514, p = 0.000). From these models it is evident that the effect of readability on investor reaction remains strong with 6-month and 12-month CAR. In unreported tests we also use Tobin's Q, a commonly used measure of long-term profitability that reflects the market's perception of current and potential profitability as an added performance measure and find that our results are comparable to those reported in Table 3 (Blose, & Shieh, 1997; Chen, Ho, Kee & Yeo, 2000; Carpenter, 2002; Erickson & Whited, 2006; Huselid, 1995; Richard, Murthi, & Ismail, 2007).

In our third robustness test, we control for potential endogeneity associated with document readability, given the possibility that the investor reaction itself may be an antecedent of the readability measures. For example, managers could use complex language to hide underperformance, in contrast profitable firms could be very forthright in their disclosures. Previous studies in accounting literature (Li, 2008; Lo, Ramos & Rogo, 2017) have documented a negative relationship between readability and level of earnings, suggesting that poor performing firms are more likely to obfuscate through complex language. However, it is not clear that the complex language is an artefact of poor performance or whether bad news

is difficult to communicate (Bloomfield, 2008). Thus, to test the potential endogeneity in our specification, we conducted Durbin-Wu-Hausman (DWH) chi-square test (test of endogeneity) as an initial check. We find that the p-value of DWH test is 0.000 (chi-squared = 15.43) rejecting the null hypothesis that the specified endogenous regressors can actually be treated as exogenous, indicating a potential endogeneity between readability and investor reaction. Additionally, previous studies such as Withers, Certo, & Semadeni (2014) and Benischke, Martin & Glaser (2018) argue that models that include interaction terms should control for endogeneity of the main effect of endogenous variable that is interacted (i.e., READ). Therefore, based on the previous studies and the DWH test we performed 2SLS regressions for our control model using Bog index and total word count as two distinct instruments to control for potential endogeneity. The motivation to choose Bog index and total word count as proxy for readability, is supported by recent studies by Bonsall, Leone, Miller & Rennekamp (2017) and Loughran & McDonald (2020). Specifically, Bonsall et al. (2017) highlight that Bog index directly captures the complex word count, sentence length, use of passive voice, weak verbs, and even jargon into its score. Therefore, we posit that Bog index is a reasonable proxy for document readability. Similarly, Loughran & McDonald (2020) highlight that total word count does not include figures, tables and other characters unlike gross file size. However, it is likely that larger 10-K documents may contain relatively more words. Thus, we use log of total word count as another proxy for readability. Models 3 and 4 of Table 4, Panel A reports the second-stage results for 2SLS. Our results remain robust to the endogeneity control. The instruments are found to be exogenous and valid as the Sargan's J-test shows that the equations are exactly identified, suggesting that the instruments used for readability are valid instruments. In addition, we also note that the Cragg-Donald Wald statistic (p-value = 0.000) rejects the null hypothesis that the instrument is weak, indicating that our specification is not plagued by weak instrument problem. Moreover, the F-statistics in our first stage models were also significant (F-statistics = 11.11, p = 0.00) confirming that instrument is not weak. The conventional rule of thumb is that an F-statistic of more than 10 indicates the instrument is not weak (see Stock & Yogo, 2002).

In a fourth robustness check, we follow previous studies (Breaugh, 2006; Glaser, Stam, & Takeuchi, 2016; Becker et al., 2016; Benischke, Martin, Glaser, 2018) that highlight that estimating the regressions with and without any control variable is important because doing so allows one to better understand the impact of the control variables on the relationship between explanatory variables (i.e., moderators) and the dependent variable. These studies also emphasize that running analysis without control variables offers important information about the utility of our explanatory variables to explain uncontrolled variance in the dependent variable (i.e., CAR). Therefore, if it turns out that our explanatory variable is highly dependent on other control variables, or a subset, the statistical significance of READ should disappear when we re-estimate our specifications without the controls. Accordingly, we begin in our control model with only READ. In the next model, we have READ and (READ x SIZE). In the next model it is

READ and READ x INVESTMENTS. Next, it is READ and READ x IO, and so on. We find that the coefficients of READ, and the interaction terms, added one at a time, are similar to those in Table 3.

7. Discussion and Conclusion

The IM literature has focused on the extent that firms might direct impressions in written communication with stakeholders. Such directing could include specific words and phrases, and imagery. However, the SEC has long encouraged a plain English style, through specific linguistic and document formatting suggestions, in any effort to attenuate any strategic attempt by corporations to misdirect shareholders. Document readability is considered vitally important in that it increases transparency and promotes greater information retention in addition to generating investor trust which then results in positive investor reaction. A fluent, clearly articulated text can lead to favorable evaluations of the company by the various stakeholders including regulators and policymakers. We add to this body of work by examining persuasive language attributes associated with a publicly traded corporation's annual 10-K filing documents and their impact on investor reaction. We proposed and found support for the notion that document readability overall matters to investors and investors react positively to 10-Ks that are more readable. We show that this relationship is moderated by a number of company specific observables like firm size, firm investments, institutional ownership and operating profitability. We also use a measure of readability that is more common in finance/accounting literature and port it over to the strategy literature. To ensure that our choice of a simple readability measure did not influence our findings we used another measure of document tonality that has been widely used in the literature and is computed from the fraction of negative words in a document (Loughran & McDonald, 2011) that has been argued to be particularly relevant to analyzing financial documents, in particular, that has been widely employed in the literature (Tetlock, 2007; Loughran & McDonald, 2011, 2014; Huang, Zang & Zheng, 2014; Demers & Vega, 2010). Also worth noting is that the construction of this document tonality measure is entirely different than our original readability measure. Our results do not change in making this substitution and speaks to the robustness of our results.

We add to the impression management literature in very significant ways by demonstrating the importance of the readability on arguably the most significant corporate communication with its stakeholders – its annual 10-K filings. While other studies have identified the importance of subtle language cues embedded in the linguistic style of managers and persuasive language attributes and language concreteness, we instead use a simple directly observed, and validated, instrument of document readability which abstains from potential complications of misidentification and misclassification of concrete language attributes associated with any text-based analysis.

