Can Long-Term Institutional Owners Improve Market Efficiency in Parsing Complex Legal Disputes?*

Paul Borochin
Miami Herbert Business School
University of Miami†

Xiaoqiong Wang
School of Business
Indiana University Kokomo‡

Siqi Wei
David Nazarian College of Business and Economics
California State University Northridge§

This version: August, 2020

Abstract

Long-horizon institutional investors can help mitigate information asymmetries around securities class action (SCA) lawsuits. We find that the machine readability of SCA complaint filings can predict the outcome and duration of class actions. Long-term institutional investor ownership leads to a more positive post-SCA announcement price reaction and increases the volatility ratio of prices as a measure of price informativeness. Furthermore, there is a significant interaction effect between long-term institutional ownership and SCA complexity on price informativeness consistent with a superior information processing ability about complex corporate events affecting portfolio firms.

Keywords: Institutional ownership; Securities class actions; Market efficiency; Natural language processing
JEL classification: G30, G14, G32, K41

*We are grateful to seminar participants at the University of Mississippi, University of Miami, and CSU Northridge for their helpful comments and suggestions. All errors are our own.
†Miami, FL 33146. Phone: (305) 284-9144. Email: paul.borochin@bus.miami.edu.
‡Kokomo, IN 46902. Phone: (765) 455-9470. Email: xw55@iu.edu.
§Los Angeles, CA 91330. Phone: (818) 677-2410. Email: siqi.wei@csun.edu.

"The plaintiffs' bar continues to file as many cases as possible, upping the odds that one will get past a motion to dismiss. Because the costs associated with surviving a motion to dismiss are not high, these firms are incentivized to continue filing these suits with the hope of higher payouts if they can survive an early-stage dismissal."

- From Nuisance to Menace: The Rising Tide of Securities Class Action Litigations, Chubb Limited, June 2019

1 Introduction

We study the ability of long-horizon institutional investors to mitigate information asymmetries around complex corporate events in the setting of securities class actions (SCAs). This setting is economically significant as the number of SCAs brought against U.S. public firms has increased markedly in the last two decades, and so has the heterogeneity of SCA outcomes. Filings ranged from 83 to 135 per year in the late 1990s, with between 62% to 70% of these being settled or judged in favor of the plaintiff. By the late 2010s, however, the number of SCA filings increased to between 164 to 411 per year, of which only 1.4% to 51% were settled or judged for the plaintiff consistent with increased opportunism by SCA filers (Chubb Limited, 2019). Settlements are costly when they occur, with an average payout of $47 million between 2009 and 2018, adjusted for inflation (Cornerstone Research, 2018). A new SCA complaint against a firm thus introduces a great deal of uncertainty about its potential direct and indirect costs.

There is substantial evidence that not all SCAs are filed in response to managerial malfeasance or negligence, poor corporate governance, or any other justified causes. The passage of the Private Securities Litigation Reform Act (PSLRA) in 1995 has curbed some abuses of the class action system in extracting payouts from firms through meritless lawsuits such as merger objections (Johnson et al., 2000). However, recent work by Kempf and Spalt (2019) finds evidence consistent with the ongoing use of meritless SCAs to "tax" innovative firms in particular, consistent with those firms' high opportunity costs of time and money. This strategic application of complex legal threats to extract settlements from targets is consistent with prior work highlighting the potential for abuse of SCAs by Alexander (1990) and Romano (1991). The lower settlement rate in more recent years reflects the reduced probability of success of each individual SCA as a consequence of their increase overall and a potential increase in meritless claims in particular (Chubb Limited, 2019).

The uncertainty about the ultimate outcome of an SCA complaint, both in terms of the merits of the case and in terms of the costs incurred by the target of the lawsuit should it be induced to settle regardless of the case's merits, creates significant costs to shareholders of targeted firms. Prior work has shown that certain types of institutional owners, particularly ones that hold a narrowly selected portfolio of stocks over a long horizon, can improve the quality of information about these portfolio firms (Borochin and Yang, 2017). In this study, we examine whether long-horizon institutional owners can mitigate the uncertainty arising out of SCA filings by helping market participants separate meritorious lawsuits from meritless ones, particularly in the case of more complex SCAs. To do this, we categorize institutional owners as purely long-term following Gaspar et al. (2005) as those in the lower tercile of investment-weighted portfolio turnover. We also use the Bushee (1998) and Bushee (2001) classification to further isolate a subsample of dedicated long-term institutional owners that have both low turnover and low portfolio diversification.

We measure SCA complexity using machine readability scores of the text of the complaint, following the approach taken in prior work on the information content of financial disclosures (Ganguly et al., 2019; Loughran and McDonald, 2014). This allows us to make several novel contributions to the securities class action and institutional ownership literature. Firstly, we demonstrate that machine readability of SCA filings is an informative measure about the ultimate outcome of the lawsuit: Individual machine readability scores are inversely related to the likelihood of settlement at the 1% statistical significance level, even when controlling for firm and case characteristics. That is, more complex SCA complaints are more likely to be settled, consistent with the Kempf and Spalt (2019) argument that they can be used as a "tax" on firms with a high opportunity cost of time and money. More complex SCA complaints also have a greater span of time between the case filing and resolution dates with statistical significance at the 1% level, suggesting that machine readability correlates with greater real case complexity.

Second, we apply principal factor analysis to SCA filing complexity to reduce the dimensionality of multiple machine readability scores to a single readability factor. We confirm that this readability factor contains the same information about settlement likelihood and case duration as the six individual readability scores, consistent with a common signal. Further, we demonstrate that less readable SCA filings produce a more muted post-announcement price reaction than more readable filings do, consistent with less efficient incorporation their content into prices. To verify that machine readability of SCA filings correlates with greater informativeness rather than simply with more negative market reactions, we also measure its relation to the abnormal return variance ratio as a proxy for new information being impounded into market prices (Chen et al., 2020; DeFond et al., 2007; Landsman and Maydew, 2002; Warner et al., 1988; Beaver, 1968). We find that more readable (less complex) SCA filings produce significantly greater variance ratios than less readable (more complex) ones up to 100 days after the filing date, consistent with greater information content of post-announcement prices. Notably, we demonstrate that this readability factor contains information not explained by mere length of the filing, in contrast to readability analysis of financial disclosures (Loughran and McDonald, 2014).

Third, we find that post-announcement price reactions are more positive for SCA targets with above-median institutional ownership in the quarter prior to the filing. This effect is statistically significant for the Gaspar et al. (2005) long-term institutional owners at all time horizons, but also obtains for the Bushee (1998) dedicated institutional owner type at the shortest ten-day time horizon. The positive relation between post-SCA price reaction and above-median institutional ownership appears to coincide with greater post-event price informativeness as indicated by the abnormal variance ratio, which is higher for all institutional ownership types at the longer-term 100- and 151-day windows at the 1% significance level.

Importantly, we find a significant interaction between the effects of the first principal component of SCA readability scores and institutional ownership on post-announcement price reaction and price informativeness, consistent with the ability of certain institutional owner types to mitigate information asymmetries around complex SCA filings. Greater stakes in the firm by the long-term and focused Bushee (1998) dedicated institutional owner type in the quarter preceding more complex SCA filings results in more positive post-announcement price reaction at all but the shortest post-event time horizon. Furthermore, greater stakes by the long-term Gaspar et al. (2005) institutional owner type in the quarter preceding more complex filings results in more informative prices as measured by the abnormal variance ratio over all but the longest time horizons. This is consistent with the superior ability of dedicated institutional owners in correcting firm misvaluation documented by Borochin and Yang (2017), and with prior work by Guay et al. (2016) that finds that managers may employ voluntary disclosure to mitigate the complexity of financial reports. We find that long-term institutional owners can do the same for complex legal disputes, controlling for firm and case characteristics, as well as the case's legal complexity, in the presence of circuit court, time, and industry effects.

In addition to improved disclosure, improvements in market liquidity offer a potential channel for this relationship: long-term institutional ownership improves liquidity in particular for firms targeted by complex SCA filings, consistent with Boone and White (2015). Higher post-SCA liquidity is consistent with a less negative market impact of SCA-related information, as well as more informative post-SCA prices. Consistent with the view that some SCA litigation is purely opportunistic (Alexander, 1990; Romano, 1991; Kempf and Spalt, 2019; Chubb Limited, 2019), we find that of the 948 law firms in our sample 20 are responsible for a majority of the filings. These firms also have higher settlement rates and lower machine readability of filings. Furthermore, this effect is additive in the number of "frequent filer" firms that join a complaint. Controlling for their effect on market reactions helps explain part, but not all, of the interactive effect between SCA complexity and long-horizon institutional ownership. This result makes sense in light of SCA filing complexity being decided by both the plaintiff's law firm as well as the actual complexity of the case, and is consistent with the strategic targeting of SCA complaints against firms with high opportunity costs of money and time in order to extract settlements (Kempf and Spalt, 2019).

Our study finds both statistical and economic significance of machine readability of the content of court filings for their subsequent outcome, adding novel evidence to a growing text analysis literature that has previously covered the economic impact of news articles (Liu and McConnell, 2013; Tetlock, 2011), financial reports and earnings disclosures (Ganguly et al., 2019; Borochin et al., 2018; Loughran and McDonald, 2014; Doran et al., 2012), and social media (Antweiler and Frank, 2004). By documenting the direct relation between SCA filing textual complexity and the likelihood of settlement in excess of legal characteristics described by McShane et al. (2012), we also add to our understanding of the strategic options available to law firms designing the filings. We thereby contribute to the recent literature on the trade-offs between the legal costs and corporate governance benefits of the SCA framework in the U.S. (Kempf and Spalt, 2019; Cheng et al., 2010).

Furthermore, we add to the literature on the relation between institutional ownership types and firm performance around significant events. On the one hand, Gaspar et al. (2005) find that long-term institutional ownership places firms in a stronger position in takeover negotiations, and Chen et al. (2007) document that only long-term institutional ownership relates to superior post-merger performance. On the other, Yan and Zhang (2007) find that long-horizon institutional trading does not predict future firm performance, although short-horizon trading does. In this study, we find additional support for an information advantage for long-horizon institutions in their superior ability to parse complex information about the legal liabilities of portfolio firms.

The rest of the paper is organized as follows. Section 2 describes the securities class action sample definition and variable construction. Section 3 presents the relation between SCA readability, institutional ownership, and case outcomes and market reactions, as well as and its consistency with prior work on the role of institutions in disseminating information on portfolio firms to the market. Section 4 tests the robustness of the relation to endogeneity concerns in institutional ownership, as well as the relation between SCA complexity and law firms filing the complaint. Section 5 concludes and suggests extensions for future work.

2 Data and Methods

2.1 Litigation sample

We aim to capture a broad sample of securities class action (SCA) lawsuits to maximize variation in case merit, filing complexity, and its interaction with institutional ownership. To identify the firms subject to SCA lawsuits as well as the complexity of the filings brought against them, we collect the data on all filings in federal courts by parsing the Security Class Action Clearinghouse (SCAC).[1] This database tracks all SCA lawsuits filed after the Private Securities Litigation Reform Act (PSLRA) of 1995. Our sample is comprised of filing dates from Jan 1, 1996 to December 31, 2018. Figure 1 presents the distribution of SCA filings by year, along with the numbers of each year's filings that are either settled or dismissed. The remainder are still ongoing.[2] The dramatic increase we observe in filings since 2016 is consistent with the potential vulnerability of the SCA

[1] The SCA filing data are collected from Security Class Action Clearinghouse website: `http://securities.stanford.edu/`.

[2] The SCAC categorizes completed cases only into "dismissed" and "settled" categories, in the first of which the plaintiff's claim is disregarded and in the second some compensation is either agreed upon or awarded.

system for abuse as described by Kempf and Spalt (2019) and Chubb Limited (2019).

The court filings in the SCAC database are not stored in a machine-readable format, and require substantial pre-processing in order to create readability scores. We first programmatically extract all the linked filing information and case details including the class period end date, the initial case filing date, case start and end dates, case status, docket number, district court, and firm ticker. We further extract all associated filing documents, focusing on the first identified complaint (FIC) of each case. We consider the FIC to be the most appropriate filing to capture the primary characteristics of the case because unlike the reference complaint (RC) it is guaranteed to be captured by the SCAC database.

We collect 5,375 SCA case instances in the SCAC from Jan 1, 1996 to December 31, 2018. Of these, we remove 163 cases associated with privately traded firms for which other characteristics are unavailable, and a further 753 firms that are missing an FIC filing. This yields a subset of 4,459 usable complaints after manual verification. Figure 2 provides a flowchart describing the raw litigation data processing.

2.2 Measuring the complexity of litigation filings

The textual data obtained from the SCAC complaint files requires natural language processing (NLP) approaches to quantify its complexity as a model input. We generate complexity measures for each of the 4,459 usable FIC filings using established NLP approaches. We first process the raw text from each FIC by removing all line and paragraph breaks and excessive white spaces to improve the parsing accuracy, and then tokenize the resulting file into individual sentences and words.

We then generate six machine readability scores as proxies for case complexity. These are the Gunning Fog Index (GF), Automated Readability Index (ARI), Smog Index (SI), Flesch Reading Ease Index (FRE), Flesch-Kincaid Grade Level Index (FKG), and the natural log of the length of the litigation document measured in words ($Log(WordCount)$). Detailed definitions of the readability scores are provided in the appended variable definitions. In the extant finance and accounting literature, studies most commonly use the Fog index to capture the readability of disclosures (Li, 2008; Miller, 2010; Bonsall and Miller, 2017). However, Loughran and McDonald

(2014) suggest that the Fog Index is superseded by a simpler readability proxy of the length of the disclosure (specifically, the size of the electronic disclosure file).

To ensure that our results are not specifically driven by one particular readability score, we consider all six measures to alleviate potential measurement error in each individual score. However, due to multicollinearity we are not able to simultaneously deploy all six of the readability scores in the regression tests. Consequently, we adopt a dimension reduction technique on the readability measures, principal component analysis (PCA), to create a composite readability measure. The advantage of the PCA approach is that it captures the maximum variability across our six dimensions of readability while projecting it onto a single-dimensional subspace. By doing so, we remove the correlated features from the readability scores while retaining the uncorrelated ones, and thereby improve the performance of the readability proxy. The higher the $PC1$, the less readable and comprehensive the security class actions filings are.[3] Figure 3 displays the scree plot of the eigenvalues after the PCA of readability scores. The eigenvalue of the first principal component is greater than 4.5, whereas the next-highest is below 1.0. This natural break between the first principal component and other components indicates that the first component $PC1$ contains the most substantial amount of common variation across all readability measures, consistent with a common underlying readability level to which each measure imperfectly relates.

To contrast our novel measure of the textual complexity of SCA filings against a measure of their purely legal complexity, we use keyword matching to create an index based on the SCA complaint topics that have been shown to have a statistical relation with the likelihood of settlement by McShane et al. (2012). These are IPOs, GAAP violations, restatements, SEC rule 10b-5, Section 11 of the Securities Act of 1933, insider trading, and transactional issues related to deals and mergers. This $LegalComplexity$ index is constructed as the total count of mentions of topics extracted from the first identified complaint text document using keyword matching. McShane et al. (2012) also identify the gap between the end of the class action period, during which the alleged harms have occurred, and the complaint filing date as significantly negatively related to SCA settlement. We use the SCAC case characteristic data to create this variable named $FilingTime$.

[3] Higher readability scores indicate higher filing complexity, except for the FRE measure. The higher the FRE, the lower the complexity of the lawsuit case filings.

2.3 Institutional Ownership and Institutional Investor Types

We define institutional investor types using data from the Thomson Reuters Institutional (13F) Holdings database. The Securities and Exchange Commission (SEC) 13F form requires institutional investment managers whose asset under management is over $100 million to report their long positions every quarter. We use the reported long equity positions to create our definitions.

To start with, we classify long-term institutional investors based on their investment horizon measures: the churn rate and portfolio turnover over the past four quarters following Gaspar et al. (2005) and Chen et al. (2007). We first calculate the churn rate CR as the frequency of an investor i rotating the positions of all stocks in the portfolio Q:

$$CR_{i,t} = \frac{\sum_{j\in Q} |N_{j,i,t}P_{j,t} - N_{j,i,t-1}P_{j,t-1} - N_{j,i,t-1}\Delta P_{j,t}|}{\sum_{j\in Q} \frac{N_{j,i,t}P_{j,t} + N_{j,i,t-1}P_{j,t-1}}{2}} \tag{1}$$

where $P_{j,t}$ and $N_{j,i,t}$ represent the price and the number of shares of company j held by institutional investor i at quarter t.

