Does financial disclosure readability predict shareholder activism?

Abhinav Anand[a], Xing Huan[b,*], Jalaj Pathak[c]

[a]*Indian Institute of Management Bangalore, India.*
[b]*EDHEC Business School, France.*
[c]*Indian Institute of Management Lucknow, India.*

Abstract

We examine the impact of financial disclosure readability on future shareholder activism, as expressed by shareholder proposals (SP). Based on a sample of 1,063 firms listed by the S&P 1500 between 2006 and 2019, we find that the likelihood of SP increases when 10-K readability declines. This effect is primarily driven by the use of complex words in (rather than the length of) 10-K files, and concentrated in corporate governance SP. We also show that M&A activities exacerbate the negative effect of 10-K readability on the likelihood of SP; restatements reduce the likelihood of SP for firms with complex disclosures by offering a communication tool for clarification; and firms improve disclosure readability subsequent to the release of negative news. Our findings are consistent with the 'incomplete revelation hypothesis' in suggesting that managers produce hard-to-read financial reports to delay the release of adverse information.

Keywords: 10-K, corporate governance, financial disclosure, readability, shareholder activism, shareholder proposals, textual analysis

JEL Classification: G14, G34, M41

*Corresponding author. Email: xing.huan@edhec.edu. Telephone: +33 493187857. Address: 393 Promenade des Anglais, Nice 06202, France.

We thank Michele Fabrizi, Cristina Grande-Herrera, April Klein, Anastasia Kopita, Selina Pei, Giulia Redigolo, and James Thewissen for helpful comments and suggestions. This paper also benefited from participants' comments at the European Accounting Association Annual Congress 2022.

1. Introduction

Shareholder activism through the proxy process entails a least costly and most common monitoring device for corporate governance (Iliev et al., 2015; Gillan and Starks, 2000). When firms' agency concerns are exacerbated, it is optimal that shareholders seek control over corporate decisions (Harris and Raviv, 2010). As an alternative mechanism to 'exit', shareholder activism is a form of 'voice' (Hirschman, 1970) when shareholders believe managerial actions are not in line with their interest. Prior literature has identified a list of antecedents and factors that explain shareholder activism, such as operating and stock market performance, cash holding, dividend policy, institutional ownership, spillover effect from peers, board monitoring, and ESG performance (Goranova and Ryan, 2014). At the outcome level, shareholder activism is shown to increase or improve firm disclosure (Flammer et al., 2021; Michelon et al., 2020; Baloria et al., 2019; Bourveau and Schoenfeld, 2017). However, it remains unknown whether financial disclosure readability influences shareholder activism. Given the growing popularity and influence of shareholder-initiated proxy proposals (Denes et al., 2017), it is important to study this research question.

This paper aims to extend our understanding of the role of financial disclosure readability in the context of shareholder proposals. Based on a sample of 1,063 firms listed by the S&P 1500 between 2006 and 2019, we examine whether the readability of 10-K filings predicts future shareholder proposals. We find that, after controlling for antecedents that prior literature has documented to explain shareholder activism and other text attributes such as tone, firms with less readable 10-K reports experience a higher likelihood of future shareholder-initiated proxy proposals. Further analyses show that this effect is primarily driven by the use of complex words in 10-K reports rather than the length of 10-K files, and concentrated in corporate governance proposals rather than social proposals. These results are robust to several additional tests including further controlling for ESG performance, using alternative readability measures, and estimation using a zero-inflated negative binomial model. We also show that M&A activities exacerbate the negative effect of 10-K readability on the likelihood of shareholder proposals; accounting restate-

ments reduce the likelihood of shareholder proposals for firms with complex disclosures by offering a communication tool for clarification; and firms improve disclosure readability subsequent to the release of negative news. Collectively, these findings are consistent with the 'incomplete revelation hypothesis' in suggesting that managers produce hard-to-read financial reports to delay the release of adverse information.

This paper makes several contributions. First, our study contributes to the literature by showing that financial disclosures' readability has a predictive power on future shareholder proposals—a previously unexplored implication of textual readability in financial disclosures. Our results are consistent with the 'incomplete revelation hypothesis' (IRH) (Bloomfield, 2008); and evidence unearthed in Li (2008), Kim et al. (2019), and Anand et al. (2022)—all of which suggest that managers produce hard-to-read financial reports to delay the release of adverse information. Second, our study contributes to the literature on the antecedents of shareholder activism (Goranova et al., 2017; Goranova and Ryan, 2014; Judge et al., 2010; Ryan and Schneider, 2002) by documenting financial disclosure readability as a plausible factor in predicting future shareholder proposals. Last but not the least, the interplay between disclosures' readability and shareholder activism has novel practical implications. Our findings thus speak to regulators who seek to improve the readability of firms' financial disclosures and practitioners who are in the position of stakeholder management to prevent shareholder activism.

The remainder of this paper proceeds as follows. Section 2 reviews background and literature, and develops hypotheses. Section 3 provides a description of our sample and empirical design. Section 4 presents the baseline results. Section 5 presents robustness test results. Section 6 concludes.

2. Background and Hypotheses Development

2.1. SEC's Plain English Regulation

With the objective of mitigating concerns over firms' unreadable financial disclosure filings, the Securities and Exchange Commission (SEC) adopted the 1998 *Plain English Mandate*, SEC Rule 421(d), complemented with a handbook entitled "*A Plain English*

Handbook: How to Create Clear SEC Disclosure Documents". The handbook encourages registrants to adopt plain English writing principles by avoiding writing constructs such as long sentences, passive voice, weak verbs, superfluous words, legal and financial jargon, numerous defined terms, abstract words, unnecessary details, and unreadable design and layout (Securities and Exchange Commission, 1998). An extensive stream of literature subsequently emerged, focusing on the impact of financial disclosures' readability on investors' behavior and welfare.[1] The SEC classified components of plain English in the following six categories: 'average sentence length', 'average word length', 'passive voice', 'legalese', 'personal pronouns', and 'negative/superfluous phrases'. While the rule officially applied only to prospectus filings, the SEC stated its clear preference for the usage of plain English in all communication with shareholders. While 10-K filings are not specifically covered by the mandate, the SEC documentation clearly favors conformance with the rule in all corporate filings.

The Plain Writing Act was signed into law on 13 October 2010. The Act requires that federal agencies use clear government communication that the public can understand and use. This legislation was not designed to make financial disclosures easier to read but highlights efforts to make the government more transparent to its citizenry. Several studies show that the Plain Writing Act improved the readability of 10-K files, leading to more effective risk management (Yin et al., 2022; Hwang and Kim, 2017).

2.2. Shareholder Proposals

Shareholder proposals have existed under the SEC Rule 14a-8 in the United States since 1942.[2] During the annual shareholder meeting, shareholders can express their concern with corporate performance and governance; pressure management for corporate reform by submitting proposals for a vote; sponsor campaigns that gain support from

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

[2] Under SEC Rule 14a-8, shareholders continuously holding shares worth USD 2000 (or 1 percent of the market value of equity) for at least one year can submit only one proposal with a 500-word supporting statement at least 120 days before the proxy statement is mailed to shareholders. A proposal may be excluded by the SEC, upon the request of the company, if it violates certain conditions or persuades the proponent to withdraw by agreeing to it. If a proposal is neither withdrawn nor excluded by the SEC, it will be included in the proxy and will be voted upon at the annual meeting.

fellow shareholders to withhold votes (in director election, for example), or make recommendations during the annual meeting. Early research questions the usefulness of shareholder proposals because of low voting support (Gillan and Starks, 2000; Gordon and Pound, 1993) and their nonbinding nature. During mid-1980s to early 1990s, shareholder proposals started to gain traction (Bauer et al., 2015; Thomas and Cotter, 2007) as a result of support from influential institutional shareholders (Denes et al., 2017); the shift to gaining majority votes, and increasing media scrutiny over the issues underlying proposals (García Osma and Grande-Herrera, 2021).

Thereafter, shareholder proposals have become one of the prevalent vehicles through which shareholders engage in activism by publicly voicing their intentions and/or dissatisfaction, in order to target the firm including its directors and management (García Osma and Grande-Herrera, 2021). Shareholder proposals entail a useful device of external control that helps counter managerial agency problems (Renneboog and Szilagyi, 2011; Bebchuk, 2005); and reduce agency costs through increasing director responsiveness to shareholder concerns (Thomas and Cotter, 2007). The market perceives proposals submitted against companies with exacerbated agency concerns as meaningful control benefits (Renneboog and Szilagyi, 2011). Shareholder proposals that win a majority votes are more likely to be implemented due to the *ex-ante* threat of reputational penalties (Ertimur et al., 2010). Shareholder proposals also increase the incidence of CEO turnover and independent board chairman appointment in target firms (Buchanan et al., 2012).

2.3. Financial Disclosure and Shareholder Activism

Disclosure helps firms reduce the chance of being targeted by means of pre-empting activism by guiding the stock price to converge to its fundamental value, enhancing communication between management and stakeholders, improving the credibility and reputation of the firm, and reducing litigation risk (Bourveau and Schoenfeld, 2017). In corporate disclosure research, 10-K filing is a frequently visited area, as it provides managers with an avenue to disclose critical inside information and managerial perspectives (Kim et al., 2019; Campbell et al., 2014; Merkley, 2014; Kravet and Muslu, 2013; Li et al., 2013). Corporate disclosure policy is jointly determined by costs and bene-

fits (Beyer et al., 2010). Grossman and Stiglitz (1980) demonstrates an efficient market equilibrium at which return to data analysis must equal the cost of analysis. Otherwise, more or fewer investors analyze the data until an equilibrium is reached. The 'incomplete revelation hypothesis' (IRH) (Bloomfield, 2002) proposes that information that is more costly to extract from public data are less completely revealed by market prices. A direct implication of IRH is that managers have incentives to strategically increase the processing cost of negative information by writing more complex financial reports to prevent stock prices from declining and associated outcomes such as discounts in their stock option-based compensation. This conjecture also referred to as 'management obfuscation hypothesis', is tested in Li (2008) and Kim et al. (2019). Li (2008) shows that managers make bad news more costly to process by writing excessively long annual reports with unnecessarily long sentences and big words. In turn, Kim et al. (2019) document that managers have both incentives and abilities to hide negative information by writing more opaque financial reports. On the other hand, managers tend to be more forthcoming in the disclosure when the firm performance is satisfactory (Schrand and Walther, 2000; Lang and Lundholm, 2000).

As discussed above, complex financial reports evoke a notion of management obfuscation of negative news in reporting entities. However, the engineered information opacity, facilitated by complex financial reports, only allows managers to delay the releases of adverse information to a certain threshold. The adverse information will be ultimately released when such threshold is surpassed (Kim et al., 2019). One could suspect that managers could simply omit negative news from the financial reports. However, given the *ex-ante* threat of reputational penalty and litigation risk, managers are less likely to be engaged in the practice of omitting key adverse information (Skinner, 1994). It is also possible that complex financial reports are produced as a result of a true depiction of firms with complex business activities and underlying economic conditions, and large and complex firms are more likely to become the target for shareholder activism (Cai and Walkling, 2011; Ertimur et al., 2011; Smith, 1996). To this end, we posit that firms with less readable financial reports are either complex entities, reporting complex man-

agerial disclosure; or those inclined to weak performance and/or corporate governance issues prior to shareholder voting; and this could lead to subsequent shareholder-initiated proposals. Therefore, we hypothesize that:

Hypothesis 1: *Financial disclosures with poor readability can lead to future shareholder-initiated proxy proposals.*

3. Data and Methodology

3.1. Data

Data used in this study are retrieved from several databases. First, the shareholder proposals' data are sourced from the Voting Analytics by Institutional Shareholder Services (ISS). Second, text and readability attribute data are sourced from SEC Analytics Suite by WRDS. Third, firm-level financial characteristics are extracted from Compustat. Fourth, board characteristics are measured using BoardEx data. Lastly, we employ institutional ownership data from Factset.