Our study could be extended into other measures of document readability into capturing other forms of corporate communication like the firm's 8-K filings which is a way for all publicly traded firms to announce significant corporate news. It would also be interesting to compare our findings with those reported by the study of Pan et al. (2018) of concrete language attributes and its impact on investors using the same transcribed earnings calls. Since our measure is free from the complexities associated with measuring subtle cues and other aspects of language concreteness, it would be nice to see if the results are close. This would provide strategy and management researchers and alternative path to capturing corporate communication through subtle and persuasive linguistic attributes.

Our study raises intriguing questions regarding the possibility of using concrete-language-free instrument to proxy for concrete language attributes. While this would certainly simplify researchers' tasks on one dimension it would raise other questions regarding what these language attributes might really be measuring.

References

Adelberg, A. H. (1979). Narrative disclosures contained in financial reports: means of communication or manipulation?. *Accounting and business research*, *9*(35), 179-190.

Andries, A. M., Balutel, D., Ihnatov, I., & Ursu, S. G. (2020). The nexus between corporate governance, risk taking, and growth. *PloS one*, *15*(2), e0228371.

Assilaméhou, Y., Lepastourel, N., & Testé, B. (2013). How the linguistic intergroup bias affects group perception: Effects of language abstraction on generalization to the group. *The Journal of Social Psychology*, *153*(1), 98-108.

Audia, P. G., & Greve, H. R. (2006). Less likely to fail: Low performance, firm size, and factory expansion in the shipbuilding industry. *Management science*, *52*(1), 83-94.

Badrinath, S. G., Kale, J. R., & Noe, T. H. (1995). Of shepherds, sheep, and the cross-autocorrelations in equity returns. *The Review of Financial Studies*, *8*(2), 401-430.

Ballinger, G. A. (2004). Using generalized estimating equations for longitudinal data analysis. *Organizational research methods*, *7*(2), 127-150.

Bamber, L. S. (1987). Unexpected earnings, firm size, and trading volume around quarterly earnings announcements. *Accounting review*, 510-532.

Bamber, L. S., Barron, O. E., & Stober, T. L. (1997). Trading volume and different aspects of disagreement coincident with earnings announcements. The Accounting Review, 72(4), 575-597.

Banz, R. W. (1981). The relationship between return and market value of common stocks. *Journal of Financial Economics*, *9*(1), 3-18.

Bao, D., Fung, S. Y. K., & Su, L. (2018). Can shareholders be at rest after adopting clawback provisions? Evidence from stock price crash risk. *Contemporary Accounting Research*, *35*(3), 1578-1615.

Bartov, E., Givoly, D., & Hayn, C. (2002). The rewards to meeting or beating earnings expectations. *Journal of accounting and economics*, *33*(2), 173-204.

Barberis, N., & Thaler, R. (2003). A survey of behavioral finance. *Handbook of the Economics of Finance*, *1*, 1053-1128.

Benischke, M. H., Martin, G. P., & Glaser, L. (2019). CEO equity risk bearing and strategic risk taking: The moderating effect of CEO personality. *Strategic Management Journal*, *40*(1), 153-177.

Bernard, V. L., & Thomas, J. K. (1990). Evidence that stock prices do not fully reflect the implications of current earnings for future earnings. *Journal of accounting and economics*, *13*(4), 305-340.

Bettis, R. A. (1983). Modern financial theory, corporate strategy and public policy: Three conundrums. *Academy of Management Review*, *8*(3), 406-415.

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.

Black, F., Jensen, M. C. & Scholes, M. S. (1972). The capital asset pricing model: Some empirical tests. Studies in the Theory of Capital Markets, Praeger Publishers Inc.

Bloomfield, R. (2008). Discussion of "annual report readability, current earnings, and earnings persistence". *Journal of Accounting and Economics*, *45*(2-3), 248-252.

Bodnaruk, A., Loughran, T., & McDonald, B. (2015). Using 10-K text to gauge financial constraints. *Journal of Financial and Quantitative Analysis*, *50*(4), 623-646.

Bolino, M. C., Kacmar, K. M., Turnley, W. H., & Gilstrap, J. B. (2008). A multi-level review of impression management motives and behaviors. *Journal of Management*, *34*(6), 1080-1109.

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

Boudoukh, J., Feldman, R., Kogan, S., & Richardson, M. (2019). Information, trading, and volatility: Evidence from firm-specific news. *Review of Financial Studies*, *32*(3), 992-1033.

Bradac, J. J., Bowers, J. W., & Courtright, J. A. (1979). Three language variables in communication research: Intensity, immediacy, and diversity. *Human Communication Research*, *5*(3), 257-269.

Breaugh, J. A. (2006). Rethinking the control of nuisance variables in theory testing. *Journal of Business and Psychology*, *20*, 429-443.