We next calculate the average churn rate over the past four quarters for each institution and sort institutional investors into terciles of average churn, classifying those institutions in the lowest tercile as long-term institutional investors. For each SCA target in our sample, we define the long-term institutional ownership (hereafter LIO) of the target firm as the ratio of the number of shares held by long-term investors to the total number of shares outstanding.

In addition to measuring institutional ownership of SCA targets by types based on their investment horizon only, we also categorize institutional ownership types using both turnover and holdings concentration or diversification following Bushee (1998), Bushee (2001), and Bushee and Noe (2000), isolating the "dedicated" (DED) institutional owners.[4] These have low portfolio turnover and concentrated portfolio holdings, and have been shown to possess an informational advantage (Borochin and Yang, 2017). For each of the SCA target firms we define its "dedicated" institutional ownership level (hereafter $DEDIO$) as the ratio of the number of shares held by "dedicated" investors to the total number of shares outstanding.

[4]This classification is obtained from http://acct.wharton.upenn.edu/faculty/bushee/IIclass.html. We thank Brian Bushee for providing the data on the website. We use the permanent classification scheme.

2.4 Summary Statistics

Table 1 presents the summary statistics for the regression sample and full COMPUSTAT and SCAC samples for the period of 1996 to 2018. Panel A presents the comparison of firm characteristics of the regression sample and COMPUSTAT sample. The average firm in our sample has a 15.1% ownership level by long-term LIO institutional types, and a 2.0% level by dedicated $DEDIO$ types. The overall institutional ownership level IO is 57.9%. These institutional ownership measures have significant variation to conduct statistical tests: LIO ranges from 0% to 50.0%, while $DEDIO$ ranges from 0% to 37.3%. The average annual return of firms associated with SCA lawsuits is -4.3% indicating a meaningful loss of value to a firm even before taking the merits of the case into account. This further underscores the variability of ex-post SCA outcomes, and the potential value that institutional investors can create by providing greater clarity about these outcomes. We further show that the ownership, size, B/M, and volume characteristics of our regression sample are similar to those of the COMPUSTAT universe and fall within one standard deviation of each other.

Panel B presents the summary statistics of the SCA filing characteristics of our sample. The average length of time, $CaseDuration$, for an SCA case to be resolved after the complaint is filed is 932 days. Thus, unlike quarterly earnings announcements and short-term corporate events like M&A, SCA events can affect a firm for over an average of 2.5 years. SCA filings can be considered challenging for investors to understand: their mean Fog Index GF is 15.822, which is classified as "difficult" according to the index's interpretation (Li, 2008). The standard deviation of GF is 1.937, which provides substantial variation and is consistent with (Li, 2008).[5] The other machine readability measures provide similar variation, with the average (standard deviation) for ARI of 20.452 (2.444), SI of 17.088 (1.236), FRE of 20.611 (6.667), and those for FKG of 18.870 (2.008). The average number of words in an SCA filing is 8930.913, ranging from 2957 to 28401 with a standard deviation of 4649.759. The SCAC database contains 3,560 downloadable complaints, 2,865 of which make it into our regression analyses due to the availability of ownership and firm data. Comparing the SCAC universe with the subsample used in the study shows that the readability measures across both samples are highly similar and within one standard deviation of each other.

[5] Li (2008) write that "the difference in the Fog index between Reader's Digest and TIME magazine is about 2."

Table 2 presents the Pearson correlations of the sample variables. The correlation between LIO and $DEDIO$ is 31.96%, suggesting that while there is some commonality, dedicated and long-term institutional investors do not necessarily hold the same firms. This providing justification for considering differential effects of both long-horizon institutional ownership types on market reactions to SCA announcements.

Furthermore, case duration is significantly negatively associated with LIO and $DEDIO$, providing preliminary univariate evidence that institutional investors can mitigate information asymmetry around SCA filings and lead to a speedier resolution. Reassuringly, machine readability measures have significant, though not perfect, correlation with each other. At the high end of magnitudes, ARI and FKG have a correlation of 96.91%, while at the low end FRE and $Log(WordCount)$ have one of -13.08%.[6] These correlations suggest that all six measures are each related to aspects of a common idea of readability, which we more formally demonstrate using PCA. Overall, larger firms tend to have more readable SCA filings, as evidenced by the negative correlation coefficients between firm size and readability measures that are decreasing in readability, and a positive one with FRE which is increasing in readability. Firms headquartered in Delaware, and therefore sued in the Delaware Court of Chancery system, also have more readable SCA filings.

3 Empirical findings

3.1 Case status and case duration

We first examine the relation between machine readability of SCA complaints and the outcome of the lawsuit as well as its duration. This motivates the economic significance of machine readability scores as proxies for the complexity of the lawsuit. We use our six machine readability scores as proxies for SCA complexity: ARI, SI, FRE, FKG, GF, and $Log(WordCount)$, as well as their first principal component. Figure 4 presents the distribution of case status of the lawsuit cases in our sample. We observe that dismissal and settlement are almost equally distributed, with only a small number of outstanding cases. This illustrates the ex-ante uncertainty about outcomes, as well as the potential perverse incentives for SCA filers to "tax" target firms (Kempf and Spalt,

[6] Unlike all other measures, the FRE index is increasing in the ease of readability and is thus negatively correlated with the other five measures.

2019; Chubb Limited, 2019) since the odds of obtaining a settlement are high.

We first fit a model for case outcomes as a function of machine readability measures to measure the information content of these proxies regarding the complexity and expected outcome of the case. We construct a dummy variable, *Settled*, which equals one if the security class action lawsuit case status is settled and zero otherwise. For each SCA filing i affecting firm f, we fit a probit model for the *Settled* indicator variable on individual readability scores as well as the overall SCA complexity measure constructed from the principal component analysis, $PC1$ in the model below:[7]

$$Settled_i = \Phi(\alpha_i + \beta MR_i + \gamma Controls_{f,i} + fe_c + fe_t + fe_j + \varepsilon_i) \quad (2)$$

Here MR_i represents each of the case-specific machine readability measures for the SCA filing, and $Controls_{f,i}$ a vector of case- and firm-specific characteristics. Following Tetlock et al. (2008) and Loughran and McDonald (2011), we control for the natural logarithm of the stock price $Price$, the natural logarithm of firm size *Size*, and the book to market ratio BM. We also control for the return volatility over the preceding year Vol as Rogers et al. (2011) document that volatility is correlated with both textual tone and the litigation risk. We further control for the past year's stock return *Ret* as SCA complaints are often brought as a result of a stock price drop (Ganguly et al., 2019; Chubb Limited, 2019) and total institutional ownership IO from the preceding quarter following Bird and Karolyi (2016). To separate the textual complexity of the SCA filings from legal complexity surrounding the case, we also control for the gap between the end of the class period during which the alleged injury occurred and the filing date of the SCA complaint, $FilingTime$, and the index of SCA topics associated with settlement, *LegalComplexity*, based on the findings of McShane et al. (2012). We include a *Delaware* incorporation indicator variable as that state has distinct and up-to-date corporate laws and a specialized Court of Chancery for speedy trials of corporate matters, making it a unique venue for SCA litigation (Li, 2008). Finally, we include circuit court fe_c, year-quarter fe_t, and industry fe_j fixed effects to control for variation in case outcomes by court, time, and firm type.[8] All variables are defined in Appendix A.

[7]The lawsuit and firm indicators are non-redundant since is possible that the same firm can be targeted by more than one SCA lawsuit.

[8]Ganguly et al. (2019) demonstrate that changes at the circuit court level can have a meaningful impact on the prospect of SCA lawsuits.

As we observe in Table 2, the six readability measures are highly correlated. Thus, we include each measure sequentially in probit models of SCA outcomes to avoid multicollinearity issues in predictive variables, as well as separate specification with the first principal component of the six individual measures $PC1$. Table 3 presents two sets of results: without and with control variables in columns 1-7 and columns 8-14, respectively. Higher values of the readability scores indicate that case filings are more difficult to read except for FRE. As shown in column 1, the coefficient of ARI is 0.049, with a t-value of 4.14. When we add control variables to the regression, the coefficient of ARI reduces to 0.047, which is significant at the 1% level. In unreported results, we calculate the average marginal effects of readability scores on the settlement likelihood. Our findings are economically significant: a one-standard-deviation increase in ARI leads to the probability of a case being settled to increase by 11.24%. These results are consistent across the rest of the readability measures. This evidence suggests that our six machine readability measures are picking up economically significant information about the complexity of SCA filings, and that SCA lawsuits with more complex and less readable complaints are more likely to be settled as opposed to being dismissed. Our evidence is consistent with the Kempf and Spalt (2019) argument that they can be used as a "tax" on firms with a high opportunity cost of time and money.

We further observe that firm size reduces the likelihood of SCA settlement. We also document a positive association between the target firm's stock price and the probability of settlement. The coefficient on $Delaware$ is negative and borderline significant, weakly consistent with its reputation for efficiently processing corporate litigation and dismissing meritless claims.

We next turn to the relation between case duration and machine readability of SCA filings. Our dependent variable is $Duration$, the logarithm of the length of time in days between the SCA filing and resolution date as obtained from the SCAC. For each SCA filing i affecting firm f, we regress case duration on each of our machine readability measures MR_i in the presence of the same set of firm-specific control variables $Controls_{f,i}$ as those used in Eq.(2), also with circuit court, time, and industry fixed effects:

$$Duration_i = \alpha_i + \beta MR_i + \gamma Controls_{f,i} + fe_c + fe_t + fe_j + \varepsilon_i \tag{3}$$

The results are reported in Table 4. As shown in column 8, the coefficient of ARI is 0.025 in the presence of control variables, and is significant at the 1% level. The evidence suggests that as the complexity of the SCA complaints increases, the length of time it takes for a case to be resolved increases accordingly. The results are economically significant: a one-standard-deviation increase in ARI contributes to an increase in case duration by 6.1%. We document a consistently positive association between machine readability scores and SCA lawsuit duration across all other measures. Less readable filings are more difficult to process and contain less information, which may result in a longer case period. We also observe a similar effect of $PC1$ on case duration. As is shown in column 14, the partial effect of our case complexity measure, $PC1$, is 0.027, which is statistically significant at the 5% level. The coefficients on $Log(LegalComplexity)$ are negative and statistically significant, indicating that multi-topic SCA complaints get resolved more quickly. Those on $Log(FilingTime)$ are positive and significant, consistent with a longer delay between the alleged damages and filing date resulting in a longer trial. Notably, there exists a significantly negative relation between institutional ownership IO and case duration. Since case duration is left-censored at zero, we check the econometric validity of an OLS approach to case duration with a Tobit regression in Table A.1. Our results hold, suggesting that they are not driven by econometric errors.

Our results suggest that the machine readability analysis of SCA filings can be used to generate informative measures with regard to the ultimate outcome and duration of SCA lawsuits controlling for the characteristics of the lawsuit and the target firm itself. However, it is important to consider whether these machine readability scores provide any information about the complexity of the legal filing beyond its overall volume or length. In prior work, Loughran and McDonald (2014) suggest that the file size of financial disclosures provides as much information on their complexity as the Fog Index, consistent with the interpretation that readability measures simply proxy for size. Furthermore, Ganguly et al. (2019) show that the word count in a firm's financial disclosures increases the odds of becoming a target for SCA. To address the concern that more complex machine readability scores may not contain additional information about SCA filing complexity in excess of the filing's length, we re-estimate the models relating ARI, SI, FRE, FKG, GF, and $PC1$ to SCA case outcome and duration while adding $Log(WordCount)$ as an additional control variable. These robustness results are reported in Appendix Tables A.2 and A.3. We show that including $Log(WordCount)$ in the regression models does not alter the significance of our other machine readability measures of the SCA filings. Thus, unlike prior work on readability in financial disclosures (Loughran and McDonald, 2014; Ganguly et al., 2019), our results suggest that document length alone does not capture the information contained in the machine readability of SCA documents, and that the structure of the text also matters for the lawsuit's outcomes.

3.2 Two sample t-test of price reaction on machine readability and institutional ownership

Having established the economic significance of machine readability for SCA outcomes, we now examine market responses to the lawsuit filings and the role of financial institutions as information intermediaries around these events.

Prior studies of market reaction to lawsuit filings, such as Karpoff and Lott Jr (1993), Bizjak and Coles (1995), Bhagat et al. (1998), Gande and Lewis (2009), Klock (2015), and Ganguly et al. (2019) document significantly negative stock price reaction around lawsuit filing dates. These negative stock price reactions are composed of several short-term and long-term influences on the target firm: direct legal costs, expected settlement, reputational costs, an increase in the probability of financial distress, and a higher chance of subsequent lawsuits (Ganguly et al., 2019; Kempf and Spalt, 2019). Given that the average SCA lawsuit in our sample lasts over 2.5 years, focusing on a short event window around the filing dates alone will likely understate the true impact of the lawsuits on the firm (Gande and Lewis, 2009). Indeed, Karpoff et al. (2017) argue that the filing date is only one event in a long process that begins when a firm becomes involved in a legal dispute, with additional news and information revealed gradually afterward. Thus, we utilize a range of event windows to capture the evolving information content of the class action lawsuits and the market reaction to this information over both short-term and long-term event windows measured in days as follows: a ten-day window following the SCA filing (2, 11), two-month (2, 61), three-month (2,101), and five-month (2, 151) windows.

We construct two variables to measure the price reaction of the lawsuits: a post-announcement price abnormal price reaction $Drift$ and an abnormal variance ratio VR. Specifically, $Drift$ is defined as the cumulative return in excess of the value-weighted market return over the event window of (2, N) where N takes on the values of 11, 61, 101, and 151. $Drift$ directly measures and captures the shareholder value changes after the filing dates. As the lawsuit continues, positive or negative news and information emerge, which eventually influences shareholder value. The post-announcement abnormal variance ratio VR is defined as the variance of the stock return in excess of the value-weighted market return over the event window of (2, N). We scale the post-event variance by the historical variance of the stock return in excess of the value-weighted market return over the estimation window of (-120, -21). Proposed by Lo and MacKinlay (1988), a higher VR implies a higher information content in post-event prices as their variability exceeds its historical level (Beaver, 1968; Bamber et al., 2000; Warner et al., 1988; Landsman and Maydew, 2002; DeFond et al., 2007; Chen et al., 2020).

We perform two-sample t-tests on $Drift$ and VR across subsamples stratified on levels of readability and institutional ownership. We first form two subsamples around the median of the first principal component of the six readability measures $PC1$, which includes ARI, SI, FRE, FKG, GF, and $Log(WordCount)$. This dimensionality reduction approach combines information from all readability scores to create simple more- and less-readable subsamples. The results of the post-announcement $Drift$ comparison between the more readable and less readable SCA filing subsamples are presented in the first panel of Table 5. Firms with more readable SCA filings experience short-term and long-term negative $Drift$ after the filing dates ranging from -0.017 to -0.005. On the other hand, $Drift$ for the less readable subsample is negative immediately after the filing date with a coefficient of -0.010 but becomes positive over time to 0.009. Overall, the $Drift$ of the more readable subsample is more negative than that of the less readable subsample. These findings suggest that less readable SCA filings produce a more muted short-term reaction in the markets than the more readable ones do.

The two-sample t-tests on VR in the first panel of Table 5 confirm that our findings are likely to be driven by information quality rather than simply by a more negative reaction to more readable SCA filings. We show that the VR is significantly higher for the more readable filing subsample, both in the short term as well as the long term. Since prior literature has shown that higher VR proxies for greater information content of prices, this suggests that the more negative $Drift$ observed for more readable SCA filings is due to a greater ease in parsing their information content.