Table 1 reports the sample selection process. The sample selection starts with retrieving the annual list of S&P 1500 firms between 2006 and 2019 from Bloomberg. By using the S&P 1500 as our population,[3] we circumvent the biases associated with the use of non-random, equal share samples in the model estimation. Our sample starts from 2006 because this is the first year shareholder voting data is available. This initial sample consists of 20,473 firm-year observations for 2,474 firms. After matching with ISS Voting Analytics, SEC Analytics, Compustat, BoardEx, and Factset data, and removing financial institutions from the sample, the resulting sample spans from 2006 to 2019 and consists of 7,786 firm-year observations for 1,063 firms. Table 2 reports the breakdown of the number of proposals by firm. The number of proposals varies from 0 to 27. Table 3 reports sample composition by year and industry.

[Tables 1, 2 and 3 about here.]

[3]The S&P 1500 index covers approximately 90% of the market capitalization of U.S. stocks.

3.2. Model

In order to test the impact of financial disclosures' readability on future shareholder-initiated proxy proposals, we estimate Model 1 using the binomial logit regression.

$$SP_{i,t} = \alpha_0 + \alpha_1 Readability_{i,t-1} + \sum_{j=2}^{18} \alpha_j Control_{i,t-1} + FE + \epsilon_{i,t} \quad (1)$$

where i indexes firms, and t indexes years. *SP* refers to shareholder proposals, which is a binary variable that takes a value of one if the firm receives shareholder proposal(s) in the proxy season, and zero otherwise. *Readability* refers to readability measures. *Fog index* is the main readability measure used in the main tests. *Fog index* is calculated as 0.4 multiply by the sum of average words per sentence and the percentage of complex words in the 10-K. Higher value of *Fog index* implies less readable texts. α_1 is the coefficient of interest. In addition to the *Fog index*, we also control for several text attributes including *tone*, *uncertain*, and *litigious* to account for the potential impact of financial disclosures' tone and sentiment on the likelihood of shareholder proposals.

The choice of firm-level covariates is motivated by prior literature on the antecedents of shareholder activism. Specifically, the natural logarithm of total assets is used as a measure of firm size. Large firms are more likely to become the target (Cai and Walkling, 2011; Ertimur et al., 2011; Smith, 1996), as these firms have greater visibility (Rehbein et al., 2004) and shareholder activists could generate more value by targeting large companies (Del Guercio and Hawkins, 1999; Strickland et al., 1996). We control for firm complexity using two measures introduced in Markarian and Parbonetti (2007): (1) External complexity, which is measured as the ratio of the firm's sales to the total sales within the industry; and (2) Internal complexity, which is calculated as the ratio of R&D expenditures to the number of employees. We control for firm cash holding as activists target cash-rich companies to extract excess cash from them (Klein and Zur, 2009), and dividend payout ratio as lower dividend payout increases the likelihood of being targeted (Brav et al., 2008). We control for financial leverage, measured as the percentage of total liabilities of total assets. Hedge fund activists also tend to target firms with a lower leverage (Klein and Zur, 2009) while the reverse holds for governance-related activism

(Ferri and Sandino, 2009; Karpoff et al., 1996). We include return on assets and market-to-book ratio to control for firms' operating performance and stock market performance, respectively. Firms with poor operating performance and suboptimal stock market performance are more likely to become the target of shareholder activism (Ertimur et al., 2011; Renneboog and Szilagyi, 2011; Karpoff et al., 1996). We also control for tangibility as a proxy for liquidation costs (Smith, 2008), and sales growth to account for the effect of business expansion.

We further control for monitoring variables, both internal and external to the firm. For internal monitoring, we include three board characteristics' variables: (1) Board size, measured as the number of directors on the board; (2) Board independence, calculated as the ratio of non-executive directors on the board; and (3) Board diversity, computed as the first principal component of the gender ratio and nationality mix of the board. Institutional ownership is included as a measure for external monitoring. Prior literature documents a positive association between institutional ownership and shareholder activism (Cai and Walkling, 2011; Renneboog and Szilagyi, 2011; Smith, 1996). Cziraki et al. (2010) find that proposal probability increases in the target company's ownership concentration, and the equity stake of institutional investors. We include industry dummies, created using the first two digits of the SIC code, to account for omitted effects at the industry level. Year and state dummies are included to control for year- and state-specific effects. State×year fixed effects are included to account for time-varying state-level effects. Table 4 reports summary statistics for all variables studied. All continuous variables are winsorized at the 1^{st} and 99^{th} percentiles of their empirical distribution. Detailed variable definition is reported in Appendix A.

[Table 4 about here.]

4. Results

4.1. The Effect of 10-K Readability on Future Shareholder Proposals

In testing for Hypothesis 1, we estimate Model 1 using logit regressions. The dependent variable is *SP*, which is a binary variable that takes a value of one if the firm receives shareholder proposal(s) in the proxy season, and zero otherwise. *Fog index* is employed as the main measure for financial disclosure readability. Table 5 reports the test results. All specifications include different fixed effects structure: Column (1) includes industry fixed effects (FE) only; Column (2) includes industry and year FE; Column (3) contains industry, year, and state FE; and Column (4) includes industry and state×year FE. The coefficient on *Fog index* is positive and statistically significant across all specifications, suggesting that less readable 10-K reports are associated with a higher probability of shareholder proposals subsequent to the 10-K date. Goodness of Fit tests are performed on all model specifications with ROC curves, which indicate good fits; and none of the F-statistics indicate any concerns about it. Therefore, we find evidence in support of Hypothesis 1.

[Table 5 about here.]

Table 6 reports the predicted probabilities[4] of shareholder proposals based on the results reported in Table 5. Probabilities are predicted at varying values for the *Fog index*, with all other variables set at their mean values.

[Figure 1 and Table 6 about here.]

It is possible that findings reported in Table 5 are driven by unobservable attributes or imperfect controls for correlated firm characteristics, rather than by the 10-K readability *per se*. To provide better identification of the effect of financial disclosure readability on future shareholder proposals, we exploit the introduction of the Plain Writing Act

[4]The logit coefficient can be transformed into probability by taking the exponential of the coefficient and then dividing the exponential value by the sum of one and the exponential value.

in 2010. This 'quasi-natural' experiment allows a more valid causal inference using a difference-in-differences (diff-in-diff) specification as follows:

$$SP_{i,t} = \beta_0 + \beta_1 Treated_i + \beta_2 Post_t + \beta_3 Treated_i \times Post_t + \sum_{j=4}^{21} \beta_j Control_{i,t-1} + FE + \epsilon_{i,t} \tag{2}$$

where *Treated* is a dummy variable that equals one (zero) if a firm's *Fog index* is above (below) the 90th (10th) percentile of the empirical distribution of the Fog index as of 2009 (i.e., The year prior to the implementation of the Plain Writing Act 2010), as depicted in Figure 2. The idea is that not all firms are affected by the Plain Writing Act equally: firms in the highest (lowest) decile of the distribution of *Fog index* are most (least) likely to be affected by the Act. *Post* is a dummy indicator that equals one after the implementation of the Plain Writing Act (from 2010), and zero before this period. The coefficient of interest is β_3, which measures the difference-in-changes in the probability of shareholder proposals for the treated firms relative to the control firms. If β_3 is statistically significant, then the Plain Writing Act has an impact on firm disclosure readability and in turn the subsequent shareholder proposals. The diff-in-diff approach ensures that model estimation is not influenced by permanent and unobserved differences between treated and control groups or by common trends. All other variables are as defined in Model 1.

[Figure 2 about here.]

We perform entropy balancing (Hainmueller, 2012) to ensure that firms in treated and control groups are comparable. Entropy balancing appropriately reweights units to obtain balance, while keeping the weights as close as possible to the base weights. It thus retains valuable information in the processed data and improves efficiency for the subsequent analysis (Hainmueller, 2012). We start with estimating the following model to identify the matching covariates for entropy balancing:

$$Treated_i = \zeta_0 + \sum_{j=1}^{14} \zeta_j Covariate_{i,t-1} + FE + \epsilon \tag{3}$$

where *Treated* is as previously defined and *Covariate* includes the set of control variables identical to that in Model 1. We also include industry, year, and state fixed effects (FE). Appendix B1 reports results from estimating Model 3 for the pre-treatment sample (i.e., 2006–2009). We identify three firm characteristics that distinguish treated from control firms: *Size*, *Market-to-book*, and $\Delta Sale$. Any significant distributional disparity between these variables for treated and control firms can potentially weaken inference from our diff-in-diff analysis. We match on the first and second moments of these three matching covariates with a tolerance level of 0.015 (Hainmueller and Xu, 2013).[5] We perform matching based on the year before the implementation of the Plain Writing Act (2009). Appendix B2 reports descriptive statistics on matching covariates for both unbalanced and entropy-balanced samples as of 2009. Panel B shows that the mean and variance of the treatment and weighted control groups are identical after entropy balancing, confirming the efficacy of our entropy balancing procedure. Statistics are similar for other years but are not reported for brevity.

Table 7 reports the diff-in-diff analysis results for the treatment effect of the Plain Writing Act on the likelihood of shareholder proposals. Columns (1) and (2) are based on unbalanced and entropy-balanced samples, respectively. The coefficient on $Treated \times Post$ is found to be significant (p=0.05) and negative in both columns. This means that firms with the most unreadable 10-Ks before the treatment receive fewer shareholder proposals in the post-treatment period due to improved 10-K readability. These results allow a more valid inference about causality running from 10-K readability to the likelihood of a shareholder proposal being filed.

[Table 7 about here.]

[5]The tolerance level refers to the maximum deviation from the moment conditions across all the variables included in the set of covariates.

4.2. The Effect of 10-K Readability on Future Shareholder Proposals in the Presence of Mergers and Acquisitions

We further test the effect of 10-K readability on the likelihood of shareholder proposals in the presence of merger and acquisition (M&A) activities. Firms involved in M&A activities tend to provide more detailed disclosure that can lead to less readable 10-K reports (Lim et al., 2022). Firms are required to disclose material M&A transaction agreements in Item 1.01 of Form 8-K. This setting thus allows us to gauge the incremental impact of M&A activities on the effect of 10-K readability on the likelihood of shareholder proposals by estimating the following model:

$$\begin{aligned} SP_{i,t} = \gamma_0 + \gamma_1 Fog\ index_{i,t-1} + \gamma_2 M\&A_{i,t-1} + \gamma_3 Fog\ index_{i,t-1} \\ \times M\&A_{i,t-1} + \sum_{j=4}^{21} \gamma_j Control_{i,t-1} + FE + \epsilon_{i,t} \end{aligned} \tag{4}$$

where $M\&A$ is a binary variable that takes a value of one if a firm reports material M&A transactions in Item 1.01 of Form 8-K prior to the 10-K date for the same fiscal year, and zero otherwise. All other variables are as defined in Model 1. The coefficient of interest is γ_3, which captures the marginal effect of M&A activities on the relationship between 10-K readability and the likelihood of shareholder proposals.

We perform entropy balancing similar to that described in the previous section. Appendix C1 reports results for identifying matching covariates for the entropy balancing while Appendix C2 reports descriptive statistics of the first three moments of the matching covariates for both unbalanced and entropy-balanced samples. We report test results based on both unbalanced and entropy-balanced samples in columns (1) and (2) of Table 8, respectively. The coefficient for *Fog index*$\times M\&A$ is positive and significant in both specifications, indicating that M&A activities exacerbate the negative impact of 10-K readability on the likelihood of shareholder proposals.