Brennan, N. M., & Merkl-Davies, D. M. (2013). Accounting narratives and impression management. In *The Routledge companion to accounting communication* (pp. 123-146). Routledge.
Brickley, J. A., Lease, R. C., & Smith Jr, C. W. (1988). Ownership structure and voting on antitakeover amendments. *Journal of financial economics*, *20*, 267-291.
Brown, N. C., Crowley, R. M., & Elliott, W. B. (2020). What are you saying? Using topic to detect financial misreporting. *Journal of Accounting Research*, *58*(1), 237-291.
Buehlmaier, M. M., & Whited, T. M. (2018). Are financial constraints priced? Evidence from textual analysis. *Review of Financial Studies*, *31*(7), 2693-2728.
Bushee, B. J., & Goodman, T. H. (2007). Which institutional investors trade based on private information about earnings and returns?. *Journal of Accounting Research*, *45*(2), 289-321.
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.
Cahan, S.F., Chang, S., Siqueira, W.Z., & Tam, K., (2022). The roles of XBRL and processed XBRL in 10-K readability, Journal of Business, Finance & Accounting, 49(1-2), 33-68.
Calomiris, C. W., & Mamaysky, H. (2019). How news and its context drive risk and returns around the world. *Journal of Financial Economics*, *133*(2), 299-336.
Campbell, J. L., Chen, H., Dhaliwal, D. S., Lu, H. M., & Steele, L. B. (2014). The information content of mandatory risk factor disclosures in corporate filings. *Review of Accounting Studies*, *19*(1), 396-455.
Campbell, J. L., Lee, H. S. G., Lu, H. M., & Steele, L. B. (2020). Express Yourself: Why Managers' Disclosure Tone Varies Across Time and What Investors Learn from It. *Contemporary Accounting Research*, *37*(2), 1140-1171.
Carpenter, M. A. (2002). The implications of strategy and social context for the relationship between top management team heterogeneity and firm performance. *Strategic Management Journal*, *23*(3), 275-284.
Cazier, R. A., & Pfeiffer, R. J. (2016). Why are 10-K filings so long?. *Accounting Horizons*, *30*(1), 1-21.
Ceccagnoli, M. (2009). Appropriability, preemption, and firm performance. *Strategic Management Journal*, *30*(1), 81-98.
Chakrabarty, B., Seetharaman, A., Swanson, Z., & Wang, X. (2018). Management risk incentives and the readability of corporate disclosures. *Financial Management*, *47*(3), 583-616.
Chakravarty, S. (2001). Stealth-trading: Which traders' trades move stock prices?. *Journal of Financial Economics*, *61*(2), 289-307.
Chen, Q. (2007). Discussion of which institutional investors trade based on private information about earnings and returns?. *Journal of Accounting Research*, *45*(2), 323-331.
Chen, J. V., Nagar, V., & Schoenfeld, J. (2018). Manager-analyst conversations in earnings conference calls. *Review of Accounting Studies*, *23*(4), 1315-1354.
Chen, S. J., Chang, T. Z., Hui-Kuang Yu, T., & Mayes, T. (2005). Firm size and book-to-market equity as risk proxy in investment decisions. *Management Research News*, *28*(4), 1-24.
Chen, S., DeFond, M. L., & Park, C. W. (2002). Voluntary disclosure of balance sheet information in quarterly earnings announcements. *Journal of Accounting and Economics*, *33*(2), 229-251.
Chris I. (2021, May 22). 'Tesla shorts are having a much, much better year than in 2020'. *CNN Business*. https://edition.cnn.com/2021/05/22/investing/tesla-shorts/index.html
Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). Applied multiple regression/correlation analysis for the behavioral sciences (3rd ed.). Mahwah, NJ: Erlbaum.
Cohen, L., Malloy, C., & Nguyen, Q. (2020). Lazy prices. *Journal of Finance*, *75*(3), 1371-1415.
Collins, D. W., Li, O. Z., & Xie, H. (2009). What drives the increased informativeness of earnings announcements over time?. *Review of Accounting Studies*, *14*, 1-30.
Conrad, J., Gultekin, M. N., & Kaul, G. (1991). Asymmetric predictability of conditional variances. *The Review of Financial Studies*, *4*(4), 597-622.
Courtis, J. K. (2002). Emerging sensitivity of accountants to the role of communication and perception engineering. *Accounting, Auditing and Accountability Journal*, *15*(4), 444-449.

Courtis, J. K. (2004). Corporate report obfuscation: artefact or phenomenon? *The British Accounting Review*, *36*(3), 291-312.
Cornett, M. M., Marcus, A. J., Saunders, A., & Tehranian, H. (2007). The impact of institutional ownership on corporate operating performance. *Journal of banking & finance*, *31*(6), 1771-1794.
Cyert, R. M., & March, J. G. (1963). A behavioral theory of the firm. *Englewood Cliffs, NJ*, *2*(4), 169-187.
D'Argembeau, A., & Van der Linden, M. (2004). Influence of affective meaning on memory for contextual information. *Emotion*, *4*(2), 173.
Das, S. R., & Chen, M. Y. (2007). Yahoo! for Amazon: Sentiment extraction from small talk on the web. *Management science*, *53*(9), 1375-1388.
D'Augusta, C., & DeAngelis, M. D. (2020). Does accounting conservatism discipline qualitative disclosure? Evidence from tone management in the MD&A. *Contemporary Accounting Research*, *37*(4), 2287-2318.
Davidson III, W. N., Jiraporn, P., Kim, Y. S., & Nemec, C. (2004). Earnings management following duality-creating successions: Ethnostatistics, impression management, and agency theory. *Academy of management journal*, *47*(2), 267-275.
Davis, A. K., Piger, J. M., & Sedor, L. M. (2012). Beyond the numbers: Measuring the information content of earnings press release language. *Contemporary Accounting Research*, *29*(3), 845-868.
Dechêne, A., Stahl, C., Hansen, J., & Wänke, M. (2010). The truth about the truth: A meta-analytic review of the truth effect. *Personality and Social Psychology Review*, *14*(2), 238-257.
Demers, E., & Vega, C. (2010). Soft information in earnings announcements: News or noise? INSEAD Working Paper No. 2010/44/FIN.
De Bondt, W. F., & Thaler, R. (1985). Does the stock market overreact?. *The Journal of finance*, *40*(3), 793-805.
Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., ... & Lautenbach, S. (2013). Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, *36*(1), 27-46.
Edelen, R. M., Ince, O. S., & Kadlec, G. B. (2016). Institutional investors and stock return anomalies. *Journal of Financial Economics*, *119*(3), 472-488.
Elliott, W. B., Koëter-Kant, J., & Warr, R. S. (2008). Market timing and the debt–equity choice. *Journal of Financial Intermediation*, *17*(2), 175-197.
Elliott, W. B., Rennekamp, K. M., & White, B. J. (2015). Does concrete language in disclosures increase willingness to invest?. *Review of Accounting Studies*, *20*, 839-865.
Elsbach, K. D. (2003). Organizational perception management. *Research in organizational behavior*, *25*, 297-332.
Ertugrul, M., Lei, J., Qiu, J., & Wan, C. (2017). Annual report readability, tone ambiguity, and the cost of borrowing. *Journal of Financial and Quantitative Analysis*, *52*(2), 811-836.
Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, *33*(1), 3-56.
Fama, E. F., & French, K. R. (1997). Industry costs of equity. *Journal of financial economics*, *43*(2), 153-193.
Fama, E. F., & French, K. R. (2015). A five-factor asset-pricing model. *Journal of Financial Economics*, *116*(1), 1-22.
Feldman, R., Govindaraj, S., Livnat, J., & Segal, B. (2010). Management's tone change, post earnings announcement drift and accruals. *Review of Accounting Studies*, *15*, 915-953.
Fisher, L. (1966). Some new stock-market indexes. *The Journal of Business*, *39*(1), 191-225.
Fiss, P. C., & Zajac, E. J. (2006). The symbolic management of strategic change: Sensegiving via framing and decoupling. *Academy of management journal*, *49*(6), 1173-1193.
Ferreira, M. A., & Matos, P. (2008). The colors of investors' money: The role of institutional investors around the world. *Journal of financial economics*, *88*(3), 499-533.
Frazier, K. B., Ingram, R. W., & Tennyson, B. M. (1984). A methodology for the analysis of narrative accounting disclosures. *Journal of Accounting Research*, 318-331.