Prior literature suggests that institutional owner types can have differential effects on firm valuation due to an informational advantage. To study the role of institutional investor types in the process of impounding information contained in more or less complex SCA complaints into market prices, we form subsamples based on above- and below-median levels of LIO and $DEDIO$. The two-sample t-tests on $Drift$ and VR across these subsamples are displayed in subsequent panels of Table 5. Long-term, focused institutional owners have comparative advantages in effectively monitoring managers since they can spread both costs and benefits associated with monitoring over a long time span (Gaspar et al., 2005; McCahery et al., 2016; Harford et al., 2018). As a result, they are motivated to collect value-enhancing information and trade on the growth potential of a firm (Huang and Petkevich, 2016; Borochin and Yang, 2017), suggesting that they should improve the informativeness of prices post-SCA.

Our univariate results suggest that institutional ownership does play a meaningful role in processing the implications of SCA filings. In the second panel of Table 5 we observe that above-median long-term institutional investors LIO result in positive $Drift$ at all horizons from 11 to 151 post-event days, whereas we observe persistently negative $Drift$ for the below-median LIO subsample over the same horizons. The differences between these two subsamples are statistically significant. The price informativeness proxied by VR is significantly higher for the above-median LIO group relative to the below-median one, especially over the long-term event window.

The test of differences between above- and below-median levels of $DEDIO$ ownership in the third panel of Table 5 are consistent with the findings above. Above-median dedicated institutional ownership is related to less negative post-announcement drift with statistical significance across three out of the four windows. Simialrly, it is related to a higher VR in two out of four windows. Taken together, these results suggest that long-term institutional owners, and dedicated owners as a subsample of them, can mitigate information asymmetry around SCA filings as the positive price reaction is accompanied by greater price informativeness in cases when the targeted firm has above-median levels of long-horizon ownership.

To ensure the robustness of our findings, we further replicate our analyses in Table 5 with the equal-weighted market returns rather than the value-weighted market returns. The results are reported in Appendix Table A.4. Using equal-weighted market returns to construct the price reaction measures provides consistent evidence.

Overall, our findings suggest that the machine readability of court filings, as well as institutional ownership types, each have a significant influence on the short-term and long-term price reaction and price informativeness after the filing of SCA complaints in univariate analysis. Our next step is to test the interaction effect between these two variables in the presence of controls in a multiple regression setting.

3.3 Price reaction, machine readability, and institutional ownership

We now turn to a multiple regression model to control for potential confounders in the univariate analysis above, and to test the interaction between SCA filing complexity for all filings i and institutional ownership of target firms f on the post-announcement price reaction. We use the following model specification for the market reaction after case filing i targeting firm f:

$$\begin{aligned} Drift_{i,(2,N)} = \alpha_i + \beta_1 IO_{f,t-1} + \beta_2 IO_{f,t-1} \times PC1_{i,t} + \beta_3 PC1_{i,t} \\ + \gamma Controls_{f,i} + fe_c + fe_t + fe_j + \varepsilon_i \end{aligned} \tag{4}$$

Here IO represents different ownership types specified earlier. $PC1$ denotes the first principal component of six established readability scores. The dependent variable is $Drift$, the cumulative return over the (2,N) day window after the SCA filing with N taking on the values of 11, 61, 101, and 151 days. The firm- and case-specific characteristics $Controls_{f,i}$ consists of the variables used in Eq. (2): overall institutional ownership IO, the natural logarithm of the stock price $Price$, the book to market equity ratio BM, and the stock return and volatility over the prior year, Ret, Vol, $Delaware$ dummy, the time gap between the end of the class period and the SCA filing $Log(FilingTime)$, the count of lawsuit topics identified as meaningful by McShane et al. (2012), $Log(LegalComplexity)$. We also augment the control set with a measure of the immediate price reaction during the (-1,1) day window around the SCA filing, $EventDayRet$, to clarify whether the long-term drift is due to a delayed response or a larger overall response. As before, we include circuit court fe_c, year-quarter fe_t, and industry fe_j fixed effects, and cluster the standard errors at the year-quarter level.

In Panels B, C, and D of Table 6, corresponding to event windows of (2, 61), (2,101), and (2,151) days respectively, we find that dedicated institutional ownership ($DEDIO$) significantly improves the post announcement abnormal price reaction for more complex SCA lawsuits, as evidenced by the positive coefficients on the interaction term, $DEDIO_{t-1} * PC1_t$. Specifically, the coefficients on the interaction term are 0.123, 0.144, and 0.185 for the three longer-term event windows respectively. The positive interaction coefficient suggests that the effect of dedicated ownership on post-announcement price reaction is stronger as the degree of the SCA filing complexity becomes greater. The findings further imply that as the event window increases, the influences of dedicated investors in improving the post-announcement drift increase accordingly. To test the robustness of this finding we use an equal-weighted market return over the event windows to construct the post-SCA announcement abnormal price reaction measures instead. The results of this alternative specification in Table A.5 support the relation between dedicated institutional ownership and market efficiency for less-readable SCA case filings. As these case filings are more likely to be difficult to comprehend, especially for less-sophisticated investors, these results are consistent with the role of long-term, dedicated investors in improving market efficiency around complex corporate events.

We also implement the same analytical framework in Eq.(4) for the abnormal variance ratio VR as a proxy for the informativeness of prices after the SCA announcement i relative to a historical level for each targeted firm f, as proposed by Lo and MacKinlay (1988):

$$\begin{aligned} VR_{i,(2,N)} = \alpha + \beta_1 IO_{f,t-1} + \beta_2 IO_{f,t-1} \times PC1_{i,t} + \beta_3 PC1_{i,t} \\ + \gamma Controls_{f,i} + fe_c + fe_t + fe_j + \varepsilon_i \end{aligned} \tag{5}$$

Here VR is defined as the variance of stock returns in excess of the value-weighted market return over the same event windows specified above, normalized by the historical variance of excess returns as described in Appendix A.

As shown in Panels A-D of Table 7, the coefficients of the interaction term, $LIO_{t-1} * PC1_t$, are positive and statistically significant across three of the four time windows. Specifically, the coefficients on the interaction term for long-term institutional ownership with case complexity are 0.393 for the (2, 11) event window in Panel A, significant at the 10% level, 0.340 for the (2,

61) event window in Panel B, significant at the 5% level, 0.396 for the (2, 101) event window in Panel C, significant at the 1% level, and statistically insignificant for the (2, 151) event window in Panel D. These findings suggest that long-term institutions can mitigate information uncertainty around complex SCA filings and increase the informativeness of stock prices over both short-term and long-term horizons. In an alternative specification we use equal-weighted market returns to construct the variance ratio to test the robustness of our findings to weighting schemes. These results in Table A.6 document a consistent statistical significance of the interaction of long-term institutional ownership with case complexity, further supporting our findings on the role of long-term institutional ownership in processing information around complex corporate events.

Figures 5 and 6 provide additional detail about the interaction between institutional ownership types and SCA complexity across all event windows. They present the marginal effects of dedicated and long-term institutional ownership on the post-SCA announcement drift and variance ratio respectively, $\beta_1 + \beta_2 \times PC1_{i,t}$, demonstrating that the coefficients on the interaction terms with dedicated ownership for are positive and significant. Furthermore, the overall effect of dedicated and long-term ownership is shown to be an increasing function of the case complexity, PC1. The results indicate that the influences of dedicated (long-term) institutional ownership on $Drift$ (price informativeness) are more pronounced when SCA filings are more complex, which supports the findings we document in Table 6 and Table 7.

3.4 Liquidity, machine readability, and institutional ownership

We next consider potential channels for the ability of long-term and dedicated institutions to improve information quality around complex SCAs. Prior work by Boone and White (2015) suggest several channels by which institutions improve the quality of information on their portfolio firms: managerial disclosure through 8-K reports, analyst following, and market liquidity. Of these, we expect market liquidity to be the most responsive to SCA announcements.

Our primary measure of liquidity is $RQSPREAD$, the relative quoted spread defined as the daily difference between bid and ask prices normalized by their midpoint. We average $RQSPREAD$ over the $(2, N)$ day window after each SCA filing i targeting firm f, where N takes the values of 11, 61, 101, and 151, and fit the model below:

$$RQSPREAD_{i,(2,N)} = \alpha + \beta_1 IO_{f,t-1} + \beta_2 IO_{f,t-1} \times PC1_{i,t} + \beta_3 PC1_{i,t} + \gamma Controls_{f,i} + fe_c + fe_t + fe_j + \varepsilon_i \quad (6)$$

Our control variable set $Controls_{f,i}$ is the same as that in the case duration model in Eq. (3), and we again include circuit court, year-quarter, and industry fixed effects.[9] We present the results in Table 8. We expect information asymmetry reduction by institutions in complex SCA cases to result in a decrease of the relative quoted spread. Consistent with this conjecture and our prior evidence on market reaction to complex SCA announcements with long-term and dedicated institutional ownership, we observe that both the $LIO_{t-1} * PC1_t$ and $DEDIO_{t-1} * PC1_t$ terms have negative coefficients for $RQSPREAD$ at the (2,11) and (2,61) day windows in Panels A and B, with significance ranging between 10% and 5% levels. The results are much weaker for the long-horizon (2,101) and (2,151) day periods in Panels C and D, but the $LIO_{t-1} * PC1_t$ is also negative and statistically significant at the 10% level for the 150-day window. These results suggest that this reduction in relative quoted spread for long-term and dedicated institutional ownership contributes to the market reaction in price trend and informativeness that we observed in Tables 6 and 7.

4 Robustness

In this section we test the robustness of our findings by controlling for potential endogeneity of institutional ownership in unobserved firm characteristics. We also consider the strategic role of SCA filing complexity as a choice made by the law firms filing the initial complaints.

4.1 Instrumental variable regression

Our analysis thus far indicates a positive relation between dedicated and long-term institutional ownership and post-SCA drift and variance ratio in the setting of complex class action lawsuits. However, the potential for endogeneity of institutional ownership with respect to future SCA law-

[9] Since this is not a test of a price reaction, we omit the *EventDayRet* variable in Eq. (4) and (5). Including it does not affect the results.

suits or market outcomes precludes a causal interpretation of this relation. While it may be argued that security class action lawsuits are relatively unanticipated events, and therefore less likely to drive reverse causality between post-SCA return performance and ownership (Bird et al., 2015), past institutional ownership could still be endogenous in future SCA filings or firm performance due to an omitted firm characteristic related to governance, valuation, or risk (see, e.g., Borochin and Yang, 2017). To mitigate this concern, we instrument for the firm-specific level of institutional ownership by type using the average level of institutional ownership of the firm's SIC2 industry peers, excluding the firm itself.[10] Since each SCA filing i is unlikely to affect the peers of the targeted firm f, we expect industry-wide levels of institutional ownership by type to be an exogenous instrument. We implement a two-stage regression estimation below:

First Stage:

$$IO_{f,t} = PeerIO_{f,t} + \gamma Controls_{f,i} + fe_t + fe_j + \varepsilon_i \tag{7}$$

Second Stage:

$$\begin{aligned} Drift_{i,(2,N)} = \alpha + \beta_1 \widehat{IO_{f,t-1}} + \beta_2 \widehat{IO_{f,t-1}} \times PC1_{i,t} + \beta_3 PC1_{i,t} \\ + \gamma Controls_{f,i} + fe_c + fe_t + fe_j + \varepsilon_i \end{aligned} \tag{8}$$

$$\begin{aligned} VR_{i,(2,N)} = \alpha + \beta_1 \widehat{IO_{f,t-1}} + \beta_2 \widehat{IO_{f,t-1}} \times PC1_{i,t} + \beta_3 PC1_{i,t} \\ + \gamma Controls_{f,i} + fe_c + fe_t + fe_j + \varepsilon_i \end{aligned} \tag{9}$$

where $Peer\, IO_{f,t}$ is the average level of institutional ownership of firm f's peers in the same SIC2 industry in quarter t. The second stage uses the predicted institutional ownership values as independent variables in the same estimation specification defined in our previous section.[11] We include the same set of control variables $Controls_{f,i}$ as in Eq. (4) and (5), and year-quarter and industry fixed effects in the first stage, and circuit court, year-quarter, and industry fixed effects in the second stage.

Our variables of interest are the interaction terms, $\widehat{IO_{f,t-1}} \times PC1_{i,t}$. Table 9 and Table 10

[10]We alternatively define the firm's industry peers using the more granular SIC4 classification, and find out that our results are consistent.

[11]For econometric accuracy, both stages are estimated jointly even though they are presented separately.

present the result from our second stage estimation for post-SCA drift and post-SCA variance ratio, respectively. Table 9 shows that dedicated institutional ownership positively contributes to the post-SCA price reaction for more complex lawsuits, as evidenced by the 5% level of statistical significance for the coefficients on the interaction terms $\widehat{DEDIO_{f,t-1}} \times PC1_{i,t}$ over the 60-, 100-, and 150-day long-term event windows, with estimates of 0.129, 0.160, and 0.197 respectively. Turning to the post-SCA variance ratio, we document similar results. Table 10 shows that long-term institutional ownership significantly improves the abnormal return variance ratio in more complex SCA filings across all event windows. The interaction term $\widehat{LIO_{f,t-1}} \times PC1_{i,t}$ is statistically significant at the 5% level for the 10- and 60-day window, at the 1% level for the 100-day window, and at the 10% level for the 15-day window. The coefficient estimates for the interaction across the four time windows are 0.381, 0.379, 0.425, and 0.245 respectively. These instrumented values of $\widehat{IO_{f,t-1}}$ are unlikely to be endogenous in the likelihood of SCA filing, and support a causal interpretation of the relation between dedicated and long-term institutional ownership and market reactions for complex securities class action complaints.

4.2 Frequent-Filer Law Firms, Case Complexity, and Outcomes

The complexity of the SCA complaint is determined by the plaintiff's law firm as well as the particulars of the case at hand. Therefore, our assertion that SCA complexity is economically meaningful relates to the Kempf and Spalt (2019) hypothesis that law firms can strategically target SCA claims at firms with high opportunity costs of time and resources. If both are correct, then we should expect two things: first, a majority of the lawsuits in our sample should be attributable to a relatively small subset of law firms that specialize in class actions, and second, these firms should explain a meaningful portion of SCA complexity, case outcomes, and market reactions. Law firm strategy thus provides a potential channel for the complexity of SCA filings. In this section, we test the importance of this channel by including fixed effects for the law firms responsible for the majority of the cases in our sample.

Similar to other case details, we obtain the name of the plaintiff's law firms from the First Identified Complaint section of each filing in the Stanford Class Action Clearing House database. Many cases are filed by multiple law firms, all of which we consider associated with each case they are mentioned in as a plantiff's firm. Since the same firm may exhibit slight variation in its title or office locations, we create fuzzy match groupings of similar law firm names with manual verification. Table 11 presents the ranking of law firms sorted by the frequency of their SCA filings, and tabulates their individual and cumulative fraction of the total filing amount. The distribution of filings across law firms validates our first expectation: of the 942 unique plantiff's firms identified in our sample, a small subset of 20 are responsible for the majority of the SCA cases brought. The most prolific of these is associated with 6.8% of the filings in the entire sample, and the frequency diminishes exponentially: the 20^{th} most frequent-filing law firm of 942 is only responsible for an additional 1.16% of the total filings. Consistent with this, the average of SCA filings per law firm is 17.49, but the median and mode filings per law firm are both 1.

Next, we test the extent to which these frequent-filer law firms explain the complexity of SCA filings and the market's reaction to this complexity. We begin by regressing SCA outcome *Settled*, as well as our six SCA filing readability measures, ARI, SI, FRE, FKG, GF, $Log(WordCount)$, and their first principal component $PC1$, on the log count of the top 20 most frequent SCA filers from Table 11 involved in the case, $Log(1 + \#Top20)$. We present the results with circuit and industry fixed effects,[12] both on their own and with the same $Controls_{f,i}$ specification as in the baseline SCA settlement model in Eq. (2), in Panels A and B of Table 12. In both cases, we observe that the number of frequent-filer law firms involved in the case is positively related to the likelihood of settlement with statistical significance at the 1% level. Furthermore, four of the five SCA filing readability measures that are positively related to filing complexity, ARI, SI, FKG, and GF, as well as the overall principal complexity factor $PC1$ are also positively related to the number of frequent-filers involved in the case with statistical significance at the 10% to 5% level. The FRE measure which is negatively related to filing complexity is also negatively related to the number of frequent-filers on the case at the 1% level. Notably, this does not appear to be simply due to longer filings as the $Log(WordCount)$ of the SCA filing is not statistically related to the involvement of these frequent-filer law firms in the lawsuit.