[Table 8 about here.]

4.3. The Effect of 10-K Readability on Future Shareholder Proposals in the Presence of Restatements

Similarly, we examine the effect of 10-K readability on the likelihood of shareholder proposals in the presence of accounting restatements.

$$SP_{i,t} = \delta_0 + \delta_1 Fog\ index_{i,t-1} + \delta_2 Restate_{i,t-1} + \delta_3 Fog\ index_{i,t-1} \times Restate_{i,t-1} + \sum_{j=4}^{21} \delta_j Control_{i,t-1} + FE + \epsilon_{i,t} \quad (5)$$

where *Restate* is a dummy variable that equals one if a firm restates the results in Form 8-K after the 10-K date for the same fiscal year and before the shareholder meeting date due to intentional misreporting, and zero otherwise. Following Hennes et al. (2008), we classify a restatement as intentional misreporting when it is associated with a subsequent investigation by the audit committee, Department of Justice, or the SEC, or if the disclosure contained the words 'fraud' or 'irregularity'. 40 out of 1,060 firms restated their financial reports during our sample period. Appendix D1 reports results for identifying matching covariates for the entropy balancing while Appendix D2 reports descriptive statistics of the first three moments of the matching covariates for both unbalanced and entropy-balanced samples. We report test results using both unbalanced and entropy-balanced samples in columns (1) and (2) of Table 9, respectively. The coefficient for *Fog index*×*Restate* is negative and significant while the coefficients for *Fog index* and *Restate* standalone are positive and significant in both specifications. This suggests that firms with low 10-K readability are subject to lower probability of receiving shareholder proposals after they restate 10-K information prior to the shareholder meeting date. Restatements, therefore, render firms a communication tool to clarify key material information.

[Table 9 about here.]

4.4. The Effect of Earnings Surprises on 10-K Readability

The underlying premise of Hypothesis 1 is the 'management obfuscation hypothesis', which infers one potential driver for the positive effect of 10-K readability on the likelihood

of shareholder proposals is that managers make disclosure more difficult to process by writing more complex 10-K reports to delay the release of negative news. If this is deemed to be true, one should observe a decline in disclosure complexity or opaqueness after the negative news is released. We test this conjecture in a setting of earnings surprises by estimating the following model using OLS:

$$\begin{aligned}\Delta Fog\ index_{i,t} = \lambda_0 + \lambda_1 Negative\ ES_{i,t-1} + \lambda_2 Positive\ ES_{i,t-1} \\ + \lambda_3 GOV\ SP_{i,t-1} + \sum_{j=4}^{17} \lambda_j Control_{i,t-1} + FE + \epsilon_{i,t}\end{aligned} \tag{6}$$

where the dependent variable is the change in the *Fog index* (Δ*Fog index*). *Negative ES* is a dummy variable that equals one if the actual earnings are smaller than the average estimates at the release of earnings, and zero otherwise. *Positive ES*, in turn, is a dummy indicator that equals one if the actual earnings are greater than the mean estimates at the release of earnings, and zero otherwise. We also control for the effect of corporate governance shareholder proposals (*GOV SP*), a dummy indicator that equals one if a firm receives corporate governance shareholder proposal(s) in a given year, on 10-K readability to account for any impact of shareholder proposals on financial disclosure readability. All other variables are as defined in Model 1. Standard errors are adjusted for clustering at the firm level. As shown in Table 10, both *Negative ES* and *Positive ES* are negatively associated with Δ*Fog index*, implying that firms' 10-K readability improves after the release of earnings surprises, an effect that is more pronounced for negative earnings surprises as suggested by both the size of coefficient and level of statistical significance for *Negative ES* relative to those for *Positive ES*, which is in line with the underlying premise of Hypothesis 1 as discussed above.

[Table 10 about here.]

5. Robustness

5.1. Alternative Readability Measures

In this section, we check the robustness of our results to the choice of readability measure. We employ a set of alternative readability measures that are either quantity-based or formula-based. For quantity-based readability measures, we employ the percentage of complex words (*Complex words %*), average words per sentence (*AWPS*), and the natural logarithm of the 10-K file size (*Log(file size)*), and the natural logarithm of the 10-K word count (*Log(#words)*). For formula-based readability measures, we use *Flesch-Kincaid index* and *SMOG index*. Table 11 reports the correlation between the Fog index and alternative readability measures. The Fog index is positively and significantly correlated with all alternative readability measures at the 1% level.

[Table 11 about here.]

Table 12 reports regression results using alternative readability measures. Columns (1) and (2) report results based on *Complex words %* and *AWPS* separately while column (3) presents results including both measures. The coefficient for *Complex words %* is positive and statistically significant at the 1% level while the coefficient for *AWPS* is positive but not statistically significant in column (2) and significant at the 10% level in column (3). Since *Complex words %* and *AWPS* are the two constructs of the *Fog index*, these results also confirm the predictability of the *Fog index* on future shareholder proposals is primarily driven by the use of complex words rather than the length of 10-K. In addition, *Log(file size)* is a significant predictor of future shareholder proposals (column 4). However, the coefficient for *Log(#words)*, as shown in column (5), is not statistically significant. This finding, together with the results based on *AWPS*, is probably due to the shortcomings of word count-based measures, which are the outcome of trading off between writing closely and succinctly given that "*writing a disclosure in plain English can sometimes increase the length of particular sections*" (Securities and Exchange Commission, 1998). Similar claims are also made in Bloomfield (2008), which discusses the classification of length-based measures based on their ontological versus obfuscation util-

ities; and specifies how the complexity of a business might require it to have lengthier disclosure (ontological argument); and this need not necessarily be done for the purpose of obfuscating information. Bonsall IV et al. (2017) note that quantity-based measures are necessarily limited metrics of plain English readability because they only capture a single plain attribute: superfluous words. Lastly, we find positive and significant results for both formula-based measures, reported in columns 6 and 7 for *Flesch-Kincaid index* and *SMOG index*, respectively.

[Table 12 about here.]

5.2. Proposal Type

We perform additional analyses to test the 10-K readability to the likelihood of observing one of the following types of shareholder proposals: corporate governance proposals, social proposals, and non-omitted proposals. As shown in Panel A of Table 13, the likelihood of observing a corporate governance proposal at the annual shareholder meeting is an increasing function of *Fog index* (p=0.01) while this effect is absent for social proposals, as reported in Panel B. This finding is in line with the premise that the proxy process mainly provides activist shareholders with a formal mechanism to raise concerns about corporate governance and corporate performance. Another way to think of this result is that activists prefer to target more opaque firms to maintain an information advantage over prices, mostly in an institutional investor setting (Bourveau and Schoenfeld, 2017; Maffett, 2012). Panel C, in turn, reports results based on a sample excluding omitted proposals.[6] The coefficients of interest remain positive with increased magnitude and significance across all specifications.

[Table 13 about here.]

[6]SEC Rule 14a-8(c) permits management to omit shareholder proposals for thirteen separate reasons.

5.3. Controlling for ESG Performance

We perform robustness tests to account for the impact of ESG performance on shareholder activism. ESG performance is known as a driver for shareholder activism. Larger companies that engage in poor ESG practices are frequently targeted by activists (Rehbein et al., 2004). We control for ESG performance using the ASSET4 ESG scores across three dimensions: environmental, social, and corporate governance. Table 14 reports results for these tests. The availability of ESG data reduces the sample size to 4,739 firm-year observations. The coefficients of interest remain significant across all specifications and preserve both their magnitude and significance.

[Table 14 about here.]

5.4. Zero-Inflated Negative Binomial Regression Results

To account for the effects of excessive zeros in our sample, we employ the zero-inflated negative binomial (ZINB) regression. The ZINB model adjusts for the difference between mean and variance and add more predictions of zeros (Long, 1997). This is conducted by assuming that the population consists of two distinct latent groups: a group that has a chance of receiving shareholder proposals (group A) and the other group that does not (group B). The ZINB model thus makes it possible to consider applications for environments with many zero occurrences such that many firms do not receive shareholder proposals in a given year. The ZINB model allows each observation to have a positive probability of being part of either group. Therefore, for each observation i, group A is selected with a probability of λ_i and group B is selected with a probability of $1 - \lambda_i$. Group A only generates zero counts while group B generates positive counts, denoted as $f(y_i|X_i)$. So the decision rule can be best described below:

$$y_i = \begin{cases} 0 & \lambda \\ f(y_i|X_i) & 1-\lambda \end{cases}$$

The probability of $Y_i = y_i|X_i$ can be formally described as:

$$P(Y_i = y_i|X_i) = \begin{cases} \lambda(\alpha Z_i) + f(0|X_i)[1 - \lambda(\alpha Z_i)] & y_i = 0 \\ f(y_i|X_i)[1 - \lambda(\alpha Z_i)] & y_i > 0 \end{cases}$$

where the probability λ_i depends on the characteristics Z that determine the inflation of zeros. The probability λ_i can be estimated using any discrete model such as the logit or probit. Thus, two separate models are used to account for two distinct latent processes. First, an inflation model (a logit model) is used to predict zeros (membership in the group that has not received any shareholder proposals). Second, the negative binomial model (the count model) is used to predict the count of shareholder proposals. The expected number of shareholder proposals ($\#SP$) is a combination of the two processes. The ZINB model uses the same set of control variables to model the count in the part of the negative binomial model and the variable *Size* in the logit part of the model.

Column (1) of Table 15 reports results for the negative binomial model, the coefficient for *Fog index* is positive and statistically significant at the 5% level, suggesting a positive effect of 10-K readability on future shareholder proposals. The expected change in $\log(\#SP)$ for a one-unit increase in *Fog index* is 0.048 holding other variables constant. In other words, if *Fog index* increases by one unit, the expected number of SP would increase by exp(0.048), which is approximately 105%. Column (2) reports results for the inflation model, which predicts the number of SP for the zero group. *Size* is a significant predictor for the number of SP being in the zero group. The negative coefficient for *Size* suggests that if firm size decreases by one unit, the odds that it would be in the zero group would increase by a factor of exp(–0.699), equivalent to approximately 50%. Therefore, the smaller the firm, the more likely the firm is in the zero group. We also include a robust option in our ZINB model specification with results reported in column (3). If the coefficient for *alpha* (i.e., the dispersion parameter) is zero, then *lnalpha* is infinity and a Poisson model is preferred. *lnalpha* equals –16.033, suggesting that ZINB model is suitable for estimating our sample.

[Table 15 about here.]

6. Conclusions

This paper aims to extend our understanding of the role of financial disclosure readability in the context of shareholder proposals. Based on a sample of 1,063 firms listed by the S&P 1500 between 2006 and 2019, we find that firms with less readable 10-K reports are subject to higher likelihood of shareholder proposals in the future. Further analyses show that this effect is primarily driven by the use of complex words in 10-K reports rather than the length of 10-K files, and concentrated in corporate governance proposals rather than social proposals. These results are robust to several additional tests including further controlling for ESG performance, using alternative readability measures, and estimation using a zero-inflated negative binomial model. We also show that M&A activities exacerbate the negative effect of 10-K readability on the likelihood of shareholder proposals; accounting restatements reduce the likelihood of shareholder proposals for firms with complex disclosures by offering a communication tool for clarification; and firms improve disclosure readability subsequent to the release of negative news. Collectively, these findings are consistent with the 'incomplete revelation hypothesis' in suggesting that managers produce hard-to-read financial reports to delay the release of adverse information.