Garcia, D. (2013). Sentiment during recessions. *Journal of Finance*, *68*(3), 1267-1300.
Goffman, E. (1959). The presentation of self in everyday life: Selections.
Gompers, P. A., & Metrick, A. (2001). Institutional investors and equity prices. The Quarterly Journal of Economics.
Graffin, S. D., Carpenter, M. A., & Boivie, S. (2011). What's all that (strategic) noise? Anticipatory impression management in CEO succession. *Strategic Management Journal*, *32*(7), 748-770.
Griffin, J. M., & Lemmon, M. L. (2002). Book-to-market equity, distress risk, and stock returns. *The Journal of Finance*, *57*(5), 2317-2336.
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.
Gunny, K., & Zhang, T. C. (2014). Do managers use meeting analyst forecasts to signal private information? Evidence from patent citations. *Journal of Business Finance & Accounting*, *41*(7-8), 950-973.
Gunning, R. (1952) The Technique of Clear Writing. *McGraw-Hill*, *New York*.
Guo, K. H., & Yu, X. (2022). Retail Investors Use XBRL Structured Data? Evidence from the SEC's Server Log. Journal of Behavioral Finance, 23(2), 166-174.
Hallen, B. L., & Pahnke, E. C. (2016). When do entrepreneurs accurately evaluate venture capital firms' track records? A bounded rationality perspective. *Academy of Management Journal*, *59*(5), 1535-1560.
Hansen, J., & Wänke, M. (2010). Truth from language and truth from fit: The impact of linguistic concreteness and level of construal on subjective truth. *Personality and Social Psychology Bulletin*, *36*(11), 1576-1588.
Harris, R., 2000. Rethinking Writing, The Athlone Press, London, p. 254.
Harvey, C. R., Liu, Y., & Zhu, H. (2016). … and the cross-section of expected returns. *Review of Financial Studies*, *29*(1), 5-68.
Healy, P. M., & Palepu, K. G. (2001). Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical disclosure literature. *Journal of accounting and economics*, *31*(1-3), 405-440.
Henderson, A. D., Miller, D., & Hambrick, D. C. (2006). How quickly do CEOs become obsolete? Industry dynamism, CEO tenure, and company performance. *Strategic Management Journal*, *27*(5), 447-460.
Hendricks, K. B., Singhal, V. R., & Zhang, R. (2009). The effect of operational slack, diversification, and vertical relatedness on the stock market reaction to supply chain disruptions. *Journal of operations management*, *27*(3), 233-246.
Hilbe, J. M., & Hardin, J. W. (2008). Generalized estimating equations for longitudinal panel analysis. *Handbook of longitudinal research: Design, measurement, and analysis*, *1*, 467-474.
Hillert, A., Jacobs, H., & Müller, S. (2014). Media makes momentum. *Review of Financial Studies*, *27*(12), 3467-3501.
Hirshleifer, D., & Teoh, S. H. (2003). Limited attention, information disclosure, and financial reporting. *Journal of accounting and economics*, *36*(1-3), 337-386.
Hoitash, R., & Hoitash, U. (2018). Measuring accounting reporting complexity with XBRL. *The Accounting Review*, *93*(1), 259-287.
Hoitash, R., & Hoitash, U. (2022). A Measure of Firm Complexity: Data and Code. *Journal of Information Systems*, *36*(2), 161-172.
Hoitash, R., Hoitash, U., & Morris, L. (2021). eXtensible Business Reporting Language (XBRL): a review and implications for future research. *Auditing: A Journal of Practice & Theory*, *40*(2), 107-132.
Horowitz, J. L., Loughran, T., & Savin, N. E. (2000). Three analyses of the firm size premium. *Journal of Empirical Finance*, *7*(2), 143-153.
Hosman, L. A. (2002). Language and persuasion. *The persuasion handbook: Developments in theory and practice*, 371-390.
Hou, K., Mo, H., Xue, C., & Zhang, L. (2021). An augmented q-factor model with expected growth. *Review of Finance*, *25*(1), 1-41.