Having established that the top 20 frequent-filer law firms responsible for the majority of the SCA filings in our sample are related to both higher settlement rates and higher SCA filing complex-

[12]We omit year-quarter fixed effects in these specifications since they absorb all variation in the relatively sparse indicator variables for the participation by the top 20 frequent filer law firms in any particular SCA filing.

ity, we next consider their effect on the market's reaction to their filings. We do this by repeating the models for price reaction and variance ratio on institutional ownership interactions with SCA complexity in Equations (4) and (5) in Tables 6 and 7, now with the inclusion of the number of frequent-filer law firms responsible for the majority of SCA filings from Table 11, $Log(1+\#Top20)$, involved in the filing. We present the results for dedicated and long-term institutional owners in Table 13.

This set of results validates our second expectation: not only do frequent-filer law firms file more complex SCA complaints and win more settlements, but the coefficients on the interaction of $DEDIO$ and $PC1$ on the post-SCA drift in Panel A of Table 13 are largely similar to, but slightly lower than, those in Table 6. The coefficients on the interaction of LIO and $PC1$ on the post-SCA variance ratio in Panel B of Table 13 also present somewhat weaker versions of those in Table 7.

The evidence on settlement rates and SCA filing complexity by frequent-filer law firms in Table 12 suggests that law firms do seem to strategically manage filing complexity. Furthermore, the reduction in the significance of the market reaction results in Panels A and B of Table 13 due to frequent-filer law firm dummies relative to the baseline results in Tables 6 and 7 confirms that this relatively small subset of law firms plays an important role in determining the complexity of SCA filings. Accounting for these frequent SCA filers explains some, but not all, of the interactive effect of SCA filing complexity with institutional ownership in determining the market's reaction to the uncertainty introduced by the lawsuits.

5 Conclusion

Security class action (SCA) litigation is an important part of corporate governance (Cheng et al., 2010), but is also open to potential abuses by law firms that file SCA complaints to win fees for themselves rather than compensation for class members (Alexander, 1990; Romano, 1991). Despite the passage of the Private Securities Litigation Reform Act of 1995 as an attempt to limit these abuses, recent analyses by litigation consultants (Cornerstone Research, 2018), insurers (Chubb Limited, 2019), and academics (Kempf and Spalt, 2019) point to increasing costs for firms targeted by an increasing number of SCA filings. These costs primarily come through settlements of claims that manage to survive a motion to dismiss. This uncertainty about SCA outcomes creates substantial losses to targeted firms at the time the complaint is filed, leading to about a 4.3% negative return on average. We find evidence confirming the ability of long-term institutional investors to help mitigate this uncertainty, consistent with Borochin and Yang (2017).

We show that machine readability of SCA complaints can be used to create informative measures of the complexity of the legal cases which can predict their outcome and duration. Furthermore, these readability measures contain information not subsumed by the overall length of the complaint, unlike related findings regarding the readability of financial disclosures (Loughran and McDonald, 2014). The predictive information from the machine readability of SCA complaints for case settlement, duration, and market reaction is not due to case characteristics previously shown to predict SCA outcomes by McShane et al. (2012). We also use a principal components dimensionality reduction approach to create a single measure of SCA readability from our initial six measures.

Ownership level in the quarter prior to the SCA filing by long-term institutional types defined following Gaspar et al. (2005), as well as a related sub-population of dedicated institutional owners defined following Bushee (1998), reduces the negative post-filing price reaction and increases the informativeness of prices prior to the filing relative to historical levels as measured by the abnormal variance ratio proposed by Lo and MacKinlay (1988).

Importantly, we find a significant interaction between case complexity as measured by machine readability and long-term institutional ownership, both for the post-announcement price reaction and for the abnormal variance ratio. These interactions have meaningful economic effects, and demonstrate the ability of long-term institutional investors to mitigate information uncertainty around complex legal disputes consistent with the informational advantage about firm valuation in general as documented by Borochin and Yang (2017).

We find a potential channel for this relationship in the evidence that long-term and dedicated institutional ownership improves market liquidity for firms targeted by complex SCA filings. We also find evidence consistent with strategic use of SCA complexity by law firms specialized in SCA filings, consistent with the Kempf and Spalt (2019) hypothesis of opportunistic SCA litigation: of the 948 law firms in our sample, 20 are responsible for a majority of the filings. These "frequent-filer" firms have an additive effect, increasing with their number involved in a particular complaint, both on the likelihood of settlement versus dismissal as well as on the complexity of the SCA filing. Controlling for their effect on market reactions helps explain part, but not all, of the interactive effect between SCA complexity and institutional ownership.

In future work, it seems worthwhile to investigate the informational advantage of long-horizon institutional investors, and their ability to reduce information asymmetries, around other complex corporate events such as acquisitions, privatization transactions, distress scenarios, and proxy fights. Furthermore, the distinct importance of machine readability for SCA filings relative to existing applications to financial disclosure, news articles, and social media, suggests avenues for further research into natural language processing of other types of materially informative text in business settings.

References

Alexander, J. C. (1990). Do the merits matter: A study of settlements in securities class actions. *Stan. L. Rev.*, 43:497.

Antweiler, W. and Frank, M. Z. (2004). Is all that talk just noise? the information content of internet stock message boards. *The Journal of Finance*, 59(3):1259–1294.

Bamber, L. S., Christensen, T. E., and Gaver, K. M. (2000). Do we really knowwhat we think we know? a case study of seminal research and its subsequent overgeneralization. *Accounting, organizations and society*, 25(2):103–129.

Beaver, W. H. (1968). The information content of annual earnings announcements. *Journal of accounting research*, pages 67–92.

Bhagat, S., Bizjak, J., and Coles, J. L. (1998). The shareholder wealth implications of corporate lawsuits. *Financial Management*, pages 5–27.

Bird, A. and Karolyi, S. A. (2016). Do institutional investors demand public disclosure? *The Review of Financial Studies*, 29(12):3245–3277.

Bird, R. C., Borochin, P. A., and Knopf, J. D. (2015). The role of the chief legal officer in corporate governance. *Journal of Corporate Finance*, 34:1–22.

Bizjak, J. M. and Coles, J. L. (1995). The effect of private antitrust litigation on the stock-market valuation of the firm. *The American Economic Review*, pages 436–461.

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

Boone, A. L. and White, J. T. (2015). The effect of institutional ownership on firm transparency and information production. *Journal of Financial Economics*, 117(3):508–533.

Borochin, P. and Yang, J. (2017). The effects of institutional investor objectives on firm valuation and governance. *Journal of Financial Economics*, 126(1):171 – 199.

Borochin, P. A., Cicon, J. E., DeLisle, R. J., and Price, S. M. (2018). The effects of conference call tones on market perceptions of value uncertainty. *Journal of Financial Markets*, 40:75–91.

Bushee, B. J. (1998). The influence of institutional investors on myopic r&d investment behavior. *The Accounting Review*, 73:305–333.

Bushee, B. J. (2001). Do institutional investors prefer near-term earnings over long-run value? *Contemporary Accounting Research*, 18(2):207–246.

Bushee, B. J. and Noe, C. F. (2000). Corporate disclosure practices, institutional investors, and stock return volatility. *Journal of Accounting Research*, 38:171–202.

Chen, X., Harford, J., and Li, K. (2007). Monitoring: Which institutions matter? *Journal of Financial Economics*, 86(2):279–305.

Chen, Y., Kelly, B., and Wu, W. (2020). Sophisticated investors and market efficiency: Evidence from a natural experiment. *Journal of Financial Economics*.

Cheng, C. A., Huang, H. H., Li, Y., and Lobo, G. (2010). Institutional monitoring through shareholder litigation. *Journal of Financial Economics*, 95(3):356–383.

Chubb Limited (2019). From nuisance to menace: The rising tide of securities class action litigations. *Available at www.chubb.com.*

Cornerstone Research (2018). Securities class action settlements: 2018 review and analysis. *Available at www.cornerstone.com.*

DeFond, M., Hung, M., and Trezevant, R. (2007). Investor protection and the information content of annual earnings announcements: International evidence. *Journal of Accounting and Economics*, 43(1):37–67.

Doran, J. S., Peterson, D. R., and Price, S. M. (2012). Earnings conference call content and stock price: the case of reits. *The Journal of Real Estate Finance and Economics*, 45(2):402–434.

Gande, A. and Lewis, C. M. (2009). Shareholder-initiated class action lawsuits: Shareholder wealth effects and industry spillovers. *Journal of Financial and Quantitative Analysis*, 44(4):823–850.

Ganguly, A., Ganguly, A., Ge, L., and Zutter, C. J. (2019). Shareholder litigation and readability in financial disclosures: Evidence from a natural experiment. *Available at SSRN 3401891.*

Gaspar, J.-M., Massa, M., and Matos, P. (2005). Shareholder investment horizons and the market for corporate control. *Journal of Financial Economics*, 76(1):135–165.

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

Harford, J., Kecskés, A., and Mansi, S. (2018). Do long-term investors improve corporate decision making? *Journal of Corporate Finance*, 50:424–452.

Huang, K. and Petkevich, A. (2016). Investment horizons and information. *Journal of Business Finance and Accounting*, 43(7-8):1017–1056.

Johnson, M. F., Kasznik, R., and Nelson, K. K. (2000). Shareholder wealth effects of the private securities litigation reform act of 1995. *Review of Accounting Studies*, 5(3):217–233.

Karpoff, J. M., Koester, A., Lee, D. S., and Martin, G. S. (2017). Proxies and databases in financial misconduct research. *The Accounting Review*, 92(6):129–163.

Karpoff, J. M. and Lott Jr, J. R. (1993). The reputational penalty firms bear from committing criminal fraud. *The Journal of Law and Economics*, 36(2):757–802.

Kempf, E. and Spalt, O. G. (2019). Litigating innovation: Evidence from securities class action lawsuits. *Available at SSRN 3143690.*

Klock, M. (2015). Do class action filings affect stock prices? the stock market reaction to securities class actions post pslra. *Journal of Business & Securities Law*, 15(2):109.

Landsman, W. R. and Maydew, E. L. (2002). Has the information content of quarterly earnings announcements declined in the past three decades? *Journal of Accounting Research*, 40(3):797–808.

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

Liu, B. and McConnell, J. J. (2013). The role of the media in corporate governance: Do the media influence managers' capital allocation decisions? *Journal of Financial Economics*, 110(1):1–17.

Lo, A. W. and MacKinlay, A. C. (1988). Stock market prices do not follow random walks: Evidence from a simple specification test. *The Review of Financial Studies*, 1(1):41–66.

Loughran, T. and McDonald, B. (2011). When is a liability not a liability? textual analysis, dictionaries, and 10-ks. *The Journal of Finance*, 66(1):35–65.

Loughran, T. and McDonald, B. (2014). Measuring readability in financial disclosures. *The Journal of Finance*, 69(4):1643–1671.

McCahery, J. A., Sautner, Z., and Starks, L. T. (2016). Behind the scenes: The corporate governance preferences of institutional investors. *The Journal of Finance*, 71(6):2905–2932.

McShane, B. B., Watson, O. P., Baker, T., and Griffith, S. J. (2012). Predicting securities fraud settlements and amounts: A hierarchical bayesian model of federal securities class action lawsuits. *Journal of Empirical Legal Studies*, 9(3):482–510.

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

Rogers, J. L., Van Buskirk, A., and Zechman, S. L. (2011). Disclosure tone and shareholder litigation. *The Accounting Review*, 86(6):2155–2183.

Romano, R. (1991). The shareholder suit: Litigation without foundation. *JL Econ. & Org.*, 7:55.

Tetlock, P. C. (2011). All the news that's fit to reprint: Do investors react to stale information? *The Review of Financial Studies*, 24(5):1481–1512.

Tetlock, P. C., Saar-Tsechansky, M., and Macskassy, S. (2008). More than words: Quantifying language to measure firms' fundamentals. *The Journal of Finance*, 63(3):1437–1467.

Warner, J. B., Watts, R. L., and Wruck, K. H. (1988). Stock prices and top management changes. *Journal of Financial Economics*, 20:461–492.

Yan, X. and Zhang, Z. (2007). Institutional investors and equity returns: Are short-term institutions better informed? *The Review of Financial Studies*, 22(2):893–924.

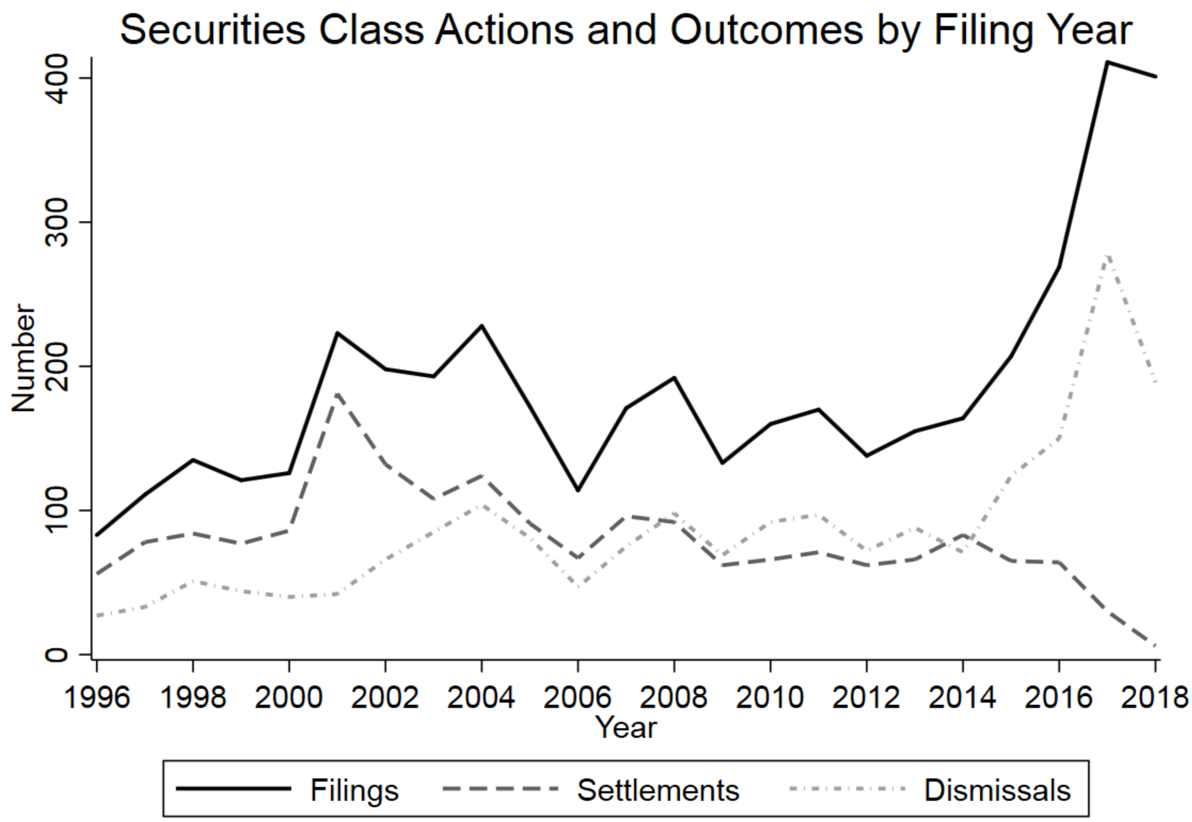

Figure 1: Number of securities class actions (SCAs) by year. The solid line represents total number of class actions filed in a given year. The dashed line represents the number of class actions filed in a given year that are settled, while the dot-dashed line represents the number that are dismissed. SCA data is from the Securities Class Action Clearinghouse database from 1996 to 2018.