Our paper contributes to the literature by examining the effect of financial disclosure readability on the likelihood of future shareholder proposals—a previously unexplored implication of textual readability in financial disclosure. Our results are consistent with the 'incomplete revelation hypothesis' (IRH) (Bloomfield, 2008); and evidence unearthed in Kim et al. (2019) and Li (2008) in suggesting that managers produce hard-to-read and complex financial reports to delay the release of adverse information. We also contribute to the literature on the antecedents of shareholder activism by showing financial disclosure readability is a significant factor explaining future shareholder activism. Given that disclosure entails a useful tool for stakeholders' communication and engagement, our findings have practical implications for regulators in improving financial reporting transparency.

References

Anand, A., Basu, S., Huan, X., Pathak, J., 2022. The semantic complexity of financial disclosures. IIM Bangalore Research Paper No. 648 .

Baloria, V., Klassen, K., Wiedman, C., 2019. Shareholder activism and voluntary disclosure initiation: The case of political spending. Contemporary Accounting Research 36, 904–933.

Bauer, R., Moers, F., Viehs, M., 2015. Who withdraws shareholder proposals and does it matter? An analysis of sponsor identity and pay practices. Corporate Governance: An International Review 23, 472–488.

Bebchuk, L., 2005. The case for increasing shareholder power. Harvard Law Review 118, 833–914.

Beyer, A., Cohen, D., Lys, T., Walther, B., 2010. The financial reporting environment: Review of the recent literature. Journal of Accounting & Economics 50, 296–343.

Bloomfield, R., 2002. The 'incomplete revelation hypothesis' and financial reporting. Accounting Horizons 16, 233–243.

Bloomfield, R., 2008. Discussion of "annual report readability, current earnings, and earnings persistence". Journal of Accounting & Economics 45, 248–252.

Bonsall IV, S., Leone, A., Miller, B., Rennekamp, K., 2017. A plain English measure of financial reporting readability. Journal of Accounting & Economics 63, 329–357.

Bourveau, T., Schoenfeld, J., 2017. Shareholder activism and voluntary disclosure. Review of Accounting Studies 22, 1307–1339.

Brav, A., Jiang, W., Partnoy, F., Thomas, R., 2008. Hedge fund activism, corporate governance, and firm performance. The Journal of Finance 63, 1729–1775.

Buchanan, B., Netter, J., Poulsen, A., Yang, T., 2012. Shareholder proposal rules and practice: Evidence from a comparison of the United States and United Kingdom. American Business Law Journal 49, 739–803.

Cai, J., Walkling, R., 2011. Shareholders' say on pay: Does it create value? Journal of Financial & Quantitative Analysis 46, 299–339.

Campbell, J., Chen, H., Dhaliwal, D., Lu, H., Steele, L., 2014. The information content of mandatory risk factor disclosures in corporate filings. Review of Accounting Studies 19, 396–455.

Cziraki, P., Renneboog, L., Szilagyi, P., 2010. Shareholder activism through proxy proposals: The European perspective. European Financial Management 16, 738–777.

Del Guercio, D., Hawkins, J., 1999. The motivation and impact of pension fund activism. Journal of Financial Economics 52, 293–340.

Denes, M., Karpoff, J., McWilliams, V., 2017. Thirty years of shareholder activism: A survey of empirical research. Journal of Corporate Finance 44, 405–424.

Ertimur, Y., Ferri, F., Muslu, V., 2011. Shareholder activism and CEO pay. The Review of Financial Studies 24, 535–592.

Ertimur, Y., Ferri, F., Stubben, S., 2010. Board of directors' responsiveness to shareholders: Evidence from shareholder proposals. Journal of Corporate Finance 16, 53–72.

Ferri, F., Sandino, T., 2009. The impact of shareholder activism on financial reporting and compensation: The case of employee stock options expensing. The Accounting Review 84, 433–466.

Flammer, C., Toffel, M., Viswanathan, K., 2021. Shareholder activism and firms' voluntary disclosure of climate change risks. Strategic Management Journal 42, 1850–1879.

García Osma, B., Grande-Herrera, C., 2021. The role of users' engagement in shaping financial reporting: Should activists target accounting more? Accounting & Business Research 51, 511–544.

Gillan, S., Starks, L., 2000. Corporate governance proposals and shareholder activism: The role of institutional investors. Journal of Financial Economics 57, 275–305.

Goranova, M., Abouk, R., Nystrom, P., Soofi, E., 2017. Corporate governance antecedents to shareholder activism: A zero-inflated process. Strategic Management Journal 38, 415–435.

Goranova, M., Ryan, L., 2014. Shareholder activism: A multidisciplinary review. Journal of Management 40, 1230–1268.

Gordon, L., Pound, J., 1993. Information, ownership structure, and shareholder voting: Evidence from shareholder-sponsored corporate governance proposals. The Journal of Finance 48, 697–718.

Grossman, S., Stiglitz, J., 1980. On the impossibility of informationally efficient markets. American Economic Review 70, 393–408.

Hainmueller, J., 2012. Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. Political Analysis 20, 25–46.

Hainmueller, J., Xu, Y., 2013. ebalance: A Stata package for entropy balancing. Journal of Statistical Software 54, 1–18.

Harris, M., Raviv, A., 2010. Control of corporate decisions: Shareholders vs. management. The Review of Financial Studies 23, 4115–4147.

Hennes, K., Leone, A., Miller, B., 2008. The importance of distinguishing errors from irregularities in restatement research: The case of restatements and CEO/CFO turnover. The Accounting Review 83, 1487–1519.

Hirschman, A., 1970. Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States. Harvard University Press, Cambridge, MA.

Hwang, B.-H., Kim, H. H., 2017. It pays to write well. Journal of Financial Economics 124, 373–394.

Iliev, P., Lins, K., Miller, D., Roth, L., 2015. Shareholder voting and corporate governance around the world. The Review of Financial Studies 28, 2167–2202.

Judge, W., Gaur, A., Muller-Kahle, M., 2010. Antecedents of shareholder activism in target firms: Evidence from a multi-country study. Corporate Governance: An International Review 18, 258–273.

Karpoff, J., Malatesta, P., Walkling, R., 1996. Corporate governance and shareholder initiatives: Empirical evidence. Journal of Financial Economics 42, 365–395.

Kim, C., Wang, K., Zhang, L., 2019. Readability of 10-K reports and stock price crash risk. Contemporary Accounting Research 36, 1184–1216.

Klein, A., Zur, E., 2009. Entrepreneurial shareholder activism: Hedge funds and other private investors. The Journal of Finance 64, 187–229.

Kravet, T., Muslu, V., 2013. Textual risk disclosures and investors' risk perceptions. Review of Accounting Studies 18, 1088–1122.

Lang, M., Lundholm, R., 2000. Voluntary disclosure and equity offerings: Reducing information asymmetry or hyping the stock? Contemporary Accounting Research 17, 623–662.

Li, F., 2008. Annual report readability, current earnings, and earnings persistence. Journal of Accounting & Economics 45, 221–247.

Li, F., Lundholm, R., Minnis, M., 2013. A measure of competition based on 10-K filings. Journal of Accounting Research 51, 399–436.

Lim, E., Chalmers, K., Hanlon, D., 2022. The influence of business strategy on annual report readability. Journal of Accounting & Public Policy 37, 65–81.

Long, J., 1997. Regression models for categorical and limited dependent variables. Sage Publications, Thousand Oaks, CA.

Loughran, T., McDonald, B., 2016. Textual analysis in accounting and finance: A survey. Journal of Accounting Research 54, 1187–1230.

Maffett, M., 2012. Financial reporting opacity and informed trading by international institutional investors. Journal of Accounting & Economics 54, 201–220.

Markarian, G., Parbonetti, A., 2007. Firm complexity and board of director composition. Corporate Governance: An International Review 15, 1224–1243.

Merkley, K., 2014. Narrative disclosure and earnings performance: Evidence from R&D disclosures. The Accounting Review 89, 725–757.

Michelon, G., Rodrigue, M., Trevisan, E., 2020. The marketization of a social movement: Activists, shareholders and CSR disclosure. Accounting, Organizations & Society 80.

Rehbein, K., Waddock, S., Graves, S., 2004. Understanding shareholder activism: Which corporations are targeted? Business & Society 43, 239–267.

Renneboog, L., Szilagyi, P., 2011. The role of shareholder proposals in corporate governance. Journal of Corporate Finance 17, 167–188.

Ryan, L., Schneider, M., 2002. The antecedents of institutional investor activism. Academy of Management Review 27, 554–573.

Schrand, C., Walther, B., 2000. Strategic benchmarks in earnings announcements: The selective disclosure of prior-period earnings components. The Accounting Review 75, 151–177.

Securities and Exchange Commission, 1998. A plain English handbook: How to create clear SEC disclosure documents. www.sec.gov/pdf/handbook.pdf.

Skinner, D., 1994. Why firms voluntarily disclose bad news. Journal of Accounting Research 32, 38–60.

Smith, M., 1996. Shareholder activism by institutional investors: Evidence from CalPERS. The Journal of Finance 51, 227–252.

Smith, M., 2008. Default risk, shareholder advantage, and stock returns. Review of Financial Studies 21, 2743–2778.

Strickland, D., Wiles, K., Zenner, M., 1996. A requiem for the USA is small shareholder monitoring effective? Journal of Financial Economics 40, 319–338.

Thomas, R., Cotter, J., 2007. Shareholder proposals in the new millennium: Shareholder support, board response, and market reaction. Journal of Corporate Finance 13, 368–391.

Yin, S., Chevapatrakul, T., Yao, K., 2022. The causal effect of improved readability of financial reporting on stock price crash risk: Evidence from the Plain Writing Act of 2010. Economics Letters 216.

Figure 1: The Effect of 10-K Readability on Future Shareholder Proposals: Predicted Probability

This figure plots predicted probabilities for the shareholder proposals as reported in Table 6.

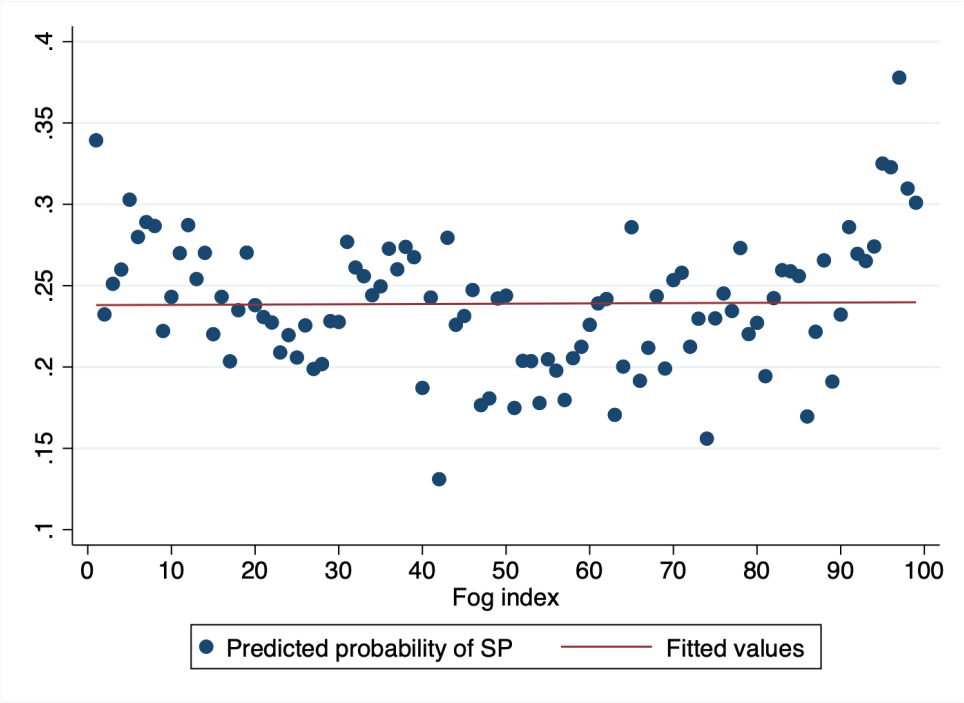

Figure 2: Firm Assignment to Treated and Control Groups

This figure depicts the firm assignment to treated and control groups, based on the empirical distribution of *Fog index*. Firms in the top decile are assigned to the treated group while firms in the bottom decile are assigned to the control group.