Hou, K., Xue, C., & Zhang, L. (2015). Digesting anomalies: An investment approach. *Review of Financial Studies*, *28*(3), 650-705.
Huang, A. H., Zang, A. Y., & Zheng, R. (2014). Evidence on the information content of text in analyst reports. The Accounting Review, 89(6), 2151-2180.
Huang, X., Teoh, S. H., & Zhang, Y. (2014). Tone management. *The Accounting Review*, *89*(3), 1083-1113.
Huselid, M. A. (1995). The impact of human resource management practices on turnover, productivity, and corporate financial performance. *Academy of management journal*, *38*(3), 635-672.
Hwang, B. H., & Kim, H. H. (2017). It pays to write well. *Journal of Financial Economics*, *124*(2), 373-394.
Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance*, *48*(1), 65-91.
Jegadeesh, N., & Wu, D. (2013). Word power: A new approach for content analysis. *Journal of financial economics*, *110*(3), 712-729.
Jennings, J., Kim, J. M., Lee, J., & Taylor, D. (2023). Measurement error, fixed effects, and false positives in accounting research. *Review of Accounting Studies*, 1-37.
Jerez-Fernandez, A., Angulo, A. N., & Oppenheimer, D. M. (2014). Show me the numbers: Precision as a cue to others' confidence. *Psychological science*, *25*(2), 633-635.
Jiang, F., Lee, J., Martin, X., & Zhou, G. (2019). Manager sentiment and stock returns. *Journal of Financial Economics*, *132*(1), 126-149.
Jones, C. M. (2002). A century of stock market liquidity and trading costs. *Available at SSRN 313681*.
Keim, D. B. (1990). A new look at the effects of firm size and E/P ratio on stock returns. *Financial Analysts Journal*, *46*(2), 56-67.
Kothari, S. P., & Shanken, J. (1997). Book-to-market, dividend yield, and expected market returns: A time-series analysis. *Journal of Financial economics*, *44*(2), 169-203.
Kravet, T., & Muslu, V. (2013). Textual risk disclosures and investors' risk perceptions. *Review of Accounting Studies*, *18*(4), 1088-1122.
Krause, R., Filatotchev, I., & Bruton, G. D. (2016). When in Rome, look like Caesar? Investigating the link between demand-side cultural power distance and CEO power. *Academy of Management Journal*, *59*(4), 1361-1384.
Lang, M., & Lundholm, R. (1993). Cross-sectional determinants of analyst ratings of corporate disclosures. Journal of Accounting Research, 31(2), 246-271.
Langacker, R. W. (1987). *Foundations of cognitive grammar: Volume I: Theoretical prerequisites* (Vol. 1). Stanford university press.
Larrimore, L., Jiang, L., Larrimore, J., Markowitz, D., & Gorski, S. (2011). Peer to peer lending: The relationship between language features, trustworthiness, and persuasion success. *Journal of Applied Communication Research*, *39*(1), 19-37.
Leary, M. R., & Kowalski, R. M. (1990). Impression management: A literature review and two-component model. *Psychological bulletin*, *107*(1), 34.
Lee, S. H., & Makhija, M. (2009). Flexibility in internationalization: is it valuable during an economic crisis?. *Strategic Management Journal*, *30*(5), 537-555.
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.
Lev, B., & Sougiannis, T. (1996). The capitalization, amortization, and value-relevance of R&D. Journal of Accounting and Economics, 21(1), 107-138.
Li, F. (2008). Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and Economics*, *45*(2-3), 221-247.
Li, J., & Zhao, X. (2015). Complexity and information content of financial disclosures: Evidence from evolution of uncertainty following 10-k filings. *Available at SSRN 2516622*.
Liang, K. Y., & Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika*, *73*(1), 13-22.

Liew, J., & Vassalou, M. (2000). Can book-to-market, size and momentum be risk factors that predict economic growth?. *Journal of Financial Economics*, *57*(2), 221-245.
Lo, K., Ramos, F., & Rogo, R. (2017). Earnings management and annual report readability. *Journal of accounting and Economics*, *63*(1), 1-25.
Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *Journal of Finance*, *66*(1), 35-65.
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.
Loughran, T., & McDonald, B. (2017). The use of EDGAR filings by investors. *Journal of Behavioral Finance*, *18*(2), 231-248.
Loughran, T., & McDonald, B. (2020). Textual analysis in finance. *Annual Review of Financial Economics*, *12*, 357-375.
Lovallo, D., Koller, T., Uhlaner, R., & Kahneman, D. (2020). Your company is too risk-averse. *Harvard Business Review*, *98*(2), 104-111.
Matsumoto, D. A. (2002). Management's incentives to avoid negative earnings surprises. *The accounting review*, *77*(3), 483-514.
Mc Laughlin, G. H. (1969). SMOG grading-a new readability formula. *Journal of reading*, *12*(8), 639-646.
McCullagh, P., & Nelder, J. A. (1989). Binary data. In *Generalized linear models* (pp. 98-148). Springer US.
McCullagh, P. (1984). Tensor notation and cumulants of polynomials. *Biometrika*, *71*(3), 461-476.
McCroskey, J. C., Wrench, J. S., & Richmond, V. P. (2003). Principles of public speaking. *Indianapolis, IN: The College Network*.
Merkl-Davies, D. M., & Brennan, N. M. (2007). Discretionary disclosure strategies in corporate narratives: incremental information or impression management?. *Journal of accounting literature*, *27*, 116-196.
Miller, B. P. (2010). The effects of reporting complexity on small and large investor trading. *The Accounting Review*, *85*(6), 2107-2143.
Novy-Marx, R. (2013). The other side of value: The gross profitability premium. *Journal of financial economics*, *108*(1), 1-28.
Olson, D. R. (2002). Roy Harris, Rethinking writing. London: Athlone Press, 2000. Pp. xvi, 254. Hb $39.95. *Language in Society*, *31*(3), 449-452.
Oppenheimer, D. M. (2006). Consequences of erudite vernacular utilized irrespective of necessity: Problems with using long words needlessly. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, *20*(2), 139-156.
Pan, L., McNamara, G., Lee, J. J., Haleblian, J., & Devers, C. E. (2018). Give it to us straight (most of the time): Top managers' use of concrete language and its effect on investor reactions. *Strategic Management Journal*, *39*(8), 2204-2225.
Puffer, S. M., & Weintrop, J. B. (1991). Corporate performance and CEO turnover: The role of performance expectations. *Administrative Science Quarterly*, 1-19.
Quigley, T. J., & Hambrick, D. C. (2012). When the former CEO stays on as board chair: Effects on successor discretion, strategic change, and performance. *Strategic Management Journal*, *33*(7), 834-859.
Richard, O. C., Murthi, B. S., & Ismail, K. (2007). The impact of racial diversity on intermediate and long-term performance: The moderating role of environmental context. *Strategic Management Journal*, *28*(12), 1213-1233.
Robert, G. (1952). The technique of clear writing. *New York*.
Rogers, J. L., Van Buskirk, A., & Zechman, S. L. (2011). Disclosure tone and shareholder litigation. *The Accounting Review*, *86*(6), 2155-2183.
Rutherford, B. A. (2003). Obfuscation, textual complexity and the role of regulated narrative accounting disclosure in corporate governance. *Journal of management and governance*, *7*, 187-210.