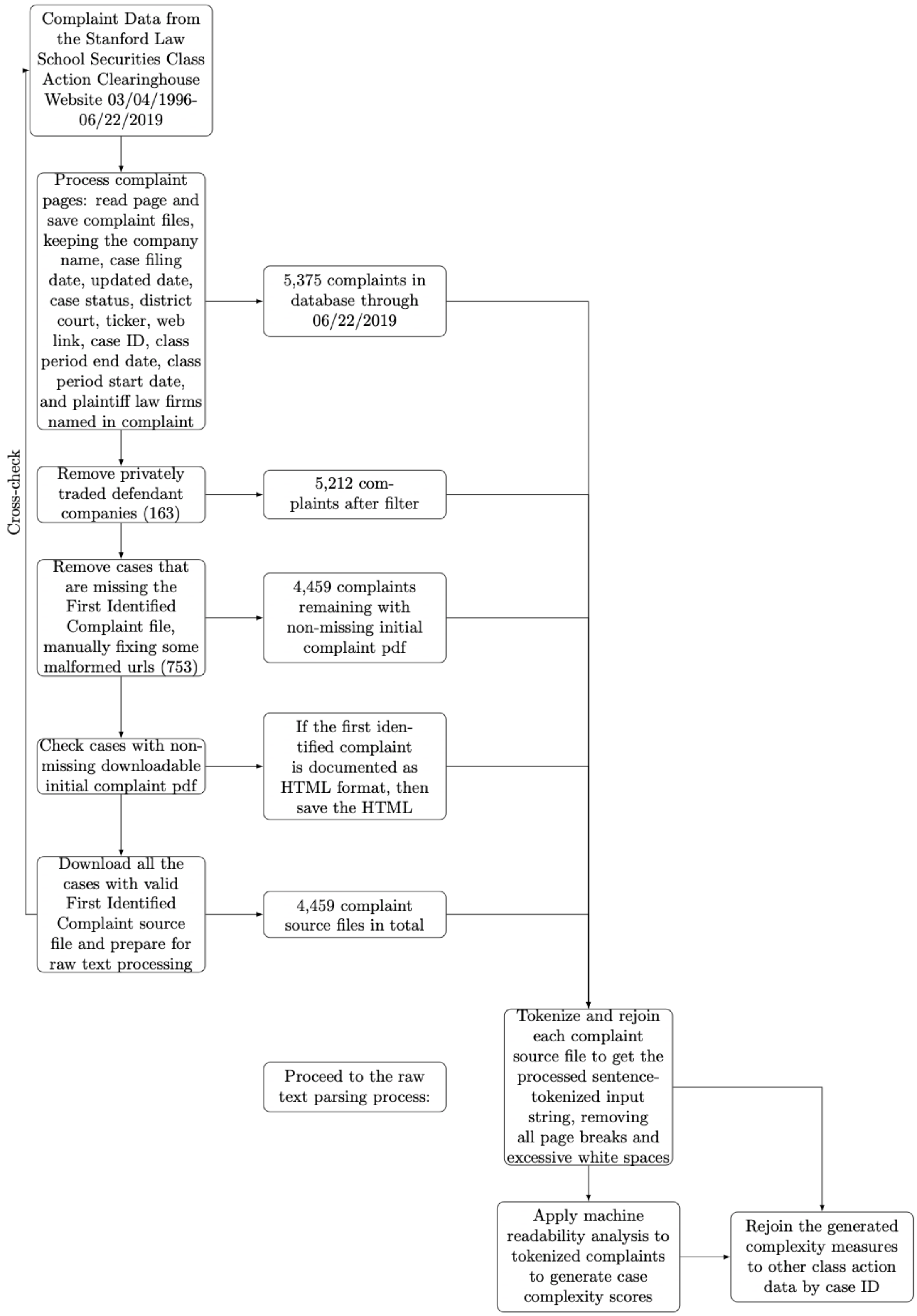

Figure 2: Data Processing and Sample Formation.

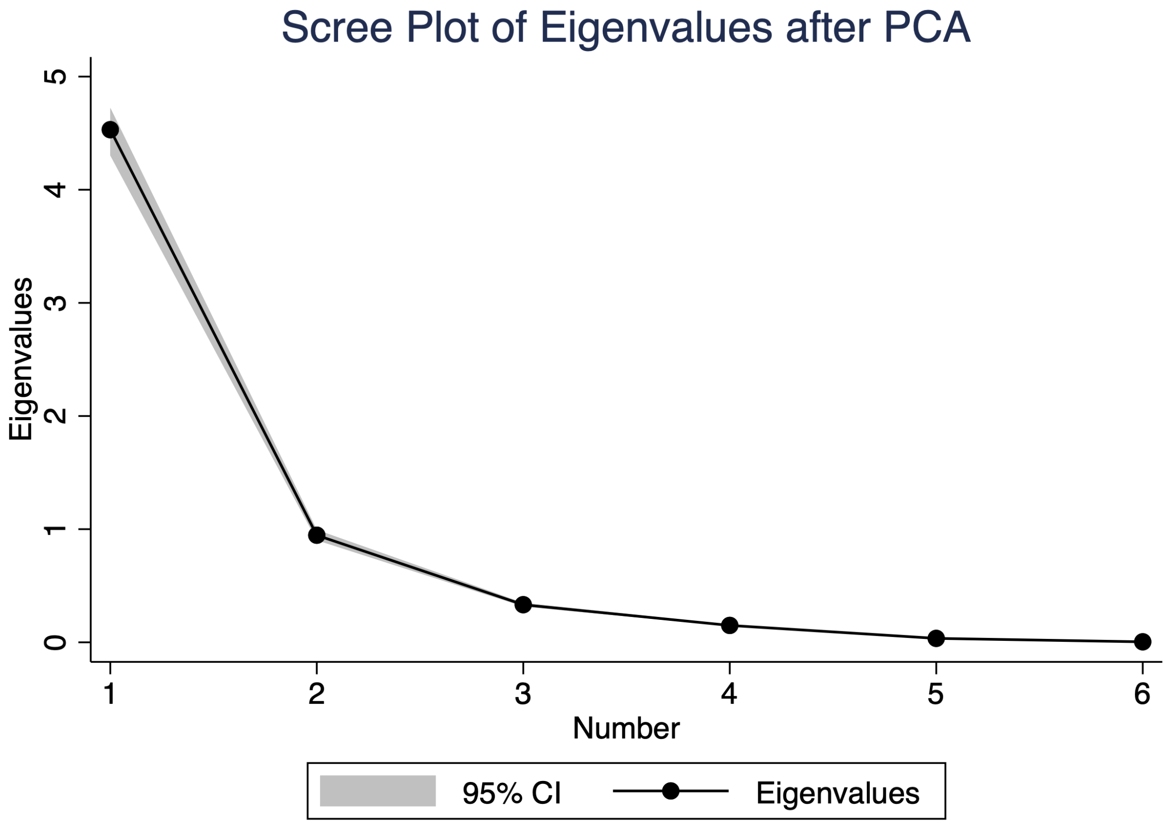

Figure 3: This figure presents the scree plot of eigenvalues from the principal component analysis of our six readability measures, ARI, SI, FRE, FKG, GF, and Log(WordCount). Eigenvalues are ordered from largest to smallest. The first principal component captures the majority of variation in the readability measures.

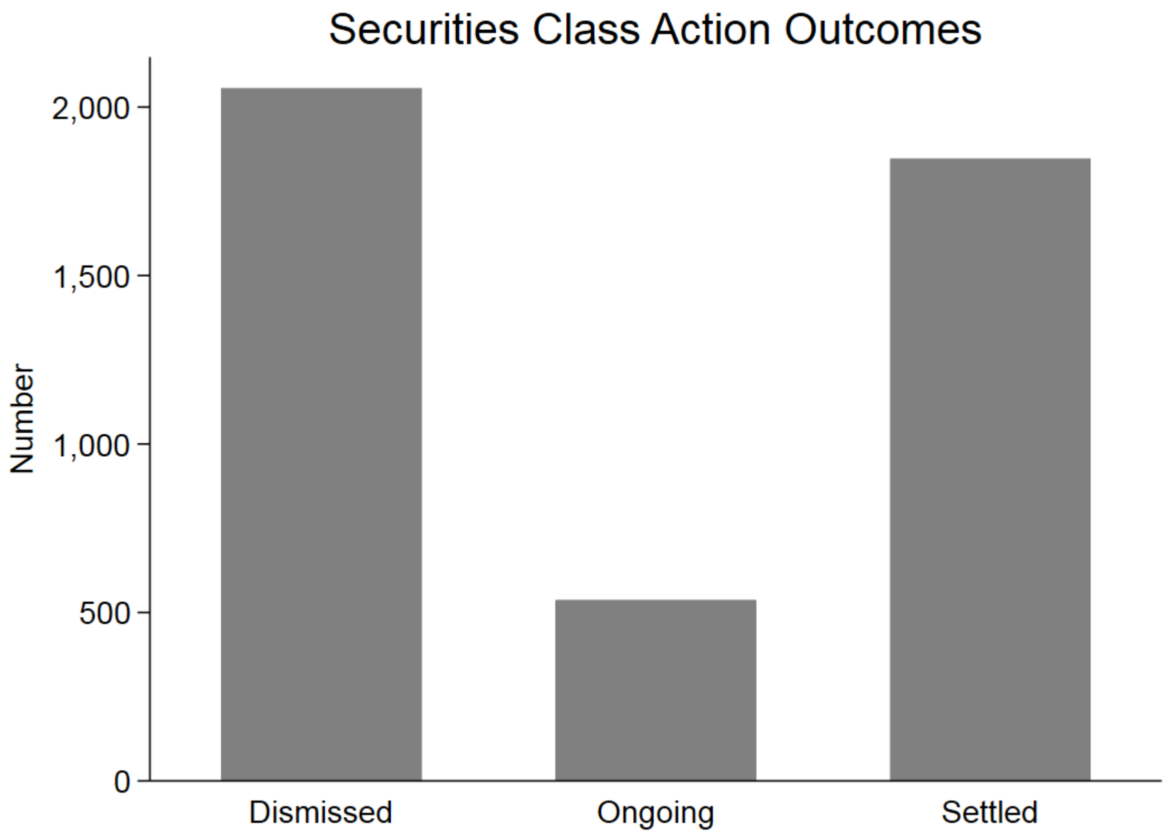

Figure 4: Summary of securities class action (SCA) outcomes. We tabulate the current status of each SCA filing in our sample, finding a mostly even split between cases that are settled and dismissed, with only a small fraction ongoing. SCA data is from the Securities Class Action Clearinghouse database from 1996 to 2018.

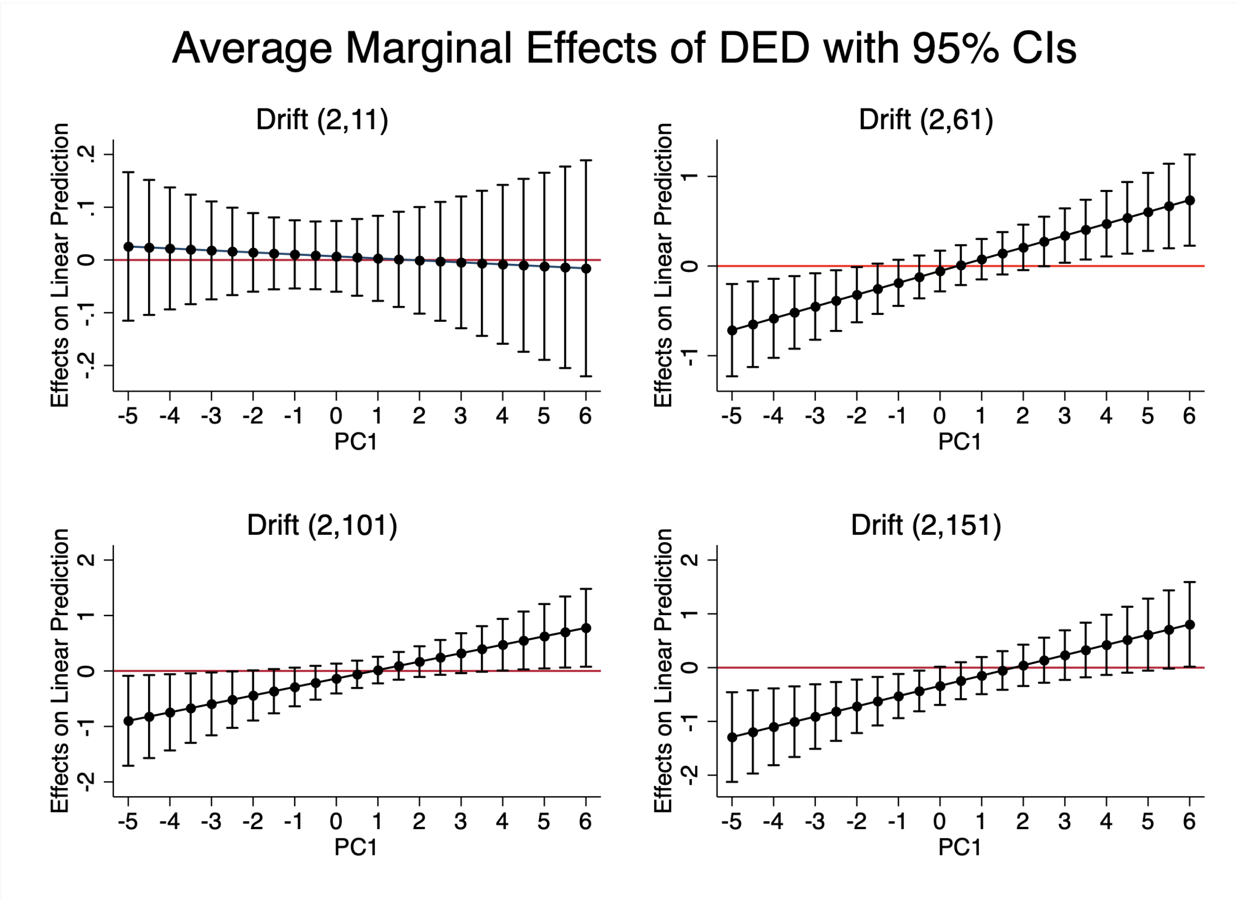

Figure 5: Average marginal effects of dedicated institutional ownership. This figure simulates the marginal effect of dedicated institutional ownership on post-SCA drift over the 10 days (2, 11), two-month (2, 61), three-month (2,101), and five-month (2, 151) event windows with respect to the changes of SCA case complexity, PC1. The lower and upper limit of the interval is assigned to the min and max value of the case complexity with a step width of 0.5. We report a 95% confidence interval (CI) around the estimates.