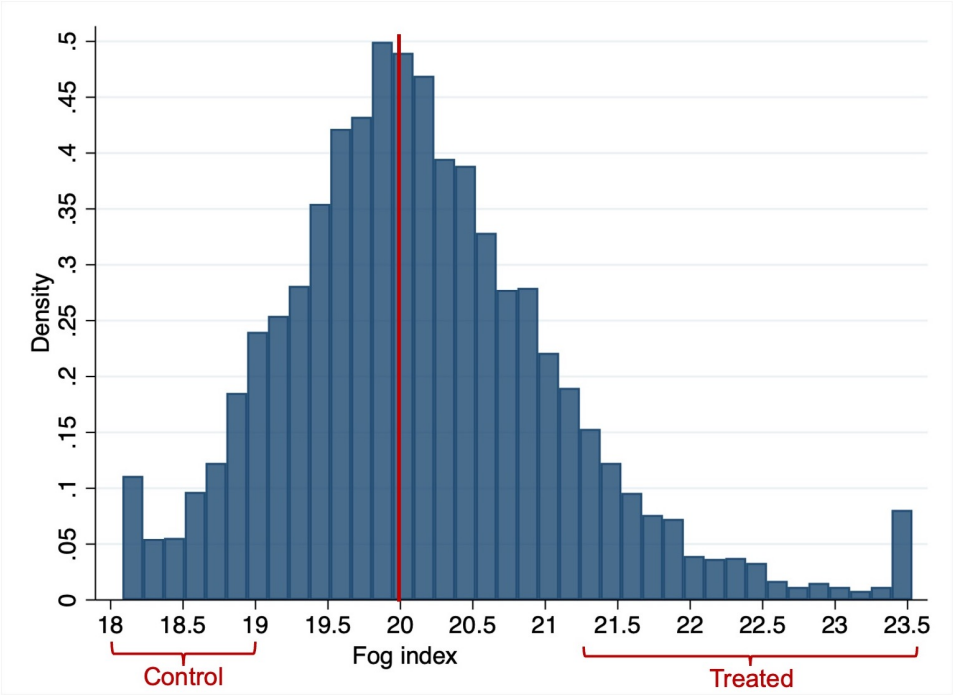

Table 1: Sample Selection

Selection criterion	Firm-year observations	Number of firms
Firm-year observations for S&P 1500 firms between 2006 and 2019	20,473	2,474
Firm-year observations with required Compustat data	9,891	1,242
Firm-year observations with required BoardEx data	9,420	1,194
Firm-year observations with required Factset data	9,195	1,158
Firm-year observations with required 10-K text attribute data	8,875	1,152
Removing financial institutions	8,730	1,132
Firm-year observations converged in the logit regression	7,786	1,063
Final sample	7,786	1,063

Table 2: Number of Proposals Per Firm

This table reports the number of shareholder proposals per firm over the 2006–2019 proxy season for U.S. firms listed by the S&P 1500.

Number of proposals per firm	Frequency	Percent
0	5,921	76.05
1	967	12.42
2	353	4.53
3	207	2.66
4	104	1.34
5	57	0.73
6	56	0.72
7	42	0.54
8	24	0.31
9	12	0.15
10	11	0.14
11	9	0.12
12	3	0.04
13	4	0.05
14	8	0.10
17	2	0.03
18	1	0.01
19	1	0.01
20	2	0.03
22	1	0.01
27	1	0.01
Total	7,786	100.00

Table 3: Sample Composition

This table reports sample composition. Panel A reports sample composition by year. Panel B reports sample composition by industry.

Panel A. Sample Composition by Year

Year	Frequency	Percent	Cumulative
2006	558	7.17	7.17
2007	564	7.24	14.41
2008	529	6.79	21.20
2009	560	7.19	28.40
2010	573	7.36	35.76
2011	565	7.26	43.01
2012	583	7.49	50.50
2013	575	7.39	57.89
2014	576	7.40	65.28
2015	542	6.96	72.25
2016	556	7.14	79.39
2017	469	6.02	85.41
2018	576	7.40	92.81
2019	560	7.19	100.00
Total	7,786	100.00	

Panel B. Sample Composition by Industry

Industry	Frequency	Percent	Cumulative
Consumer nondurables	369	4.74	4.74
Consumer durables	314	4.03	8.77
Manufacturing	1,405	18.05	26.82
Energy, oil, gas, and coal	136	1.75	28.56
Chemicals and allied products	431	5.54	34.10
Business equipment	2,424	31.13	65.23
Telecommunications	40	0.51	65.75
Wholesale and retail	1,326	17.03	82.78
Healthcare	1,180	15.16	97.93
Other	161	2.07	100.00
Total	7,786	100.00	

Table 4: Summary Statistics

This table presents descriptive statistics of the variables studied in this paper. N and SD refer to the number of observations and the standard deviation, respectively. Min and Max refer to the minimum and maximum values, respectively. All variables are as defined in Appendix A.

	N	Mean	Median	SD	Min	Max
SP	7,786	0.240	0.000	0.427	0.000	1.000
Fog index	7,786	20.146	20.056	0.978	18.082	23.533
Tone	7,786	-0.893	-0.880	0.403	-2.052	0.007
Uncertain	7,786	1.469	1.469	0.256	0.839	2.103
Litigious	7,786	1.012	0.898	0.452	0.387	2.843
Size	7,786	7.759	7.600	1.577	4.806	12.029
External complexity	7,786	0.019	0.003	0.041	0.000	0.253
Internal complexity	7,786	0.058	0.003	0.177	0.000	1.310
Cash holding	7,786	0.183	0.131	0.167	0.004	0.725
Dividend payout	7,786	0.197	0.017	0.435	-1.410	2.582
Leverage	7,786	0.496	0.495	0.221	0.085	1.148
Market-to-book	7,786	2.166	1.813	1.187	0.822	7.132
Tangibility	7,786	0.198	0.149	0.157	0.014	0.709
Return on assets	7,786	0.082	0.084	0.102	-0.331	0.338
ΔSales	7,786	0.079	0.067	0.170	-0.410	0.757
Board size	7,786	9.248	9.000	2.083	5.000	15.000
Board independence	7,786	1.047	0.889	0.386	0.600	2.400
Board diversity	7,786	0.000	-0.055	1.031	-1.376	3.857
Institutional ownership	7,786	0.859	0.889	0.132	0.427	1.000
Industry spillover	7,786	26.838	24.000	19.242	1.000	71.000

Table 5: The Effect of 10-K Readability on Future Shareholder Proposals: Baseline Results

This table reports regression results related to the effect of 10-K readability on future shareholder proposals from estimating Model 1. The logit regressions include industry, state, year, and state×year fixed effects in different combinations (not reported). *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. Robust z-statistics are reported in parentheses. All variables are as defined in Appendix A.

	(1) SP	(2) SP	(3) SP	(4) SP
Fog index	**0.091****	**0.090****	**0.109****	**0.133*****
	(2.297)	**(2.197)**	**(2.524)**	**(2.914)**
Tone	-0.099	-0.109	-0.170	-0.185
	(-0.807)	(-0.892)	(-1.328)	(-1.368)
Uncertain	-0.084	-0.026	-0.019	-0.009
	(-0.497)	(-0.150)	(-0.104)	(-0.049)
Litigious	0.059	0.061	0.009	-0.024
	(0.533)	(0.544)	(0.080)	(-0.197)
Size	0.935***	0.938***	1.018***	1.076***
	(21.006)	(20.877)	(21.345)	(21.334)
External complexity	11.192***	11.121***	10.692***	11.963***
	(6.949)	(6.879)	(6.336)	(6.674)
Internal complexity	0.707*	0.711*	0.910**	1.062***
	(1.943)	(1.948)	(2.474)	(2.796)
Cash holding	0.005	0.036	0.205	0.038
	(0.014)	(0.110)	(0.592)	(0.104)
Dividend payout	0.079	0.083	0.089	0.067
	(0.942)	(0.978)	(1.051)	(0.729)
Leverage	-0.367*	-0.345*	-0.492**	-0.585***
	(-1.825)	(-1.702)	(-2.328)	(-2.611)
Market-to-book	0.259***	0.254***	0.305***	0.325***
	(6.611)	(6.221)	(7.294)	(7.361)
Tangibility	1.330***	1.322***	1.183***	1.138***
	(3.907)	(3.864)	(3.314)	(3.029)
Return on assets	0.636	0.651	0.300	0.334
	(1.241)	(1.248)	(0.562)	(0.597)
ΔSales	-1.074***	-1.170***	-1.175***	-1.331***
	(-4.555)	(-4.708)	(-4.650)	(-4.979)
Board size	0.036	0.035	0.030	0.034
	(1.633)	(1.602)	(1.286)	(1.408)
Board independence	0.179**	0.175*	0.207**	0.224**
	(1.967)	(1.908)	(2.151)	(2.237)
Board diversity	-0.097**	-0.099**	-0.096**	-0.087**
	(-2.483)	(-2.465)	(-2.325)	(-2.013)
Institutional ownership	0.056	0.017	-0.037	0.001
	(0.178)	(0.053)	(-0.113)	(0.002)
Industry spillover	-0.002	-0.008*	-0.008	-0.010*
	(-0.532)	(-1.704)	(-1.608)	(-1.949)
Constant	-12.138***	-11.909***	-12.904***	-14.110***
	(-11.519)	(-11.061)	(-8.461)	(-5.083)
Industry FE	Yes	Yes	Yes	Yes
Year FE	No	Yes	Yes	No
State FE	No	No	Yes	No
State×Year FE	No	No	No	Yes
Observations	7,786	7,786	7,786	7,786
Pseudo R^2	0.349	0.350	0.365	0.401
ROC curve	0.869	0.870	0.877	0.894

Table 6: The Effect of 10-K Readability on Future Shareholder Proposals: Predicted Probability

This table presents the predicted probabilities for the shareholder proposals from results reported in column (4) of Table 5. Probabilities are predicted with values of the *Fog index* varied and all other variables set at their mean values.

	Mean predicted probability of SP
Fog index (1st decile)	27.06%
Fog index (2nd decile)	24.92%
Fog index (3rd decile)	21.74%
Fog index (4th decile)	25.49%
Fog index (5th decile)	22.00%
Fog index (6th decile)	19.86%
Fog index (7th decile)	22.36%
Fog index (8th decile)	22.86%
Fog index (9th decile)	22.92%
Fog index (10th decile)	30.33%

Table 7: The Effect of 10-K Readability on Future Shareholder Proposals: The Plain Writing Act 2010

This table reports the diff-in-diff analysis results concerning the effect of 10-K readability on future shareholder proposals based on Model 2. The logit regressions include industry and state×year fixed effects. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. Robust z-statistics are reported in parentheses. All variables are as defined in Appendix A.