Salancik, G. R., & Meindl, J. R. (1984). Corporate attributions as strategic illusions of management control. *Administrative science quarterly*, 238-254.
Schendel, D. E., & Hofer, C. W. (1979). A new view of business policy and planning. *Strategic Management, Boston: Little, Brown Boston*.
Scholes, M., & Williams, J. (1977). Estimating betas from nonsynchronous data. *Journal of financial economics*, *5*(3), 309-327.
Schrand, C. M., & Walther, B. R. (2000). Strategic benchmarks in earnings announcements: the selective disclosure of prior-period earnings components. *The Accounting Review*, *75*(2), 151-177.
SEC. (2017). EDGAR public dissemination service (PDS) technical specification. SEC Public Dissemination Service.
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.
Semadeni, M., Withers, M. C., & Trevis Certo, S. (2014). The perils of endogeneity and instrumental variables in strategy research: Understanding through simulations. *Strategic Management Journal*, *35*(7), 1070-1079.
Shah, A. K., & Oppenheimer, D. M. (2007). Easy does it: The role of fluency in cue weighting. *Judgment and Decision making*, *2*(6), 371-379.
Shefrin, H. (2002). *Beyond greed and fear: Understanding behavioral finance and the psychology of investing*. Oxford University Press on Demand.
Sias, R. W. (2004). Institutional herding. Review of Financial Studies, 17(1), 165-206.
Sias, R. W., Starks, L. T., & Titman, S. (2006). Changes in institutional ownership and stock returns: Assessment and methodology. *The Journal of Business*, *79*(6), 2869-2910.
Sloan, R. G. (1996). Do stock prices fully reflect information in accruals and cash flows about future earnings?. *Accounting review*, 289-315.
Snefjella, B., & Kuperman, V. (2015). Concreteness and psychological distance in natural language use. *Psychological science*, *26*(9), 1449-1460.
Stanton, P., Stanton, J., & Pires, G. (2004). Impressions of an annual report: an experimental study. *Corporate Communications: An International Journal*, *9*(1), 57-69.
Staw, B. M., McKechnie, P. I., & Puffer, S. M. (1983). The justification of organizational performance. *Administrative science quarterly*, 582-600.
Stillman, S. (2003). Review of generalized estimating equations by Hardin and Hilbe. *The Stata Journal*, *3*(2), 208-210.
Stock, J. H., Wright, J. H., & Yogo, M. (2002). A survey of weak instruments and weak identification in generalized method of moments. *Journal of Business & Economic Statistics*, *20*(4), 518-529.
Sydserff, R., & Weetman, P. (1999). A texture index for evaluating accounting narratives: An alternative to readability formulas. *Accounting, Auditing & Accountability Journal*.
Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *The Journal of finance*, *62*(3), 1139-1168.
Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. The Journal of Finance, 62(3), 1139-1168.
Tetlock, P. C., Saar-Tsechansky, M., & Macskassy, S. (2008). More than words: Quantifying language to measure firms' fundamentals. *Journal of Finance*, *63*(3), 1437-1467.
Toma, C. L., & D'Angelo, J. D. (2015). Tell-tale words: Linguistic cues used to infer the expertise of online medical advice. *Journal of Language and Social Psychology*, *34*(1), 25-45.
Toubiana, D., & Maruenda, H. (2021). Guidelines for correlation coefficient threshold settings in metabolite correlation networks exemplified on a potato association panel. *BMC bioinformatics*, *22*, 1-14.
Washburn, M., & Bromiley, P. (2014). Managers and analysts: An examination of mutual influence. *Academy of Management Journal*, *57*(3), 849-868.
Westphal, J. D., & Graebner, M. E. (2010). A matter of appearances: How corporate leaders manage the impressions of financial analysts about the conduct of their boards. *Academy of Management Journal*, *53*(1), 15-44.

Westphal, J. D., & Zajac, E. J. (1998). The symbolic management of stockholders: Corporate governance reforms and shareholder reactions. *Administrative science quarterly*, 127-153.
Womack, K. L. (1996). Do brokerage analysts' recommendations have investment value?. *The journal of finance*, *51*(1), 137-167.
Xu, N., Jiang, X., Chan, K. C., & Yi, Z. (2013). Analyst coverage, optimism, and stock price crash risk: Evidence from China. *Pacific-Basin Finance Journal*, *25*, 217-239.
You, H., & Zhang, X. J. (2009). Financial reporting complexity and investor underreaction to 10-K information. *Review of Accounting studies*, *14*, 559-586.
Zajac, E. J., & Westphal, J. D. (1995). Accounting for the explanations of CEO compensation: Substance and symbolism. *Administrative science quarterly*, 283-308.

Table 1 – Descriptive Statistics

Panel A: Firm characteristics (XBRL sample – 2012 – 2022)

Variables	Mean	SD [a]	Median
CAR [b]	5.3	47.0	1.0
XBRL characters [c]	3.4	1.6	3.2
READ (Log of XBRL characters)	14.9	0.5	15.0
SIZE [c]	9,831.7	38,618.7	1,526.0
ASSETS [c]	6,374.9	22,076.8	1,174.1
INV	-0.1	0.3	-0.1
IO [b]	70	30	80
OP	0.18	1.79	0.241
ROA	-0.2	34.0	0.1
TURNOVER	27.2	44.5	18.8
VOLATILITY	0.1	0.2	0.1
LEVERAGE	0.3	0.4	0.2
AGE [d]	25.9	18.9	21.0
ACCRUALS	-21.7	522.2	0.0

Panel B: Firm characteristics (pre-XBRL sample – 1994-2011)

Variables	Mean	SD [a]	Median
CAR [b]	2.3	13.9	1.5
Gross File Size [e]	11.5	8.7	8.3
READ (Log of Gross File Size)	2.2	0.6	2.1
SIZE [c]	15,022.6	37,083.9	3,898.2
ASSETS [c]	11,106.2	26,576.9	3,177.9
ROA	0.2	0.1	0.2
TURNOVER	30.7	24.2	25.0
VOLATILITY	0.1	0.2	0.1
LEVERAGE	0.2	0.2	0.2
AGE [d]	29.6	21.2	21.0
ACCRUALS	-22.5	507.5	0.2

Notes.

a. Standard deviation
b. Percentage
c. Millions
d. Years
e. Megabytes (MB)