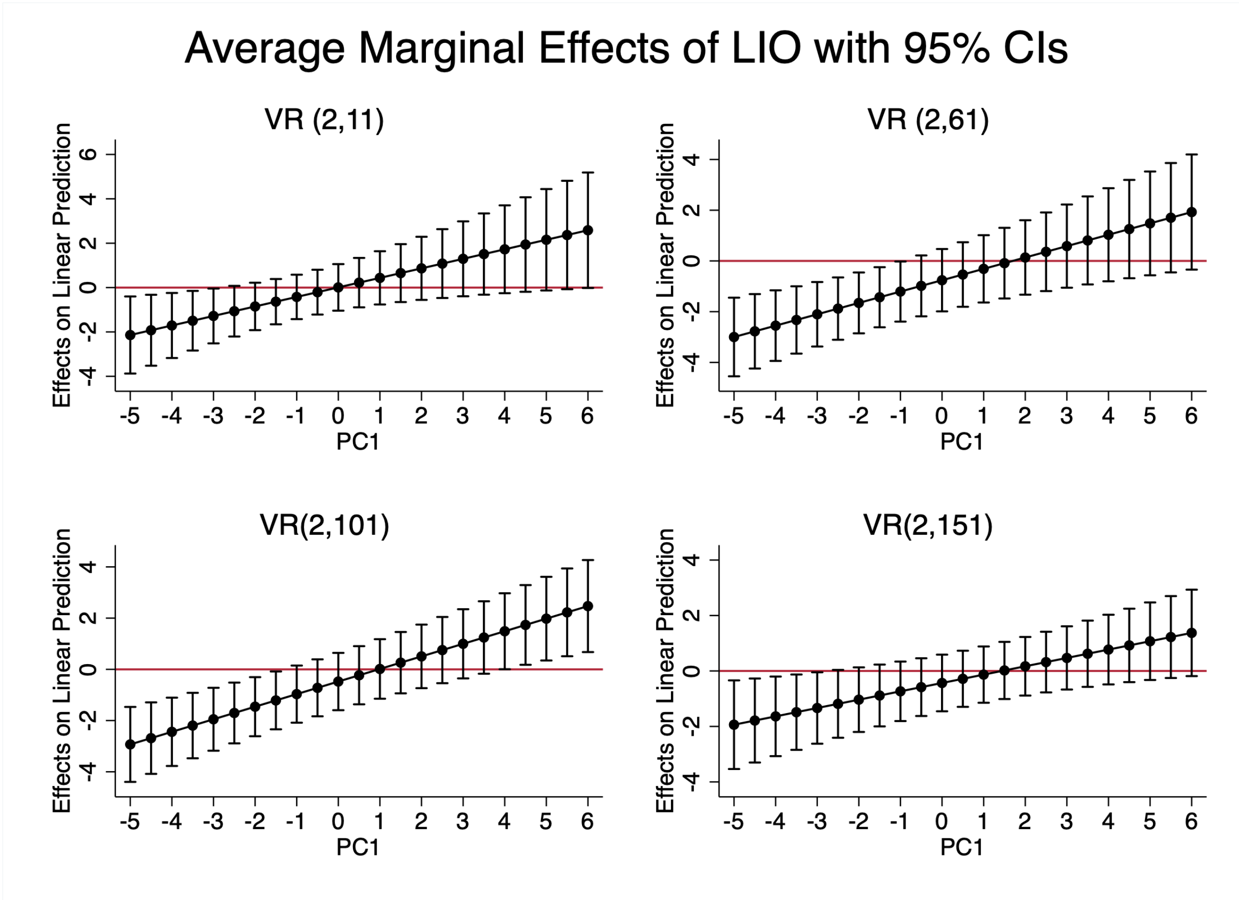

Figure 6: Average marginal effects of long-term institutional ownership. This figure simulates the marginal effect of long-term institutional ownership on post-SCA abnormal return variance ratio over the 10 days (2, 11), two-month (2, 61), three-month (2,101), and five-month (2, 151) event windows with respect to the changes of SCA case complexity, PC1. The lower and upper limit of the interval is assigned to the min and max value of the case complexity with a step width of 0.5. We report a 95% confidence interval (CI) around the estimates.

Table 1: Summary Statistics. This table reports the summary statistics for firm-level characteristics including institutional ownership levels by type, and SCA filing characteristics including their readability measures. We also present summary statistics of the Compustat universe and the SCA filing universe to evaluate the representativeness of our sample. The sample period ranges from 1996 to 2018. Variable definitions are given in Appendix A.

Panel A: Firm Characteristics								
	Regression Sample					COMPUSTAT Sample		
Variable	N	Mean	Std	Min	Max	N	Mean	Std
LIO	2,994	0.151	0.118	0.000	0.500	439,468	0.112	0.111
DEDIO	2,994	0.020	0.056	0.000	0.373	439,468	0.016	0.046
IO	2,994	0.579	0.312	0.000	1.183	439,468	0.440	0.324
Price	3,435	2.816	1.200	-0.470	6.031	435,772	2.508	1.193
Size	3,435	6.768	2.030	2.313	11.896	435,772	5.605	2.095
BM	3,353	0.789	1.342	0.000	10.808	429,118	0.729	0.746
Ret	3,427	-0.043	0.374	-1.095	1.222	433,770	0.029	0.285
Vol	3,411	0.155	0.129	0.009	0.714	431,430	0.125	0.111

Panel B: Security Class Action Lawsuits Characteristics								
	Regression Sample					SCA Sample		
Variable	N	Mean	Std	Min	Max	N	Mean	Std
ARI	2,865	20.452	2.444	15.000	27.800	3,560	20.447	2.437
SI	2,865	17.088	1.236	14.000	20.200	3,560	17.106	1.240
FRE	2,865	20.611	6.667	3.260	36.420	3,560	20.601	6.635
FKG	2,865	18.870	2.008	14.400	24.900	3,560	18.860	2.001
GF	2,865	15.822	1.937	11.620	21.780	3,560	15.808	1.920
WordCount	2,865	8930.913	4649.759	2957.000	28401.000	3,560	8862.896	4618.186
CaseDuration	2,863	932.196	825.506	0.000	5903.000	3,558	924.169	814.050
LegalComplexity	2,865	18.185	34.998	0.000	1390.000	3,560	18.126	32.889
FilingTime	2,851	85.395	163.196	0.000	1775.000	3,544	81.421	163.915

Table 2: Correlation Matrix. This table reports the pairwise correlation matrix of our variables including firm characteristics, institutional ownership, and SCA filing complexity measures. The sample period ranges from 1996 to 2018. Variable definitions are given in Appendix A.

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(1)	LIO	1.0000								
(2)	DEDIO	0.3196***	1.0000							
(3)	Log(CaseDuration)	-0.2513***	-0.0971***	1.0000						
(4)	ARI	-0.0722***	-0.0323	0.1376***	1.0000					
(5)	SI	0.0152	0.0073	-0.0159	0.8618***	1.0000				
(6)	FRE	0.0655***	0.0346	-0.0648***	-0.8176***	-0.8436***	1.0000			
(7)	FKG	-0.0763***	-0.0400*	0.1210***	0.9691***	0.8655***	-0.8858***	1.0000		
(8)	GF	-0.0864***	-0.0401*	0.1062***	0.9373***	0.7745***	-0.7040***	0.9377***	1.0000	
(9)	Log(WordCount)	0.0167	0.0083	0.1852***	0.2910***	0.2086***	-0.1330***	0.2697***	0.1653***	1.0000
(10)	IO	0.6956***	0.2164***	-0.1047***	-0.0194	0.0171	0.0309	-0.0305	-0.0394*	0.0607**
(11)	Price	0.3041***	0.0025	-0.0316	-0.0541**	-0.0498**	0.0638***	-0.0616**	-0.0548**	-0.0051
(12)	Size	0.4329***	0.0201	-0.0016	-0.0773***	-0.0859***	0.1168***	-0.0883***	-0.0767***	0.0546**
(13)	BM	-0.0335	-0.0145	0.1173***	-0.0076	-0.0057	-0.0250	-0.0007	-0.0347	0.0912***
(14)	Ret	0.0979***	0.0232	-0.1809***	0.0003	0.0522**	-0.0035	0.0029	0.0183	-0.0860***
(15)	Vol	-0.2787***	-0.0619***	0.1446***	0.0420*	0.0015	-0.0275	0.0369*	0.0497**	-0.0171
(16)	Delaware	0.1231***	0.1404***	-0.2265***	-0.1391***	-0.0608**	0.0699***	-0.1369***	-0.1411***	-0.1083***
(17)	Log(LegalComplexity)	0.0787***	0.0519**	-0.1949***	0.0043	0.1100***	-0.0468*	-0.0024	-0.0745***	0.2272***
(18)	Log(FilingTime)	-0.1523***	-0.0789***	0.4302***	0.1047***	-0.0804***	-0.0038	0.0791***	0.0763***	0.2241***
		(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
(10)	IO	1.0000								
(11)	Price	0.3896***	1.0000							
(12)	Size	0.4747***	0.6555***	1.0000						
(13)	BM	-0.1105***	-0.2868***	-0.1945***	1.0000					
(14)	Ret	0.0988***	0.1756***	0.1544***	-0.2300***	1.0000				
(15)	Vol	-0.2681***	-0.2807***	-0.3369***	0.0751***	0.1609***	1.0000			
(16)	Delaware	0.0748***	0.0337	0.0120	-0.0296	0.0796***	-0.0339	1.0000		
(17)	Log(LegalComplexity)	0.0383*	-0.0684***	-0.0432*	0.0752***	0.0637***	-0.0215	0.1463***	1.0000	
(18)	Log(FilingTime)	-0.0999***	-0.1170***	-0.0585**	0.1423***	-0.1754***	0.1090***	-0.2094***	-0.2386***	1.0000

Table 3: Case Status and Machine Readability. This table reports the panel probit regressions of *Settled* on readability measures of the security class action lawsuit filings and firm-level control characteristics for the period of 1996-2018. *Settled* equals one if the security class action lawsuit case status is settled, and zero otherwise. The readability measures include ARI, SI, FRE, FKG, GF, and Log(WordCount). PC1 is the first principal component of the six readability measures. Firm-level control characteristics include IO, Price, Size, BM, Ret, Vol, Delaware, Log(LegalComplexity), and Log(FilingTime). Variable definitions are given in Appendix A. We control for circuit, year-quarter, and industry fixed effects. t-statistics based on standard errors clustered at the year-quarter level are reported in parentheses. Statistical significance is denoted by *, **, and *** at the 10%, 5%, and 1% levels, respectively.

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
ARI_t	0.049***							0.047***						
	(4.14)							(3.44)						
SI_t		0.097***							0.082***					
		(4.36)							(3.29)					
FRE_t			-0.018***							-0.015***				
			(-4.21)							(-3.03)				
FKG_t				0.059***							0.056***			
				(3.98)							(3.22)			
GF_t					0.043***							0.044**		
					(2.74)							(2.51)		
$Log(WordCount)_t$						0.189***							0.180**	
						(3.10)							(2.33)	
$PC1_t$							0.057***							0.052***
							(4.19)							(3.30)
IO_{t-1}								-0.124	-0.126	-0.120	-0.118	-0.113	-0.112	-0.122
								(-1.12)	(-1.13)	(-1.08)	(-1.06)	(-1.03)	(-1.02)	(-1.10)
$Price_{t-1}$								0.082**	0.083**	0.081**	0.082**	0.083**	0.082**	0.082**
								(2.15)	(2.17)	(2.13)	(2.15)	(2.17)	(2.16)	(2.15)
$Size_{t-1}$								-0.065***	-0.063***	-0.063***	-0.064***	-0.067***	-0.075***	-0.064***
								(-2.66)	(-2.62)	(-2.58)	(-2.63)	(-2.76)	(-3.12)	(-2.62)
BM_{t-1}								0.049	0.048	0.046	0.047	0.048	0.042	0.048
								(0.95)	(0.93)	(0.91)	(0.92)	(0.94)	(0.82)	(0.92)
Ret_{t-1}								-0.154	-0.155	-0.153	-0.156	-0.153	-0.124	-0.155
								(-1.58)	(-1.59)	(-1.56)	(-1.59)	(-1.56)	(-1.30)	(-1.59)
Vol_{t-1}								0.330	0.324	0.330	0.338	0.328	0.283	0.332
								(1.31)	(1.30)	(1.32)	(1.35)	(1.31)	(1.15)	(1.32)
$Delaware_t$								-0.400	-0.402	-0.402	-0.402	-0.421	-0.444*	-0.399
								(-1.55)	(-1.56)	(-1.57)	(-1.56)	(-1.63)	(-1.69)	(-1.55)
$Log(LegalComplexity)_t$								0.042	0.038	0.042	0.044	0.051	0.017	0.041
								(1.27)	(1.13)	(1.25)	(1.32)	(1.55)	(0.48)	(1.23)
$Log(FilingTime)_t$								-0.029*	-0.024	-0.025	-0.028	-0.027	-0.037**	-0.027
								(-1.71)	(-1.39)	(-1.44)	(-1.62)	(-1.60)	(-2.12)	(-1.59)
N	3355	3355	3355	3355	3355	3355	3355	2596	2596	2596	2596	2596	2596	2596
PseudoR^2	0.210	0.210	0.210	0.209	0.207	0.207	0.210	0.205	0.204	0.204	0.205	0.203	0.202	0.205
Circuit FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 4: Case Duration and Machine Readability. This table reports the panel OLS regressions of case duration on readability measures of the security class action lawsuit filings and firm-level control characteristics for the period of 1996-2018. Case duration is defined as the log of the number of days between the filing date and the status date. Our filing readability measures are ARI, SI, FRE, FKG, GF, and Log(WordCount). PC1 is the first principal component of the six readability measures. Firm-level control characteristics include IO, Price, Size, BM, Ret, Vol, Delaware, Log(LegalComplexity), and Log(FilingTime). Variable definitions are given in Appendix A. We control for circuit, year-quarter, and industry fixed effects. t-statistics, based on standard errors clustered at the year-quarter level, are reported in parentheses. Statistical significance is denoted by *, **, and *** at the 10%, 5%, and 1% levels, respectively.

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
ARI_t	0.028***							0.025***						
	(2.94)							(2.69)						
SI_t		0.016							0.036**					
		(1.00)							(2.18)					
FRE_t			-0.003							-0.006*				
			(-1.08)							(-1.80)				
FKG_t				0.033***							0.030**			
				(2.78)							(2.62)			
GF_t					0.032**							0.024*		
					(2.49)							(1.97)		
$Log(WordCount)_t$						0.226***							0.205***	
						(4.57)							(3.07)	
$PC1_t$							0.026**							0.027**
							(2.48)							(2.56)
IO_{t-1}								-0.169**	-0.169**	-0.167**	-0.167**	-0.164**	-0.166**	-0.168**
								(-2.37)	(-2.37)	(-2.34)	(-2.33)	(-2.29)	(-2.32)	(-2.35)
$Price_{t-1}$								0.059***	0.060***	0.059***	0.059***	0.060***	0.059***	0.059***
								(2.92)	(2.93)	(2.90)	(2.91)	(2.92)	(2.88)	(2.91)
$Size_{t-1}$								0.063***	0.063***	0.063***	0.063***	0.062***	0.056***	0.063***
								(4.07)	(4.09)	(4.10)	(4.08)	(4.03)	(3.69)	(4.08)
BM_{t-1}								0.065***	0.064***	0.063***	0.063***	0.064***	0.057***	0.064***
								(2.98)	(3.00)	(3.01)	(2.96)	(2.98)	(2.76)	(2.97)
Ret_{t-1}								-0.265***	-0.265***	-0.264***	-0.266***	-0.266***	-0.240***	-0.266***
								(-4.78)	(-4.77)	(-4.72)	(-4.78)	(-4.78)	(-4.45)	(-4.77)
Vol_{t-1}								0.552***	0.552***	0.554***	0.557***	0.553***	0.528***	0.554***
								(3.41)	(3.41)	(3.41)	(3.43)	(3.40)	(3.26)	(3.42)
$Delaware_t$								-0.233	-0.241	-0.249	-0.235	-0.244	-0.244	-0.236
								(-1.29)	(-1.33)	(-1.37)	(-1.30)	(-1.36)	(-1.38)	(-1.30)
$Log(LegalComplexity)_t$								-0.064**	-0.065**	-0.062**	-0.063**	-0.058**	-0.092***	-0.064**
								(-2.37)	(-2.43)	(-2.34)	(-2.34)	(-2.18)	(-3.12)	(-2.39)
$Log(FilingTime)_t$								0.102***	0.105***	0.105***	0.103***	0.104***	0.092***	0.104***
								(7.28)	(7.28)	(7.30)	(7.28)	(7.28)	(6.51)	(7.28)
N	3558	3558	3558	3558	3558	3558	3558	2847	2847	2847	2847	2847	2847	2847
Adjusted R^2	0.399	0.396	0.396	0.398	0.398	0.402	0.398	0.453	0.452	0.451	0.453	0.452	0.455	0.452
Circuit FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 5: Two-Sample Tests of Price Reaction on Machine Readability and Institutional Ownership. This table presents the comparison of means of the filing date price reaction for two subsamples by medians of the principal component of the six machine readability measures PC1 and of institutional ownership levels by type. Post-announcement drift is defined as the cumulative return in excess of the value-weight market return over the event window of (2, N). Post-announcement variance ratio is defined as the variance of stock return in excess of the value-weighted market return over the event window of (2, N), scaled by the variance of stock return in excess of the value-weighted market return over the estimation window of (-120, -21). N ranges from 11, 61, 101, and 151. All variables are defined in Appendix A. The table reports the means of price reaction of the two subsamples, the t-values, and p-value of the differences in the two subsample means.