	(1) SP	(2) SP
Treated	3.398***	4.336***
	(4.428)	(4.325)
Post	0.054	0.158
	(0.012)	(0.114)
Treated×Post	**-1.557****	**-1.942****
	(-2.051)	**(-2.177)**
Tone	-1.349*	-2.115***
	(-1.912)	(-2.678)
Uncertain	-0.628	-1.124
	(-0.685)	(-1.117)
Litigious	-0.260	-0.712
	(-0.489)	(-0.967)
Size	2.312***	2.706***
	(7.748)	(5.959)
External complexity	9.969	8.064
	(0.906)	(0.505)
Internal complexity	-0.518	0.923
	(-0.091)	(0.164)
Cash holding	-4.843*	-8.070**
	(-1.725)	(-2.495)
Dividend payout	0.942**	0.887**
	(2.526)	(2.146)
Leverage	-1.181	-0.275
	(-0.774)	(-0.122)
Market-to-book	1.424***	1.595***
	(5.137)	(5.620)
Tangibility	5.469**	4.508
	(2.065)	(1.611)
Return on assets	3.449	5.791
	(0.984)	(1.405)
ΔSales	-5.428***	-5.456***
	(-4.111)	(-3.967)
Board size	0.080	0.172
	(0.675)	(1.247)
Board independence	-0.946*	-1.043
	(-1.723)	(-1.470)
Board diversity	-0.553**	-0.717***
	(-2.230)	(-2.668)
Institutional ownership	-2.143	-2.711
	(-1.110)	(-1.363)
Industry spillover	-0.012	-0.018
	(-0.492)	(-0.606)
Constant	-20.575***	-23.898***
	(-4.386)	(-4.464)
Sample	Unbalanced	Entropy balanced
Industry FE	Yes	Yes
State×Year FE	Yes	Yes
Observations	898	898
Pseudo R^2	0.681	0.681
ROC curve	0.967	0.967

Table 8: The Effect of 10-K Readability on Future Shareholder Proposals: Mergers and Acquisitions

This table reports regression results concerning the effect of 10-K readability on future shareholder proposals in the presence of mergers and acquisitions from estimating Model 4. The logit regressions include industry and state×year fixed effects. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. Robust z-statistics are reported in parentheses. All variables are as defined in Appendix A.

	(1) SP	(2) SP
Fog index	0.099**	0.104
	(2.005)	(1.551)
M&A	-3.316*	-4.588**
	(-1.845)	(-2.354)
Fog index×M&A	**0.159***	**0.220****
	(1.796)	**(2.291)**
Tone	-0.186	-0.264
	(-1.374)	(-1.498)
Uncertain	-0.007	0.281
	(-0.036)	(1.075)
Litigious	-0.033	-0.032
	(-0.264)	(-0.201)
Size	1.085***	1.139***
	(21.321)	(16.837)
External complexity	11.941***	15.897***
	(6.669)	(5.212)
Internal complexity	1.078***	1.513***
	(2.836)	(2.749)
Cash holding	0.007	-0.265
	(0.019)	(-0.517)
Dividend payout	0.072	0.057
	(0.793)	(0.553)
Leverage	-0.596***	-0.696**
	(-2.657)	(-2.084)
Market-to-book	0.324***	0.366***
	(7.334)	(5.663)
Tangibility	1.086***	1.586***
	(2.867)	(3.193)
Return on assets	0.272	0.097
	(0.484)	(0.114)
ΔSales	-1.289***	-1.538***
	(-4.750)	(-4.581)
Board size	0.033	0.083**
	(1.346)	(2.567)
Board independence	0.221**	0.141
	(2.208)	(1.147)
Board diversity	-0.087**	-0.074
	(-2.022)	(-1.306)
Institutional ownership	-0.006	0.263
	(-0.017)	(0.575)
Industry spillover	-0.011**	-0.004
	(-1.974)	(-0.644)
Constant	-13.470***	-16.570***
	(-4.794)	(-7.585)
Sample	Unbalanced	Entropy balanced
Industry FE	Yes	Yes
State×Year FE	Yes	Yes
Observations	7,786	7,786
Pseudo R^2	0.401	0.473
ROC curve	0.894	0.894

Table 9: The Effect of 10-K Readability on Future Shareholder Proposals: Restatements

This table reports regression results concerning the effect of 10-K readability on future shareholder proposals in the presence of restatements from estimating Model 5. The logit regressions include industry and state×year fixed effects. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. Robust z-statistics are reported in parentheses. All variables are as defined in Appendix A.

	(1) SP	(2) SP
Fog index	0.185***	0.210***
	(3.813)	(2.677)
Restate	21.389**	24.590***
	(2.084)	(3.740)
Fog index×Restate	**-1.035****	**-1.194*****
	(-2.006)	**(-3.674)**
Tone	-0.093	0.318
	(-0.644)	(1.391)
Uncertain	0.163	1.624***
	(0.794)	(4.329)
Litigious	0.130	0.568***
	(1.000)	(2.771)
Size	1.016***	0.946***
	(19.420)	(10.922)
External complexity	8.801***	16.865***
	(5.540)	(5.493)
Internal complexity	1.197***	1.321***
	(3.062)	(3.090)
Cash holding	0.204	0.652
	(0.523)	(1.061)
Dividend payout	0.037	0.331***
	(0.373)	(2.748)
Leverage	-0.432*	-0.250
	(-1.784)	(-0.630)
Market-to-book	0.268***	0.190***
	(5.608)	(2.578)
Tangibility	1.450***	2.967***
	(3.568)	(4.666)
Return on assets	-0.044	0.921
	(-0.074)	(0.942)
ΔSales	-1.065***	-2.614***
	(-3.775)	(-5.122)
Board size	0.012	0.153***
	(0.472)	(3.183)
Board independence	0.473***	-0.041
	(4.566)	(-0.193)
Board diversity	-0.087*	-0.018
	(-1.885)	(-0.280)
Institutional ownership	0.058	0.651
	(0.155)	(1.089)
Industry spillover	-0.008	0.001
	(-1.410)	(0.137)
Constant	-17.341***	-24.123***
	(-6.247)	(-8.484)
Sample	Unbalanced	Entropy balanced
Industry FE	Yes	Yes
State×Year FE	Yes	Yes
Observations	7,649	7,649
Pseudo R^2	0.387	0.612
ROC curve	0.891	0.891

Table 10: The Effect of Earnings Surprises on 10-K Readability

This table reports results concerning the effect of earnings surprises on subsequent 10-K readability from estimating Model 6. The OLS regressions include industry and state×year fixed effects. Standard errors are adjusted for clustering at the firm level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. Robust t-statistics are reported in parentheses. All variables are as defined in Appendix A.

	(1) ΔFog index	(2) ΔFog index	(3) ΔFog index	(4) ΔFog index
Negative ES	**-0.352*** **	**-0.275** **	**-0.282** **	**-0.312** **
	(-2.657)	**(-2.165)**	**(-2.172)**	**(-2.395)**
Positive ES	**-0.295** **	**-0.204***	**-0.207***	**-0.210***
	(-2.535)	**(-1.818)**	**(-1.794)**	**(-1.830)**
GOV SP	-0.023	0.034	0.026	0.058
	(-0.196)	(0.297)	(0.219)	(0.482)
Size	-0.005	-0.015	-0.026	-0.028
	(-0.143)	(-0.494)	(-0.812)	(-0.861)
External complexity	0.279	0.351	0.647	0.425
	(0.339)	(0.411)	(0.741)	(0.477)
Internal complexity	-0.218	-0.190	-0.205	-0.212
	(-1.236)	(-1.201)	(-1.229)	(-1.266)
Cash holding	0.198	0.044	0.012	0.048
	(0.869)	(0.199)	(0.052)	(0.202)
Dividend payout	-0.063	0.009	0.015	0.022
	(-0.710)	(0.099)	(0.163)	(0.234)
Leverage	-0.292*	-0.105	-0.121	-0.108
	(-1.790)	(-0.659)	(-0.733)	(-0.632)
Market-to-book	0.032	0.059*	0.065*	0.059*
	(0.880)	(1.661)	(1.795)	(1.674)
Tangibility	-0.027	-0.124	-0.258	-0.141
	(-0.106)	(-0.485)	(-0.912)	(-0.500)
Return on assets	-0.424	-0.538	-0.550	-0.523
	(-0.939)	(-1.251)	(-1.257)	(-1.247)
ΔSales	0.186	-0.204	-0.226	-0.288
	(0.744)	(-0.814)	(-0.896)	(-1.213)
Board size	-0.007	-0.017	-0.012	-0.017
	(-0.360)	(-0.866)	(-0.591)	(-0.845)
Board independence	-0.007	-0.026	-0.026	-0.031
	(-0.092)	(-0.378)	(-0.367)	(-0.461)
Board diversity	-0.025	0.006	0.006	0.014
	(-0.871)	(0.222)	(0.223)	(0.460)
Institutional ownership	-0.189	-0.037	-0.026	-0.028
	(-0.936)	(-0.189)	(-0.124)	(-0.134)
Constant	0.805***	0.693**	0.753**	0.811***
	(2.628)	(2.338)	(2.468)	(2.650)
Industry FE	Yes	Yes	Yes	Yes
Year FE	No	Yes	Yes	No
State FE	No	No	Yes	No
State×Year FE	No	No	No	Yes
Observations	8,630	8,630	8,630	8,630
R^2	0.002	0.072	0.073	0.138

Table 11: Correlation Between Readability Measures

This table reports the correlation between the Fog index and alternative readability measures.

	Fog	(1)	(2)	(3)	(4)	(5)
(1) Complex words %	0.258***	1				
(2) AWPS	0.889***	-0.210***	1			
(3) Log(#words)	0.503***	-0.072***	0.544***	1		
(4) Log(file size)	0.114***	0.107***	0.067***	0.391***	1	
(5) Flesch-Kincaid index	0.960***	0.130***	0.909***	0.547***	0.149***	1
(6) SMOG index	0.998***	0.258***	0.887***	0.501***	0.107***	0.959***

Table 12: The Effect of 10-K Readability on Future Shareholder Proposals: Alternative Readability Measures

This table reports regression results related to the effect of 10-K readability on future shareholder proposals using alternative readability measures. The logit regressions include industry, state, year, and state×year fixed effects in different combinations (not reported). *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. Robust z-statistics are reported in parentheses. All variables are as defined in Appendix A.