Panel B: Pairwise Correlations

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
(1) CAR	1.00												
(2) READ (Log of XBRL characters)	0.06***	1.00											
(3) log SIZE	-0.01	-0.17***	1.00										
(4) log ASSETS	0.00	-0.23***	0.77***	1.00									
(5) INV	-0.04***	-0.06***	-0.02**	0.01	1.00								
(6) IO	-0.01	-0.26***	-0.04***	-0.03***	-0.04***	1.00							
(7) OP	0.00	-0.02*	0.00	0.00	0.01	0.02*	1.00						
(8) ROA	-0.08***	-0.04***	0.00	0.00	0.01	0.02*	0.16***	1.00					
(9) TURNOVER	0.22***	-0.01	-0.03***	-0.02*	-0.03***	0.10***	0.00	0.00	1.00				
(10) VOLATILITY	0.07***	0.13***	-0.07***	-0.06***	-0.04***	-0.11***	-0.01	0.00	0.13***	1.00			
(11) LEVERAGE	-0.01	-0.11***	0.03***	0.03***	0.12***	0.04***	0.00	-0.02*	0.03***	0.00	1.00		
(12) Log AGE	-0.07***	-0.20***	0.22***	0.29***	0.09***	-0.11***	-0.01	0.01	-0.11***	-0.10***	-0.03***	1.00	
(13) ACCRUALS	-0.01*	0.03***	-0.26***	-0.19***	0.02*	0.01	0.00	0.00	-0.01	0.01	0.00	-0.03***	1.00

Note. ***, ** and * indicate p<0.01, ** p<0.05 and * p<0.1 respectively

Table 2 - Moderating effect of firm characteristics upon relationship between readability and investor reaction

Characteristics	Theoretical logic	Predicted moderation effect	Supported
Firm Size	Decreases the effect of readability on investor reaction	Attenuation	Yes
Investments	Decreases the effect of readability on investor reaction	Attenuation	Yes
Institutional ownership	Decreases the effect of readability on investor reaction	Attenuation	Yes
Operating profitability	Increases the effect of readability on investor reaction	Accentuation	Yes

Table 3 – GEE estimates of readability on investor reaction

	Control Model	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)
Dependent:	3-month Cumulative Abnormal Returns (CAR)					
Intercept	0.80***	2.00***	0.08	0.41**	1.24***	0.523*
	(0.002)	(0.004)	(0.683)	(0.045)	(0.000)	(0.057)
ASSETS	-0.285		-0.522	-0.171	-0.295	
	(0.471)		(0.157)	(0.375)	(0.453)	
ROA	-4.948**	-8.306**	-5.607**	-3.270**	-5.009**	-6.565***
	(0.032)	(0.023)	(0.016)	(0.031)	(0.040)	(0.006)
TURNOVER	9.627***	9.231***	6.679***	2.771***	9.496***	2.856***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
VOLATILITY	6.732***	6.759***	5.007***	2.985**	6.856***	4.186***
	(0.000)	(0.001)	(0.002)	(0.039)	(0.000)	(0.003)
LEVERAGE	-4.295***	-5.404***	-2.498*	-0.130	-4.438***	-1.128
	(0.006)	(0.009)	(0.053)	(0.901)	(0.005)	(0.346)
AGE [a]	-1.827***	-1.314***	-0.975**	-0.220	-1.882***	0.115
	(0.000)	(0.009)	(0.033)	(0.509)	(0.000)	(0.730)
ACCRUALS	-0.001	-0.001	-0.001	-0.001***	-0.001	-0.001***
	(0.107)	(0.300)	(0.146)	(0.001)	(0.103)	(0.009)
NEGTONE	-2.893**	-2.863**	-1.385	0.655	-2.926**	1.236*
	(0.029)	(0.033)	(0.242)	(0.343)	(0.026)	(0.068)
POSTONE	-2.480	-1.486	0.024	3.562**	-2.303	2.834
	(0.457)	(0.672)	(0.994)	(0.050)	(0.488)	(0.110)
SIZE		-20.582*				
		(0.056)				
INVESTMENTS			-1.505**			
			(0.044)			
IO				-0.858***		
				(0.004)		
OP dummy					0.990***	6.993
					(0.001)	(0.509)
READ	6.156***	14.096***	1.118	2.895**	8.893***	4.176**
	(0.001)	(0.003)	(0.451)	(0.045)	(0.000)	(0.027)
READ x SIZE		-1.337**				
		(0.05)				
READ x INV			-0.098**			
			(0.047)			
READ x IO				-0.053***		
				(0.007)		
READ x OP dummy [b]					0.064***	0.427
					(0.001)	(0.550)
Firm, Year and Industry FE	Included	Included	Included	Included	Included	Included
N	13,256	12,523	13,116	11,096	13,256	11,440
Wald Chi^2	296.91	265.99	268.12	300.02	298.12	105.45
(p-value)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Notes. Robust p-values are reported in parentheses. [a] we used natural logarithm of firm age; [b] the dummy variable is defined as 1 for negative change in OP and 0 otherwise.

Table 4: Supplementary Tests

Panel A: Robustness and endogeneity tests

	Control Model				2SLS	
	Model 1	Model 2	Model 3	Model 4		
Dependent:	3-month CAR	3-month CAR	6-month CAR	12-month CAR	12-month CAR:	12-month CAR:
Readability	Log (Gross file size)	Log (HTML Characters)	Log (Gross file size)		Instrument: Log of Bog Index	Instrument: Log of total word count
Intercept	3.536*	0.76***	1.32***	1.90***		
	(0.062)	(0.000)	(0.002)	(0.000)		
READ	0.437**	5.707***	9.938***	13.514***	4.730**	1.057**
	(0.037)	(0.000)	(0.001)	(0.000)	(0.011)	(0.031)
ASSETS [a]	-0.170	-0.091	-0.887	-1.046	-4.904***	-4.663***
	(0.104)	(0.802)	(0.192)	(0.162)	(0.000)	(0.000)
ROA	0.560	-4.795**	-2.157*	0.020***	3.077	4.294
	(0.625)	(0.030)	(0.074)	(0.000)	(0.248)	(0.134)
TURNOVER	0.143	9.757***	15.035***	16.891***	-5.282***	-0.539
	(0.486)	(0.000)	(0.000)	(0.000)	(0.000)	(0.290)
VOLATILITY	0.332	5.727***	8.813***	9.971***	7.331***	3.380
	(0.506)	(0.001)	(0.004)	(0.002)	(0.000)	(0.128)
LEVERAGE	1.256	-4.147***	-5.413**	-4.261*	2.529	-0.621
	(0.164)	(0.008)	(0.010)	(0.056)	(0.209)	(0.779)
AGE [b]	-0.619**	-2.031***	-3.762***	-4.719***	-11.557**	-2.921*
	(0.011)	(0.000)	(0.000)	(0.000)	(0.012)	(0.062)
ACCRUALS	-0.000	-0.001	-0.002*	-0.003	0.000	0.000
	(0.437)	(0.127)	(0.095)	(0.131)	(0.971)	(0.855)
NEGTONE	1.121***	-2.925**	-1.772	-2.565	-2.491	1.305
	(0.007)	(0.028)	(0.397)	(0.306)	(0.303)	(0.172)
POSTONE	-1.928*	-2.390	-5.444	-5.616	-10.303***	-3.964*
	(0.050)	(0.492)	(0.283)	(0.377)	(0.009)	(0.072)
Firm, Year and Industry FE	Included				Firm FE	Firm FE
N	20,496	13,375	13,365	13,456	34,315	34,315
Wald Chi^2 (p-value)	599.20 (0.000)	303.78 (0.000)	183.64 (0.000)	347.28 (0.000)		
First-stage F-stat p-value					0.000	0.000

Notes. Robust p-values are reported in parentheses. [a] we used natural logarithm of total assets; [b] we used natural logarithm of firm age.