Readability (PC1)								
	Drift				Abnormal Return Variance			
			Difference in Mean				Difference in Mean	
	More Readable	Less Readable	t-value	p-value	More Readable	Less Readable	t-value	p-value
(2, 11)	-0.017	-0.010	-1.58	0.0570	1.802	1.459	3.44	0.0003
(2, 61)	-0.005	0.004	-0.75	0.2268	1.915	1.645	2.81	0.0025
(2, 101)	-0.010	0.013	-1.55	0.0608	1.884	1.654	2.57	0.0052
(2, 151)	-0.017	0.009	-1.42	0.0776	1.820	1.694	1.47	0.0710
LIO								
	Drift				Abnormal Return Variance			
			Difference in Mean				Difference in Mean	
	Below Median	Above Median	t-value	p-value	Below Median	Above Median	t-value	p-value
(2, 11)	-0.024	-0.001	-4.70	0.0000	1.663	1.544	1.26	0.1040
(2, 61)	-0.030	0.020	-4.14	0.0000	1.656	1.722	-0.78	0.2167
(2, 101)	-0.020	0.020	-2.63	0.0043	1.577	1.762	-2.35	0.0094
(2, 151)	-0.039	0.033	-3.91	0.0000	1.525	1.737	-2.88	0.0020
DEDIO								
	Drift				Abnormal Return Variance			
			Difference in Mean				Difference in Mean	
	Below Median	Above Median	t-value	p-value	Below Median	Above Median	t-value	p-value
(2, 11)	-0.020	-0.004	-3.35	0.0004	1.597	1.610	-0.14	0.4447
(2, 61)	-0.012	0.002	-1.16	0.1225	1.639	1.751	-1.31	0.0956
(2, 101)	0.000	-0.003	0.20	0.4223	1.555	1.810	-3.23	0.0006
(2, 151)	-0.011	0.002	-0.66	0.2535	1.508	1.780	-3.67	0.0001

Table 6: Post-Announcement Drift, Machine Readability, and Institutional Ownership. This table reports the panel OLS regressions of Post-announcement drift on readability measures of the security class action lawsuit filings, institutional ownership measures, and firm-level control characteristics for the period of 1996-2018. Post-announcement drift is defined as the cumulative return in excess of the value-weight market return over the event window of (2, N). N ranges from 11, 61, 101, and 151. PC1 is the first principal component of the six readability measures, ARI, SI, FRE, FKG, GF, and Log(WordCount). All variables are defined in Appendix A. We control for circuit, year-quarter, and industry fixed effects. t-statistics based on standard errors clustered at the year-quarter level are reported in parentheses. Statistical significance is denoted by *, **, and *** at the 10%, 5%, and 1% levels, respectively.

	Panel A: (2, 11)		Panel B: (2, 61)		Panel C: (2, 101)		Panel D: (2, 151)	
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
LIO_{t-1}	-0.001		0.267***		0.278**		0.361**	
	(-0.05)		(2.80)		(2.38)		(2.44)	
$LIO_{t-1}*PC1_t$	0.007		0.012		0.026		0.030	
	(0.84)		(0.43)		(0.77)		(0.78)	
$DEDIO_{t-1}$		0.002		-0.051		-0.130		-0.314
		(0.05)		(-0.36)		(-0.76)		(-1.44)
$DEDIO_{t-1}*PC1_t$		-0.006		0.123**		0.144*		0.185**
		(-0.31)		(2.36)		(1.74)		(2.19)
$PC1_t$	-0.001	0.000	-0.005	-0.005	-0.007	-0.006	-0.008	-0.006
	(-0.48)	(0.18)	(-0.80)	(-1.32)	(-0.99)	(-1.14)	(-0.92)	(-1.11)
$EventDayRet_t$	-0.011	-0.012	-0.209***	-0.207***	-0.220***	-0.220***	-0.308***	-0.307***
	(-0.29)	(-0.30)	(-2.89)	(-2.85)	(-2.71)	(-2.66)	(-3.13)	(-3.07)
IO_{t-1}	0.027**	0.027**	0.008	0.066**	0.001	0.065*	0.018	0.107***
	(1.99)	(2.28)	(0.21)	(2.42)	(0.01)	(1.96)	(0.34)	(2.63)
$Price_{t-1}$	-0.005	-0.005	-0.044***	-0.046***	-0.046***	-0.048***	-0.065***	-0.068***
	(-1.53)	(-1.53)	(-5.36)	(-5.53)	(-4.52)	(-4.66)	(-6.16)	(-6.29)
$Size_{t-1}$	0.008***	0.008***	0.009*	0.011**	0.007	0.009	0.009	0.012
	(4.78)	(4.82)	(1.75)	(2.29)	(0.97)	(1.34)	(1.20)	(1.56)
BM_{t-1}	0.004	0.004	-0.005	-0.003	0.027*	0.029*	0.031*	0.033**
	(0.85)	(0.85)	(-0.54)	(-0.36)	(1.82)	(1.88)	(1.96)	(2.07)
Ret_{t-1}	0.000	0.001	0.011	0.007	0.018	0.013	0.003	-0.004
	(0.05)	(0.06)	(0.43)	(0.28)	(0.57)	(0.41)	(0.08)	(-0.11)
Vol_{t-1}	-0.021	-0.021	-0.023	-0.027	-0.027	-0.030	0.045	0.041
	(-0.72)	(-0.71)	(-0.24)	(-0.28)	(-0.28)	(-0.31)	(0.42)	(0.39)
$Delaware_t$	0.020	0.019	-0.029	-0.015	0.009	0.025	-0.005	0.020
	(1.50)	(1.39)	(-0.68)	(-0.37)	(0.17)	(0.46)	(-0.07)	(0.31)
$Log(LegalComplexity)_t$	0.001	0.001	-0.009	-0.009	-0.017**	-0.017**	-0.028***	-0.029***
	(0.38)	(0.39)	(-1.41)	(-1.50)	(-2.30)	(-2.39)	(-3.09)	(-3.22)
$Log(FilingTime)_t$	-0.003**	-0.003**	-0.002	-0.001	0.003	0.004	0.001	0.003
	(-2.42)	(-2.42)	(-0.58)	(-0.32)	(0.60)	(0.81)	(0.19)	(0.45)
N	2761	2761	2403	2403	2279	2279	2165	2165
Adjusted R^2	0.035	0.035	0.069	0.068	0.102	0.101	0.122	0.121
Circuit FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 7: Post-Announcement Variance Ratio, Machine Readability, and Institutional Ownership. This table reports the panel OLS regressions of Post-announcement variance ratio on readability measures of the security class action lawsuit filings, institutional ownership measures, and firm-level control characteristics for the period of 1996-2018. Post-announcement variance ratio is defined as the variance of stock return in excess of the value-weighted market return over the event window of (2, N), scaled by the variance of stock return in excess of the value-weighted market return over the estimation window of (-120, -21). N ranges from 11, 61, 101, and 151. PC1 is the first principal component of the six readability measures, ARI, SI, FRE, FKG, GF, and Log(WordCount). All variables are defined in Appendix A. We control for circuit, year-quarter, and industry fixed effects. t-statistics based on standard errors clustered at the year-quarter level are reported in parentheses. Statistical significance is denoted by *, **, and *** at the 10%, 5%, and 1% levels, respectively.

	Panel A: (2, 11)		Panel B: (2, 61)		Panel C: (2, 101)		Panel D: (2, 151)	
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
LIO_{t-1}	0.608		-0.342		-0.030		0.002	
	(0.98)		(-0.43)		(-0.04)		(0.00)	
$LIO_{t-1}*PC1_t$	0.393*		0.340**		0.396***		0.196	
	(1.94)		(2.18)		(3.13)		(1.42)	
$DEDIO_{t-1}$		-0.594		-0.061		0.042		0.042
		(-0.78)		(-0.06)		(0.05)		(0.04)
$DEDIO_{t-1}*PC1_t$		0.115		0.019		0.057		-0.018
		(0.37)		(0.04)		(0.15)		(-0.03)
$PC1_t$	0.007	0.061**	-0.013	0.033	-0.026	0.028	-0.005	0.022
	(0.20)	(2.34)	(-0.40)	(1.47)	(-0.89)	(1.49)	(-0.16)	(1.22)
$EventDayRet_t$	-4.022***	-4.031***	-2.342***	-2.360***	-1.883***	-1.904***	-1.549***	-1.562***
	(-4.79)	(-4.78)	(-4.13)	(-4.17)	(-3.70)	(-3.75)	(-3.36)	(-3.40)
IO_{t-1}	-0.458**	-0.289	0.125	0.073	0.005	0.017	-0.028	-0.020
	(-2.01)	(-1.45)	(0.62)	(0.40)	(0.04)	(0.11)	(-0.16)	(-0.13)
$Price_{t-1}$	0.044	0.036	0.005	0.005	0.024	0.023	0.038	0.038
	(0.78)	(0.64)	(0.09)	(0.09)	(0.53)	(0.52)	(0.88)	(0.88)
$Size_{t-1}$	0.034	0.037	-0.011	-0.015	-0.011	-0.013	-0.011	-0.011
	(1.02)	(1.06)	(-0.43)	(-0.60)	(-0.43)	(-0.50)	(-0.40)	(-0.41)
BM_{t-1}	-0.106*	-0.104*	-0.066	-0.069	-0.025	-0.026	-0.046	-0.046
	(-1.79)	(-1.74)	(-0.95)	(-0.98)	(-0.45)	(-0.44)	(-0.96)	(-0.94)
Ret_{t-1}	-0.279	-0.280	-0.216	-0.208	-0.084	-0.078	-0.038	-0.034
	(-1.56)	(-1.56)	(-1.60)	(-1.53)	(-0.70)	(-0.64)	(-0.37)	(-0.33)
Vol_{t-1}	-1.228***	-1.244***	-1.693***	-1.698***	-2.159***	-2.164***	-2.175***	-2.179***
	(-2.84)	(-2.86)	(-3.27)	(-3.26)	(-5.62)	(-5.58)	(-5.89)	(-5.86)
$Delaware_t$	-0.201	-0.219	-0.563	-0.622*	-0.414	-0.456*	-0.412*	-0.429*
	(-0.76)	(-0.82)	(-1.65)	(-1.81)	(-1.65)	(-1.73)	(-1.76)	(-1.72)
$Log(LegalComplexity)_t$	-0.203***	-0.203***	-0.073	-0.073	-0.043	-0.043	-0.023	-0.023
	(-3.83)	(-3.81)	(-1.34)	(-1.32)	(-0.76)	(-0.74)	(-0.41)	(-0.40)
$Log(FilingTime)_t$	-0.127***	-0.124***	-0.113***	-0.114***	-0.105***	-0.106***	-0.108***	-0.109***
	(-3.79)	(-3.72)	(-3.99)	(-4.19)	(-3.99)	(-4.13)	(-4.53)	(-4.65)
N	2738	2738	2384	2384	2260	2260	2147	2147
Adjusted R^2	0.134	0.133	0.110	0.109	0.176	0.174	0.196	0.195
Circuit FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 8: Post-Announcement Relative Quoted Spread, Machine Readability, and Institutional Ownership. This table reports the panel OLS regressions of Post-announcement relative quoted spread on readability measures of the security class action lawsuit filings, institutional ownership measures, and firm-level control characteristics for the period of 1996-2018. Post-announcement spread is defined as the average of daily relative quoted spread (defined as the ratio of the absolute value of the difference between the bid and ask prices to the sum of the bid and ask prices) over the event window of (2, N). N ranges from 11, 61, 101, and 151. PC1 is the first principal component of the six readability measures, ARI, SI, FRE, FKG, GF, and Log(WordCount). All variables are defined in Appendix A. We control for circuit, year-quarter, and industry fixed effects. t-statistics based on standard errors clustered at the year-quarter level are reported in parentheses. Statistical significance is denoted by *, **, and *** at the 10%, 5%, and 1% levels, respectively.

	Panel A: (2, 11)		Panel B: (2, 61)		Panel C: (2, 101)		Panel D: (2, 151)	
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
LIO_{t-1}	0.015***		0.013***		0.010***		0.009***	
	(3.45)		(3.72)		(2.80)		(2.64)	
LIO_{t-1}*$PC1_t$	-0.002*		-0.002**		-0.001		-0.002*	
	(-1.87)		(-2.19)		(-1.57)		(-1.91)	
$DEDIO_{t-1}$		-0.002		0.001		0.001		0.002
		(-0.62)		(0.46)		(0.29)		(0.57)
$DEDIO_{t-1}$*$PC1_t$		-0.003**		-0.002**		-0.002		-0.002
		(-2.54)		(-2.02)		(-1.43)		(-1.45)
$PC1_t$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	(1.34)	(0.93)	(1.56)	(1.00)	(0.75)	(0.14)	(1.04)	(0.35)
IO_{t-1}	-0.008***	-0.005***	-0.008***	-0.005***	-0.007***	-0.005***	-0.007***	-0.005***
	(-5.65)	(-5.62)	(-6.89)	(-6.79)	(-6.38)	(-7.40)	(-6.24)	(-7.66)
$Price_{t-1}$	-0.002***	-0.002***	-0.001***	-0.001***	-0.001***	-0.001***	-0.001**	-0.001**
	(-4.45)	(-4.49)	(-3.44)	(-3.51)	(-3.04)	(-3.13)	(-2.48)	(-2.56)
$Size_{t-1}$	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***
	(-8.24)	(-8.46)	(-9.65)	(-9.96)	(-10.04)	(-10.33)	(-10.33)	(-10.55)
BM_{t-1}	0.000	0.000	0.001**	0.001**	0.001***	0.001***	0.001***	0.001***
	(0.05)	(0.38)	(2.15)	(2.38)	(3.43)	(3.57)	(3.42)	(3.55)
Ret_{t-1}	-0.002	-0.002	-0.001	-0.001	-0.000	-0.000	-0.000	-0.000
	(-1.28)	(-1.38)	(-0.94)	(-1.04)	(-0.40)	(-0.52)	(-0.31)	(-0.46)
Vol_{t-1}	0.002	0.001	-0.005*	-0.005*	-0.006**	-0.006**	-0.006**	-0.006**
	(0.57)	(0.54)	(-1.84)	(-1.85)	(-2.23)	(-2.25)	(-2.06)	(-2.05)
$Delaware_t$	-0.002	-0.002	-0.002*	-0.001	-0.002*	-0.002	-0.002	-0.002
	(-1.58)	(-1.33)	(-1.77)	(-1.26)	(-1.88)	(-1.59)	(-1.47)	(-1.30)
$Log(LegalComplexity)_t$	0.000	0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.31)	(0.39)	(-0.49)	(-0.59)	(-0.95)	(-1.04)	(-0.91)	(-1.06)
$Log(FilingTime)_t$	0.000*	0.000**	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	(1.75)	(2.09)	(-0.55)	(-0.06)	(-1.28)	(-0.84)	(-1.43)	(-0.96)
N	2741	2741	2398	2398	2281	2281	2180	2180
Adjusted R^2	0.510	0.506	0.535	0.532	0.559	0.557	0.566	0.563
Circuit FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 9: Post-Announcement Drift, Machine Readability, and Instrumented Institutional Ownership. This table reports 2SLS regressions of Post-announcement drift on readability measures of the security class action lawsuit filings, instrumented institutional ownership measures, and firm-level control characteristics for the period of 1996-2018. The instrumental variable is a firm's average industry peer institutional ownership measures. Post-announcement drift is defined as the cumulative return in excess of the value-weight market return over the event window of (2, N). N ranges from 11, 61, 101, and 151. PC1 is the first principal component of the six readability measures, ARI, SI, FRE, FKG, GF, and Log(WordCount). All variables are defined in Appendix A. We control for circuit, year-quarter, and industry fixed effects. t-statistics based on standard errors clustered at the year-quarter level are reported in parentheses. Statistical significance is denoted by *, **, and *** at the 10%, 5%, and 1% levels, respectively.