	(1) SP	(2) SP	(3) SP	(4) SP	(5) SP	(6) SP	(7) SP
Complex words %	**0.127*** (3.364)**		**0.141*** (3.661)**				
AWPS		**0.023 (1.245)**	**0.037* (1.912)**				
Log(file size)				**0.280*** (3.248)**			
Log(#words)					**-0.092 (-0.695)**		
Flesch-Kincaid index						**0.103** (2.136)**	
SMOG index							**0.195*** (3.037)**
Tone	-0.164	-0.166	-0.183	-0.147	-0.159	-0.171	-0.182
	(-1.217)	(-1.228)	(-1.350)	(-1.091)	(-1.180)	(-1.264)	(-1.346)
Uncertain	-0.158	-0.035	-0.095	-0.007	-0.116	-0.041	-0.010
	(-0.828)	(-0.183)	(-0.489)	(-0.036)	(-0.592)	(-0.216)	(-0.050)
Litigious	0.163	0.024	0.076	0.083	0.107	0.004	-0.023
	(1.357)	(0.185)	(0.596)	(0.708)	(0.873)	(0.029)	(-0.187)
Size	1.073***	1.078***	1.072***	1.040***	1.085***	1.074***	1.075***
	(21.222)	(21.371)	(21.217)	(20.150)	(21.034)	(21.279)	(21.322)
External complexity	11.893***	11.959***	11.915***	12.078***	11.897***	11.936***	11.972***
	(6.624)	(6.680)	(6.634)	(6.739)	(6.642)	(6.664)	(6.678)
Internal complexity	1.094***	1.074***	1.077***	1.104***	1.096***	1.058***	1.059***
	(2.882)	(2.838)	(2.831)	(2.925)	(2.900)	(2.788)	(2.787)
Cash holding	0.129	0.062	0.081	0.165	0.108	0.048	0.033
	(0.356)	(0.172)	(0.223)	(0.454)	(0.298)	(0.132)	(0.090)
Dividend payout	0.059	0.061	0.064	0.065	0.053	0.066	0.067
	(0.646)	(0.673)	(0.705)	(0.709)	(0.577)	(0.720)	(0.735)
Leverage	-0.576**	-0.594***	-0.571**	-0.653***	-0.594***	-0.586***	-0.585***
	(-2.569)	(-2.650)	(-2.548)	(-2.899)	(-2.645)	(-2.611)	(-2.612)
Market-to-book	0.325***	0.326***	0.324***	0.329***	0.326***	0.325***	0.325***
	(7.342)	(7.390)	(7.317)	(7.468)	(7.402)	(7.363)	(7.365)
Tangibility	1.112***	1.091***	1.150***	1.125***	1.060***	1.119***	1.142***
	(2.944)	(2.903)	(3.047)	(2.979)	(2.815)	(2.978)	(3.039)
Return on assets	0.233	0.254	0.322	0.285	0.127	0.302	0.340
	(0.417)	(0.454)	(0.574)	(0.511)	(0.225)	(0.540)	(0.608)
ΔSales	-1.316***	-1.316***	-1.331***	-1.300***	-1.293***	-1.322***	-1.335***
	(-4.915)	(-4.927)	(-4.972)	(-4.862)	(-4.829)	(-4.946)	(-4.991)
Board size	0.032	0.033	0.033	0.031	0.033	0.034	0.034
	(1.307)	(1.375)	(1.356)	(1.270)	(1.340)	(1.386)	(1.414)
Board independence	0.229**	0.231**	0.224**	0.228**	0.235**	0.227**	0.226**
	(2.286)	(2.308)	(2.234)	(2.273)	(2.347)	(2.269)	(2.252)
Board diversity	-0.087**	-0.089**	-0.085**	-0.091**	-0.090**	-0.087**	-0.086**
	(-2.013)	(-2.062)	(-1.967)	(-2.116)	(-2.092)	(-2.013)	(-2.004)
Institutional ownership	0.022	0.046	-0.015	0.050	0.086	0.022	-0.004
	(0.064)	(0.133)	(-0.044)	(0.145)	(0.252)	(0.064)	(-0.011)
Industry spillover	-0.010*	-0.011**	-0.010*	-0.010*	-0.011**	-0.011**	-0.010*
	(-1.927)	(-1.970)	(-1.918)	(-1.807)	(-1.990)	(-1.964)	(-1.950)
Constant	-14.886***	-12.159***	-15.986***	-15.621***	-10.756***	-13.108***	-14.849***
	(-5.425)	(-4.525)	(-5.675)	(-5.305)	(-3.640)	(-4.785)	(-5.206)
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State×Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7,786	7,786	7,786	7,786	7,786	7,786	7,786
Pseudo R^2	0.401	0.400	0.402	0.401	0.400	0.400	0.401
ROC curve	0.894	0.894	0.894	0.894	0.893	0.894	0.894

Table 13: The Effect of 10-K Readability on Future Shareholder Proposals: Proposal Type

This table reports regression results related to the effect of 10-K readability on future shareholder proposals based on proposal type. The logit regressions include industry, state, year, and state×year fixed effects in different combinations (not reported). *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. Robust z-statistics are reported in parentheses. All variables are as defined in Appendix A.

Panel A. Corporate governance proposals

	(1) SP	(2) SP	(3) SP	(4) SP
Fog index	**0.133*** **	**0.141*** **	**0.146*** **	**0.174*** **
	(3.167)	**(3.245)**	**(3.206)**	**(3.618)**
Constant	-14.530***	-14.519***	-14.660***	-17.144***
	(-12.825)	(-12.562)	(-9.159)	(-6.173)
Controls	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Year FE	No	Yes	Yes	No
State FE	No	No	Yes	No
State×Year FE	No	No	No	Yes
Observations	7,649	7,649	7,649	7,649
Pseudo R^2	0.328	0.331	0.348	0.386
ROC curve	0.861	0.862	0.872	0.891

Panel B. Social proposals

	(1) SP	(2) SP	(3) SP	(4) SP
Fog index	**-0.025**	**-0.028**	**-0.007**	**0.013**
	(-0.499)	**(-0.548)**	**(-0.137)**	**(0.220)**
Constant	-10.544***	-10.521***	-9.769***	-10.532***
	(-7.986)	(-7.805)	(-4.816)	(-3.253)
Controls	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Year FE	No	Yes	Yes	No
State FE	No	No	Yes	No
State×Year FE	No	No	No	Yes
Observations	7,335	7,335	7,335	7,335
Pseudo R^2	0.390	0.392	0.407	0.441
ROC curve	0.898	0.899	0.904	0.916

Panel C. Excluding omitted proposals

	(1) SP	(2) SP	(3) SP	(4) SP
Fog index	**0.134*** **	**0.152*** **	**0.167*** **	**0.212*** **
	(2.976)	**(3.239)**	**(3.395)**	**(4.043)**
Constant	-13.503***	-13.517***	-13.206***	-15.790***
	(-11.155)	(-10.916)	(-7.869)	(-5.362)
Controls	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Year FE	No	Yes	Yes	No
State FE	No	No	Yes	No
State×Year FE	No	No	No	Yes
Observations	7,530	7,530	7,530	7,530
Pseudo R^2	0.372	0.375	0.394	0.434
ROC curve	0.887	0.888	0.896	0.912

Table 14: The Effect of 10-K Readability on Future Shareholder Proposals: Controlling for ESG

This table reports regression results related to the effect of 10-K readability on future shareholder proposals further controlling for ESG performance. The logit regressions include industry, state, year, and state×year fixed effects in different combinations (not reported). *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. Robust z-statistics are reported in parentheses. All variables are as defined in Appendix A.

	(1) SP	(2) SP	(3) SP	(4) SP
Fog index	**0.118*** **	**0.117** **	**0.132*** **	**0.174*** **
	(2.628)	**(2.497)**	**(2.691)**	**(3.303)**
Tone	0.064	0.056	0.001	0.007
	(0.458)	(0.402)	(0.003)	(0.041)
Uncertain	0.266	0.288	0.347*	0.398*
	(1.380)	(1.449)	(1.660)	(1.799)
Litigious	0.191	0.196	0.139	0.080
	(1.489)	(1.517)	(1.032)	(0.549)
Size	0.917***	0.909***	0.981***	1.072***
	(15.745)	(15.476)	(15.657)	(15.995)
External complexity	11.079***	10.947***	10.166***	11.317***
	(6.219)	(6.124)	(5.462)	(5.724)
Internal complexity	0.916*	0.954**	1.388***	1.443***
	(1.940)	(2.013)	(2.900)	(2.820)
Cash holding	0.029	0.051	0.328	0.210
	(0.079)	(0.137)	(0.824)	(0.500)
Dividend payout	0.178*	0.181*	0.171*	0.166
	(1.898)	(1.912)	(1.782)	(1.586)
Leverage	-0.226	-0.214	-0.468**	-0.579**
	(-1.015)	(-0.956)	(-1.981)	(-2.280)
Market-to-book	0.253***	0.251***	0.302***	0.316***
	(5.650)	(5.410)	(6.298)	(6.204)
Tangibility	1.111***	1.087***	1.061**	1.049**
	(2.827)	(2.745)	(2.542)	(2.360)
Return on assets	1.316**	1.259**	0.974	1.098*
	(2.218)	(2.087)	(1.566)	(1.655)
ΔSales	-1.078***	-1.154***	-1.097***	-1.337***
	(-3.927)	(-3.981)	(-3.708)	(-4.202)
Board size	0.015	0.013	0.009	0.004
	(0.583)	(0.513)	(0.354)	(0.144)
Board independence	0.167*	0.149	0.188*	0.191*
	(1.685)	(1.489)	(1.767)	(1.707)
Board diversity	-0.163***	-0.166***	-0.165***	-0.161***
	(-3.600)	(-3.613)	(-3.483)	(-3.184)
Institutional ownership	-0.101	-0.104	-0.171	-0.028
	(-0.269)	(-0.277)	(-0.435)	(-0.067)
Industry spillover	-0.003	-0.010*	-0.009	-0.010
	(-0.627)	(-1.688)	(-1.582)	(-1.548)
ESG	0.520*	0.651**	1.087***	1.068***
	(1.945)	(2.258)	(3.584)	(3.271)
Constant	-12.932***	-12.559***	-14.375***	-17.958***
	(-10.268)	(-9.697)	(-7.767)	(-3.856)
Industry FE	Yes	Yes	Yes	Yes
Year FE	No	Yes	Yes	No
State FE	No	No	Yes	No
State×Year FE	No	No	No	Yes
Observations	4,739	4,739	4,739	4,739
Pseudo R^2	0.304	0.306	0.321	0.365
ROC curve	0.845	0.846	0.853	0.875

Table 15: The Effect of 10-K Readability on Future Shareholder Proposals: Zero-Inflated Negative Binomial Regression Results

This table reports results related to the effect of 10-K readability on future shareholder proposals based on a zero-inflated negative binomial (ZINB) regression analysis. The dependent variable is the number of shareholder proposals (#SP) that a firm receives in a given year. The regressions include year, industry, and state fixed effects (not reported). *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. Robust z-statistics are reported in parentheses. All variables are as defined in Appendix A.

	(1) #SP	(2) inflate	(3)
Fog index	**0.048****		
	(2.374)		
Tone	-0.219***		
	(-3.107)		
Uncertain	-0.023		
	(-0.230)		
Litigious	0.102*		
	(1.771)		
Size	0.729***	**-0.699*****	
	(20.979)	**(-10.369)**	
External complexity	2.323***		
	(3.545)		
Internal complexity	0.808***		
	(2.828)		
Cash holding	0.273		
	(1.236)		
Dividend payout	0.059		
	(1.359)		
Leverage	-0.393***		
	(-2.767)		
Market-to-book	0.229***		
	(8.685)		
Tangibility	0.919***		
	(4.436)		
Return on assets	0.190		
	(0.544)		
ΔSales	-0.508***		
	(-3.515)		
Board size	0.001		
	(0.118)		
Board independence	0.053		
	(1.134)		
Board diversity	-0.072***		
	(-3.173)		
Institutional ownership	-0.304		
	(-1.428)		
Industry spillover	-0.002		
	(-0.798)		
lnalpha			**-16.033*****
			(-104.619)
Constant	-10.169***	4.742***	
	(-9.955)	(7.275)	
Industry FE		Yes	
State×Year FE		Yes	
Observations		7,786	