Panel B: Repeating Table 3 without any control variables

	Control Model	Model (1)	Model (2)	Model (3)	Model (4)
Dependent:			3-month CAR		
Intercept	0.74*** (0.000)	0.80*** (0.000)	0.25* (0.057)	0.41** (0.048)	1.37*** (0.000)
READ	5.016*** (0.000)	5.176*** (0.001)	1.431 (0.104)	2.928** (0.048)	1.16*** (0.000)
SIZE		10.76 (0.262)			
READ x SIZE		-0.013 (0.414)			
INV			-1.75** (0.031)		
READ x INV			-0.11** (0.035)		
IO				-0.89*** (0.004)	
READ x IO				-0.05*** (0.007)	
OP dummy [b]					0.38** (0.013)
READ x OP dummy [b]					0.074*** (0.000)
Firm, Year and Industry FE	Included	Included	Included	Included	Included
N	13,256	12,523	13,116	11,096	13,256
Wald Chi^2 (p-value)	279.81 (0.000)	253.83 (0.000)	252.54 (0.000)	263.36 (0.000)	280.48 (0.000)

Notes. Robust p-values are reported in parentheses. [a] we used natural logarithm of firm age; [b] the dummy variable is defined as 1 for positive change in OP and 0 otherwise.

Figure 1: Readability x SIZE

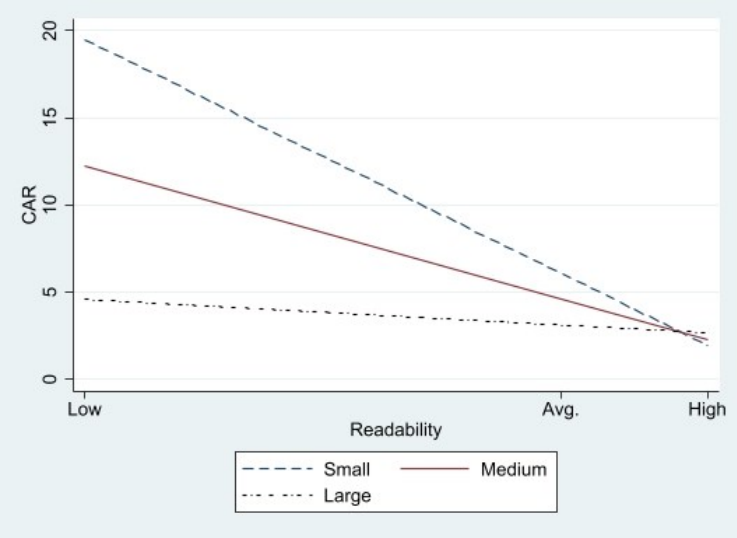

Figure 3: Readability x INV

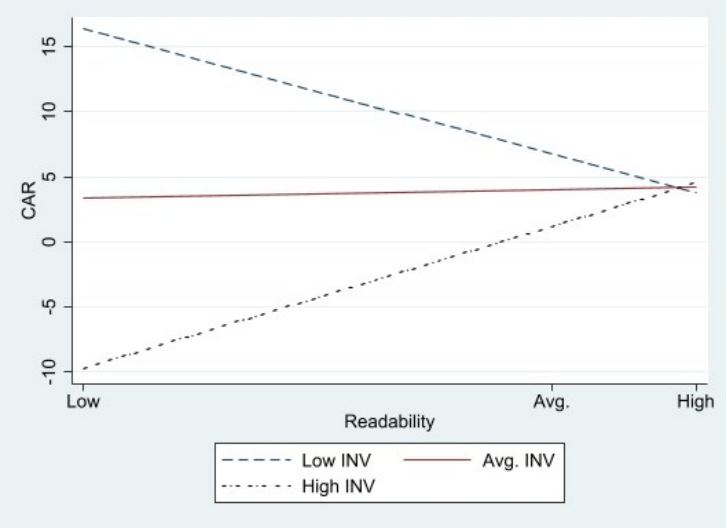

Figure 4: Readability x IO

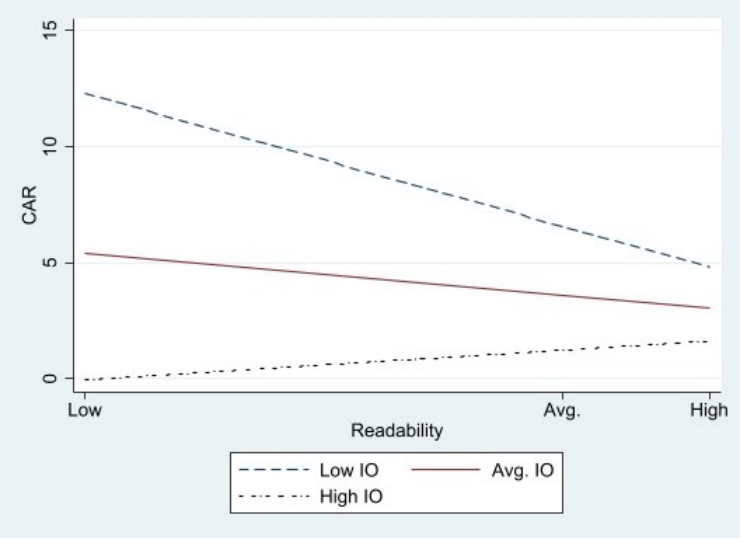

Figure 5: Readability x OP