Appendix A. Variable Definition

Variable	Definition	Source
Dependent Variables		
SP	A binary variable that takes a value of one if a firm receives shareholder proposal(s) in a given year, and zero otherwise.	ISS
#SP	The number of shareholder proposal(s) that a firm receives in a given year.	ISS
Text and Readability Attributes		
AWPS	The total number of words divided by the total number of sentences. Words are only counted if they exist occur in a sentence of five words or more.	SEC Analytics
Complex words %	The number of words that contain three or more syllables divided by the number of words in the 10-K filing.	SEC Analytics
Fog index	0.4×(AWPS + Complex words %). Higher value of Fog index implies less readable text.	SEC Analytics
Tone	The difference between the number of Loughran-McDonald financial-positive words and the number of Loughran-McDonald financial-negative words divided by the total number of words in the 10-K filing that occur in the master dictionary.	SEC Analytics
Uncertain	The number of Loughran-McDonald financial-uncertainty words divided by the total number of words in the 10-K filing that occur in the master dictionary.	SEC Analytics
Litigious	The number of Loughran-McDonald financial-litigious words divided by the total number of words in the 10-K filing that occur in the master dictionary.	SEC Analytics
Log(#words)	The natural logarithm of the word count from the 10-K filing.	SEC Analytics
Log(file size)	The natural logarithm of the file size in megabytes of the 10-K filing.	SEC Analytics
Flesch-Kincaid index	Calculated as $0.39\times\text{AWPS}+11.8\times\frac{\#Syllables}{\#Words}-15.59$	SEC Analytics
SMOG index	Calculated as $1.043\times\sqrt{\frac{\#Complex\ words\times 30}{\#Sentences}}+3.1291$	SEC Analytics
Experiment Variables		
Treated	A dummy variable that equals one (zero) if a firm's Fog index is above (below) the 90th (10th) percentile of the empirical distribution of the Fog index as of 2009 (i.e., The year prior to the implementation of the Plain Writing Act 2010).	
Post	A dummy variable that equals one after the implementation of the Plain Writing Act (from 2010), and zero before this period.	
M&A	A dummy variable that equals one if a firm reports material merger and acquisition transactions in Item 1.01 of Form 8-K prior to the 10-K date for the same fiscal year, and zero otherwise..	Audit Analytics

Variable	Definition	Source
Restate	A dummy variable that equals one if a firm restates the results in Form 8-K after the 10-K date for the same fiscal year and before the shareholders meeting date due to intentional misreporting, and zero otherwise. A restatement is classified as intentional misreporting when it is associated with a subsequent investigation by the audit committee, Department of Justice, or the SEC, or if the disclosure contained the words 'fraud' or 'irregularity' (Hennes et al., 2008).	Audit Analytics
Negative ES	A dummy variable that equals one if the actual earnings are smaller than the mean estimates at the release of earnings to the marketplace, and zero otherwise.	I/B/E/S
Positive ES	A dummy variable that equals one if the actual earnings are greater than the mean estimates at the release of earnings to the marketplace, and zero otherwise.	I/B/E/S
Control Variables		
Size	The natural logarithm of a firm's market value of equity.	Compustat
External complexity	Firm's sales (Compustat item "sale") divided by the total sales within the industry.	Compustat
Internal complexity	R&D expenditures (Compustat item "rnd") divided by the number of employees (Compustat item "emp").	Compustat
Cash holding	Cash and short-term investments (Compustat item "che") divided by total assets (Compustat item "at").	Compustat
Dividend payout	Dividends paid to ordinary shares (Compustat item "dvc") divided by income before extraordinary items (Compustat item "ib").	Compustat
Leverage	Total liabilities (Compustat item "lt") divided by total assets (Compustat item "at").	Compustat
Market-to-book	The ratio of the market value of equity to the book value of equity.	Compustat
Tangibility	Total net property, plant and equipment divided (Compustat item "ppent") by total assets (Compustat item "at").	Compustat
Return on assets	Income before extraordinary items (Compustat item "ib") divided by total assets (Compustat item "at").	Compustat
ΔSales	Annual growth in sales revenue (Compustat item "sale").	Compustat
Board size	Number of directors on the board.	BoardEx
Board independence	The ratio of number of non-executive directors to total number of directors.	BoardEx
Board diversity	The first principal component of the gender ratio and nationality mix of the board.	BoardEx
Institutional ownership	Sum of all institutional holdings in a firm's stock divided by market capitalization at the end of the calendar year.	Factset
Industry spillover	The number of shareholder proposals for each industry each year preceding the proxy proposal date of the focal company.	ISS
ESG	The sum of rating scores in three attributes: environmental, social, and corporate governance.	ASSET4

Appendix B1. The Plain Writing Act 2010: Identification of Matching Covariates

This table identifies firm characteristics that distinguish between treated and control firms for the Plain Writing Act 2010. A logit model with industry, year, and state fixed effects is estimated based on the pre-treatment sample (i.e., 2006–2009), where the dependent variable is *Treated*, which takes a value of one (zero) if a firm's Fog index is above (below) the 90th (10th) percentile of the empirical distribution of the Fog index as of 2009. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. Robust z-statistics are reported in parentheses. All variables are as defined in Appendix A.

	Treated
Size	2.674***
	(3.650)
External complexity	15.957
	(0.675)
Internal complexity	-6.543
	(-0.229)
Cash holding	-7.020
	(-1.394)
Dividend payout	0.073
	(0.091)
Leverage	2.852
	(0.731)
Market-to-book	1.603**
	(2.566)
Tangibility	2.942
	(0.507)
Return on assets	-1.832
	(-0.365)
ΔSales	-6.006***
	(-2.633)
Board size	0.090
	(0.351)
Board independence	-1.566
	(-1.553)
Board diversity	-0.245
	(-0.479)
Institutional ownership	-2.828
	(-0.859)
Constant	-23.919**
	(-2.527)
Industry FE	Yes
Year FE	Yes
State FE	Yes
Observations	270
Pseudo R^2	0.609
ROC Curve	0.956

Appendix B2. The Plain Writing Act 2010: Descriptive Statistics of Unbalanced and Entropy-Balanced Samples

This table reports descriptive statistics of unbalanced and entropy-balanced samples. The balancing is based on the year 2009, uses the first two moments of the empirical distribution of matching covariates identified in Appendix B1, and a tolerance level of 0.015.

Panel A. Unbalanced				
	Treated		Control	
	Mean	Variance	Mean	Variance
Size	8.174	2.746	7.736	2.805
Market-to-book	1.541	0.636	1.606	0.467
ΔSales	0.085	0.029	0.092	0.021

Panel B. Entropy balanced				
	Treated		Control	
	Mean	Variance	Mean	Variance
Size	8.174	2.746	8.173	2.747
Market-to-book	1.541	0.636	1.541	0.636
ΔSales	0.085	0.029	0.085	0.029

Appendix C1. Mergers and Acquisitions: Identification of Matching Covariates

This table identifies firm characteristics that distinguish between firms that are involved in mergers and acquisitions and firms that are not. A logit model with industry and state×year fixed effects is estimated, where the dependent variable is $M\&A$, which takes a value of one if a firm reports material merger and acquisition transactions in Item 1.01 of Form 8-K prior to the 10-K date for the same fiscal year, and zero otherwise. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. Robust z-statistics are reported in parentheses. All variables are as defined in Appendix A.

	M&A
Size	0.340***
	(9.254)
External complexity	2.477*
	(1.876)
Internal complexity	-0.784***
	(-2.801)
Cash holding	-2.306***
	(-7.874)
Dividend payout	0.061
	(0.784)
Leverage	-1.044***
	(-5.181)
Market-to-book	-0.133***
	(-3.419)
Tangibility	-3.953***
	(-11.536)
Return on assets	-3.297***
	(-7.635)
ΔSales	3.059***
	(14.458)
Board size	-0.001
	(-0.029)
Board independence	-0.109
	(-1.205)
Board diversity	-0.098***
	(-2.758)
Institutional ownership	-0.557**
	(-2.084)
Constant	-5.151**
	(-2.004)
Industry FE	Yes
State×Year FE	Yes
Observations	7,349
Pseudo R^2	0.200
ROC Curve	0.800

Appendix C2. Mergers and Acquisitions: Descriptive Statistics of Unbalanced and Entropy-Balanced Samples

This table reports descriptive statistics of unbalanced and entropy-balanced samples. The balancing is based on the first three moments of the empirical distribution of matching covariates identified in Appendix C1, and a tolerance level of 0.015.

Panel A. Unbalanced

	M&A=1			M&A=0		
	Mean	Variance	Skewness	Mean	Variance	Skewness
Size	8.253	2.828	0.387	7.610	2.290	0.491
External complexity	0.021	0.002	3.437	0.018	0.002	3.766
Internal complexity	0.043	0.017	6.617	0.063	0.036	4.890
Cash holding	0.158	0.020	1.299	0.190	0.030	1.198
Leverage	0.513	0.042	0.236	0.491	0.051	0.399
Market-to-book	2.021	0.946	1.982	2.209	1.540	1.799
Tangibility	0.147	0.015	1.792	0.213	0.026	1.135
Return on assets	0.066	0.009	-1.069	0.087	0.011	-0.859
ΔSales	0.119	0.034	0.883	0.067	0.027	0.688
Board diversity	0.104	1.059	0.444	-0.031	1.060	0.605
Institutional ownership	0.853	0.017	-0.827	0.861	0.017	-1.039

Panel B. Entropy balanced

	M&A=1			M&A=0		
	Mean	Variance	Skewness	Mean	Variance	Skewness
Size	8.253	2.828	0.387	8.253	2.829	0.387
External complexity	0.021	0.002	3.437	0.021	0.002	3.437
Internal complexity	0.043	0.017	6.617	0.043	0.017	6.610
Cash holding	0.158	0.020	1.299	0.158	0.021	1.300
Leverage	0.513	0.042	0.236	0.513	0.042	0.236
Market-to-book	2.021	0.946	1.982	2.021	0.948	1.983
Tangibility	0.147	0.015	1.792	0.147	0.015	1.793
Return on assets	0.066	0.009	-1.069	0.066	0.009	-1.067
ΔSales	0.119	0.034	0.883	0.119	0.034	0.883
Board diversity	0.104	1.059	0.444	0.104	1.059	0.444
Institutional ownership	0.853	0.017	-0.827	0.853	0.017	-0.827

Appendix D1. Restatements: Identification of Matching Covariates

This table identifies firm characteristics that distinguish between treated and control firms for the Plain Writing Act 2010. A logit model with industry, year, and state fixed effects is estimated based on the pre-treatment sample (i.e., 2006–2009), where the dependent variable is *Restate*, which takes a value of one if a firm restates the results in Form 8-K after the 10-K date for the same fiscal year and before the shareholders meeting date due to intentional misreporting, and zero otherwise. A restatement is classified as intentional misreporting when it is associated with a subsequent investigation by the audit committee, Department of Justice, or the SEC, or if the disclosure contained the words "fraud" or "irregularity" (Hennes et al., 2008). *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. Robust z-statistics are reported in parentheses. All variables are as defined in Appendix A.

	Restate
Size	-0.086
	(-0.476)
External complexity	-4.125
	(-0.503)
Internal complexity	-0.370
	(-0.431)
Cash holding	-0.937
	(-0.768)
Dividend payout	-0.361
	(-0.799)
Leverage	-0.551
	(-0.603)
Market-to-book	-0.016
	(-0.098)
Tangibility	-3.596**
	(-2.099)
Return on assets	-2.949*
	(-1.743)
ΔSales	0.892
	(1.131)
Board size	-0.019
	(-0.183)
Board independence	-0.637
	(-1.239)
Board diversity	0.197
	(1.213)
Institutional ownership	1.940
	(1.606)
Constant	-1.003
	(-0.412)
Industry FE	Yes
State×Year FE	Yes
Observations	1,381
Pseudo R^2	0.153
ROC Curve	0.804

Appendix D2. Restatements: Descriptive Statistics of Unbalanced and Entropy-Balanced Samples

This table reports descriptive statistics of unbalanced and entropy-balanced samples. The balancing is based on the first three moments of the empirical distribution of matching covariates identified in Appendix D1, and a tolerance level of 0.015.

Panel A. Unbalanced

	Restate=1			Restate=0		
	Mean	Variance	Skewness	Mean	Variance	Skewness
Tangibility	0.149	0.013	1.388	0.198	0.024	1.249
Return on assets	0.049	0.010	-0.645	0.082	0.010	-0.893

Panel B. Entropy balanced

	Restate=1			Restate=0		
	Mean	Variance	Skewness	Mean	Variance	Skewness
Tangibility	0.149	0.013	1.388	0.149	0.013	1.413
Return on assets	0.049	0.010	-0.645	0.049	0.010	-0.645