	Panel A: (2, 11)		Panel B: (2, 61)		Panel C: (2, 101)		Panel D: (2, 151)	
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\widehat{LIO_{t-1}}$	0.003		0.262***		0.293**		0.391***	
	(0.08)		(2.80)		(2.46)		(2.79)	
$\widehat{LIO_{t-1}}*PC1_t$	0.006		0.011		0.022		0.021	
	(0.58)		(0.43)		(0.72)		(0.58)	
$\widehat{DEDIO}_{t-1}$		0.004		-0.073		-0.133		-0.302
		(0.11)		(-0.56)		(-0.75)		(-1.47)
$\widehat{DEDIO}_{t-1}*PC1_t$		-0.006		0.129**		0.160**		0.197**
		(-0.33)		(2.11)		(1.97)		(2.26)
$PC1_t$	-0.001	0.000	-0.005	-0.006*	-0.007	-0.006	-0.006	-0.006
	(-0.38)	(0.07)	(-1.01)	(-1.72)	(-1.12)	(-1.53)	(-0.83)	(-1.25)
N	2690	2690	2388	2388	2266	2266	2153	2153
Adjusted R^2	0.100	0.099	0.140	0.139	0.173	0.172	0.192	0.191
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Circuit FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 10: Post-Announcement Variance Ratio, Machine Readability, and Instrumented Institutional Ownership. This table reports 2SLS regressions of Post-announcement variance ratio on readability measures of the security class action lawsuit filings, instrumented institutional ownership measures, and firm-level control characteristics for the period of 1996-2018. The instrumental variable is a firm's average industry peer institutional ownership measures. Post-announcement variance ratio is defined as the variance of stock return in excess of the value-weighted market return over the event window of (2, N), scaled by the variance of stock return in excess of the value-weighted market return over the estimation window of (-120, -21). N ranges from 11, 61, 101, and 151. PC1 is the first principal component of the six readability measures, ARI, SI, FRE, FKG, GF, and Log(WordCount). All variables are defined in Appendix A. We control for circuit, year-quarter, and industry fixed effects. t-statistics based on standard errors clustered at the year-quarter level are reported in parentheses. Statistical significance is denoted by *, **, and *** at the 10%, 5%, and 1% levels, respectively.

	Panel A: (2, 11)		Panel B: (2, 61)		Panel C: (2, 101)		Panel D: (2, 151)	
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\widehat{LIO_{t-1}}$	0.483		-0.087		0.222		0.111	
	(0.78)		(-0.13)		(0.37)		(0.19)	
$\widehat{LIO_{t-1}}*PC1_t$	0.381**		0.379**		0.425***		0.245*	
	(1.99)		(2.45)		(2.92)		(1.72)	
$\widehat{DEDIO_{t-1}}$		-0.631		0.014		0.123		0.142
		(-0.76)		(0.01)		(0.15)		(0.14)
$\widehat{DEDIO_{t-1}}*PC1_t$		0.143		0.031		0.065		-0.019
		(0.43)		(0.08)		(0.20)		(-0.04)
$PC1_t$	0.011	0.062***	-0.021	0.031	-0.032	0.026	-0.012	0.022
	(0.33)	(2.60)	(-0.75)	(1.37)	(-1.25)	(1.28)	(-0.47)	(1.14)
N	2667	2667	2369	2369	2247	2247	2135	2135
Adjusted R^2	0.192	0.191	0.179	0.178	0.247	0.244	0.266	0.265
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Circuit FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 11: Frequency Distribution of SCA Filings by Law Firm. The majority of SCA complaints in our sample are filed by 20 "frequent filer" law firms. This table displays the filing frequency counts, as well as firm-specific fractions of the total SCA filing sample and cumulative percentages of the total SCA filing sample for the law firms associated with the majority of the filings by the SCAC.

No.	Law Firm Name	Freq.	Percent %	Cum. Percent %
1	Milberg Weiss Bershad & Schulman LLP	1197	6.80	6.80
2	Schiffrin & Barroway LLP	716	4.06	10.86
3	Law Offices of Charles J. Piven, P.A.	606	3.44	14.30
4	Wolf Haldenstein Adler Freeman & Herz LLP	586	3.33	17.63
5	Stull, Stull & Brody	549	3.12	20.75
6	The Rosen Law Firm P.A.	543	3.08	23.83
7	Coughlin Stoia Geller Rudman & Robbins LLP	540	3.07	26.90
8	Cauley Geller Bowman Coates & Rudman LLP	531	3.01	29.91
9	Bernstein Liebhard & Lifshitz, LLP	487	2.76	32.67
10	Pomerantz LLP	458	2.60	35.27
11	Robbins Geller Rudman & Dowd LLP	367	2.08	37.35
12	Scott & Scott LLC	335	1.9	39.25
13	Schatz & Nobel, P.C.	301	1.71	40.96
14	Glancy Binkow & Goldberg LLP	271	1.54	42.5
15	Brodsky & Smith, LLC	267	1.52	44.02
16	Weiss & Yourman	241	1.37	45.39
17	Faruqi & Faruqi LLP	212	1.2	46.59
18	Rigrodsky & Long, P.A.	207	1.18	47.77
19	Berger & Montague PC	205	1.16	48.93
20	Cohen Milstein Hausfeld & Toll PLLC	204	1.16	50.09

Table 12: Frequent Filers, SCA Outcomes, and Filing Complexity. This table reports regression results on the role of the most prolific SCA-filing law firms. The main independent variable is the natural log of one plus the number of the top 20 law firms by filing frequency in our sample, accounting for over 50% of the total SCAs filed, that are involved in the case. PC1 is the first principal component of the six filing readability measures: ARI, SI, FRE, FKG, GF, and Log(WordCount). Variable definitions are given in Appendix A. We control for the circuit, year-quarter, and industry fixed effects. t-statistics based on standard errors clustered at the year-quarter level are reported in parentheses. Statistical significance is denoted by *, **, and *** at the 10%, 5%, and 1% levels, respectively.

Panel A: Without Controls								
	Settled	PC1	ARI	SI	FRE	FKG	GF	Log(WordCount)
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$Log(1+\#Top20)_t$	0.719***	0.153*	0.164*	0.028	-0.733***	0.168**	0.122*	-0.003
	(5.79)	(1.87)	(1.86)	(0.57)	(-2.94)	(2.21)	(1.74)	(-0.16)
N	2777	2822	2822	2822	2822	2822	2822	2822
Adjusted R^2	0.069	0.044	0.044	0.026	0.014	0.061	0.108	0.026
Controls	No	No	No	No	No	No	No	No
Circuit FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Panel B: With Controls								
	Settled	PC1	ARI	SI	FRE	FKG	GF	Log(WordCount)
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$Log(1+\#Top20)_t$	0.633***	0.138*	0.128	0.050	-0.734***	0.146*	0.101	-0.021
	(6.11)	(1.69)	(1.43)	(1.07)	(-2.85)	(1.89)	(1.48)	(-0.97)
N	2658	2708	2708	2708	2708	2708	2708	2708
Adjusted R^2	0.104	0.061	0.064	0.058	0.034	0.079	0.125	0.179
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Circuit FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 13: Frequent Filer Law Firms, Post-Announcement Price Reaction, Machine Readability, and Institutional Ownership. This table reports the panel OLS regressions of Post-announcement price reaction on readability measures of the security class action lawsuit filings, institutional ownership measures, and firm-level control characteristics for the period of 1996-2018. Post-announcement price reaction is measured by Post-announcement drift and Post-announcement variance ratio. Post-announcement drift is defined as the cumulative return in excess of the value-weight market return over the event window of (2, N). Post-announcement variance ratio is defined as the variance of stock return in excess of the value-weighted market return over the event window of (2, N), scaled by the variance of stock return in excess of the value-weighted market return over the estimation window of (-120, -21). N ranges from 11, 61, 101, and 151. PC1 is the first principal component of the six readability measures: ARI, SI, FRE, FKG, GF, and Log(WordCount). Variable definitions are given in Appendix A. We also include the number of frequent-filer law firms responsible for the majority of SCA filings for each firm, as well as circuit, year-quarter, and industry fixed effects. t-statistics based on standard errors clustered at the year-quarter level are reported in parentheses. Statistical significanceis denoted by *, **, and *** at the 10%, 5%, and 1% levels, respectively.

Panel A: Post-Announcement Drift								
	Panel A: (2, 11)		(Panel B:2, 61)		(Panel C: 2, 101)		Panel D: (2, 151)	
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
LIO_{t-1}	-0.002		0.268***		0.278**		0.366**	
	(-0.05)		(2.80)		(2.37)		(2.47)	
LIO_{t-1}*$PC1_t$	0.007		0.012		0.026		0.030	
	(0.86)		(0.43)		(0.77)		(0.76)	
$DEDIO_{t-1}$		0.002		-0.051		-0.130		-0.313
		(0.05)		(-0.36)		(-0.76)		(-1.44)
$DEDIO_{t-1}$*$PC1_t$		-0.005		0.123**		0.144*		0.183**
		(-0.31)		(2.36)		(1.74)		(2.18)
$PC1_t$	-0.001	0.000	-0.005	-0.005	-0.007	-0.006	-0.008	-0.006
	(-0.50)	(0.16)	(-0.80)	(-1.32)	(-0.99)	(-1.15)	(-0.87)	(-1.06)
$Log(1 + \#Top20)_t$	0.005	0.005	-0.001	-0.001	0.002	0.003	-0.032	-0.031
	(0.70)	(0.68)	(-0.10)	(-0.05)	(0.12)	(0.19)	(-1.58)	(-1.53)
N	2761	2761	2403	2403	2279	2279	2165	2165
Adjusted R^2	0.035	0.035	0.069	0.067	0.102	0.101	0.122	0.121
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Circuit FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 13: Frequent Filer Law Firms, Post-Announcement Price Reaction, Machine Readability, and Institutional Ownership (cont'd).

Panel B: Post-Announcement Variance Ratio								
	Panel A: (2, 11)		Panel B: (2, 61)		Panel C: (2, 101)		Panel D: (2, 151)	
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
LIO_{t-1}	0.607		-0.346		-0.037		-0.009	
	(0.98)		(-0.44)		(-0.05)		(-0.01)	
LIO_{t-1}*$PC1_t$	0.393*		0.341**		0.397***		0.198	
	(1.95)		(2.19)		(3.15)		(1.42)	
$DEDIO_{t-1}$		-0.594		-0.063		0.039		0.040
		(-0.78)		(-0.06)		(0.05)		(0.04)
$DEDIO_{t-1}$*$PC1_t$		0.116		0.023		0.059		-0.014
		(0.37)		(0.05)		(0.16)		(-0.03)
$PC1_t$	0.007	0.061**	-0.014	0.033	-0.026	0.027	-0.005	0.022
	(0.19)	(2.33)	(-0.42)	(1.46)	(-0.90)	(1.47)	(-0.18)	(1.18)
$Log(1 + \#Top20)$	0.023	0.021	0.053	0.050	0.045	0.042	0.068	0.067
	(0.18)	(0.16)	(0.46)	(0.43)	(0.44)	(0.40)	(0.73)	(0.73)
N	2738	2738	2384	2384	2260	2260	2147	2147
Adjusted R^2	0.134	0.133	0.110	0.109	0.176	0.174	0.195	0.195
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Circuit FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Appendix A

We detail the construction of the variables used in our study below.

LIO: Fraction of shares owned by institutional investors that are long-term investors based on Gasper et al. (2005). Data source: CRSP, Thomson Reuters

DEDIO: Fraction of shares owned by institutional investors that are "Dedicated" investors based on Bushee (2001). Data source: CRSP, Thomson Reuters

Case Duration: The number of days between the filing date and the case resolution date. Data source: Stanford Securities Class Action Clearinghouse

GF: The Gunning FOG index is calculated in the following procedure. First, determine the average sentence length and divide the number of words by the number of sentences; Second, count the "complex" words consisting of three or more syllables. Then, add the average sentence length and the percentage of complex words. It is constructed as $0.4\big[(\frac{words}{sentences}) + 100(\frac{complex\ words}{words})\big]$. Data source: Stanford Securities Class Action Clearinghouse

ARI: The Automated Readability Index is calculated as $4.71 * (characters/words) + 0.5 * (words/sentences) - 21.43$, which outputs a number that approximates the grade level needed to comprehend the text. Data source: Stanford Securities Class Action Clearinghouse

SI: The Smog Index is calculated as $1.0430 \times \sqrt{number\ of\ pollysyllables \times 30\ number\ of\ sentences} + 3.1291$. Data source: Stanford Securities Class Action Clearinghouse

FRE: The Flesch Reading Ease index is calculated as $206.835 - 1.015 \times (\frac{total\ words}{total\ sentences}) - 84.6 \times (\frac{total\ syllables}{total\ words})$. Higher scores indicate that the case filings that are easier to read. Data source: Stanford Securities Class Action Clearinghouse

FKG: The Flesch-Kincaid Grade Level is calculated as $0.39 \times (\frac{total\ words}{total\ sentences}) + 11.8 \times (\frac{total\ syllables}{total\ words})$, which outputs a number that corresponds with a U.S. grade level. It represents the number of years of education generally required to understand this context of the case filing. Data source: Stanford Securities Class Action Clearinghouse

WordCount: The number of words in the text of Security Class Action filing, without punctuation. Data source: Stanford Securities Class Action Clearinghouse

FilingTime: The length of time from the security class end date to the filing date. Data source: Stanford Securities Class Action Clearinghouse

LegalComplexity: The legal complexity index is constructed as the total count of mentions of statistically significant class action complaint topics for case settlement identified by McShane et al. (2012). We extract these from the first identified complaint text using keyword matching. The significant security class action case topics are: (1) IPOs, (2) GAAP violations, (3) restatements of financials, (4) Rule 10b - 5, (5) Securities Act Section 11, (6) insider trading, (7) transactions (i.e., deals or mergers). Data source: Stanford Securities Class Action Clearinghouse

#Top20: The number of law firms from the top 20 by overall SCA filing frequency involved in the SCA filing. These 20 "frequent-filer" law firms represent a majority of all SCA filings in our sample. Data source: Stanford Securities Class Action Clearinghouse

Delaware: An indicator variable that equals one if the firm is headquartered, and therefore the case is tried, in Delaware. Data source: Stanford Securities Class Action Clearinghouse

EventDayRet: The cumulative stock return in excess of the value-weighted market return over the event window of (-1,1) around the filing date of the SCA complaint. Data source: Stanford Securities Class Action Clearinghouse, CRSP

Price: The log of stock prices. Data source: CRSP

Size: The log of market capitalization. Data source: CRSP, COMPUSTAT

BM: The ratio of book equity to market equity. Data source: CRSP, COMPUSTAT

Ret: Annualized stock returns. Data source: CRSP

Vol: Annualized standard deviation of monthly stock returns. Data source: CRSP

IO: Fraction of shares owned by institutional investors. Data source: CRSP, Thomson Reuters

PC1: The first principal component of the six readability measures ARI, SI, FRE, FKG, GF, and Log(WordCount). Data source: Stanford Securities Class Action Clearinghouse

Drift: The cumulative return in excess of the value-weighted market return over the event window of (2, N). N ranges from 11, 61, 101, and 151. Data source: CRSP

VR: The variance of stock return in excess of the value-weighted market return over the event window of (2, N), scaled by the variance of stock return in excess of the value-weighted market return over the estimation window of (-120, -21). N ranges from 11, 61, 101, and 151. Data source: CRSP

RQSPREAD: The difference between bid and ask prices divided by the midpoint of bid and ask prices. Data source: CRSP

Internet Appendix

This Appendix contains robustness checks to our main results using Tobit regression for left-truncated case duration, the importance of filing length as a measure of complexity, and alternative estimations of market reaction using equal-weighting.

Table A.1: Case Duration and Machine Readability Tobit Regression. This table reports the panel Tobit regressions of case duration on readability measures of the security class action lawsuit filings and firm-level control characteristics for the period of 1996-2018. Case duration is defined as the log of the number of days between the filing date and the status date. The readability measures include ARI, SI, FRE, FKG, GF, and Log(WordCount). PC1 is the first principal component of the six readability measures. Variable definitions are given in Appendix A. We control for the circuit, year-quarter, and industry fixed effects. t-statistics based on standard errors clustered at the year-quarter level are reported in parentheses. Statistical significance is denoted by *, **, and *** at the 10%, 5%, and 1% levels, respectively.

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
ARI_t	0.028***							0.025***						
	(3.02)							(2.78)						
SI_t		0.016							0.036**					
		(1.03)							(2.26)					
FRE_t			-0.003							-0.006*				
			(-1.11)							(-1.86)				
FKG_t				0.033***							0.030***			
				(2.86)							(2.71)			
GF_t					0.032**							0.024**		
					(2.55)							(2.04)		
$Log(WordCount)_t$						0.226***							0.205***	
						(4.69)							(3.17)	
$PC1_t$							0.026**							0.027***
							(2.55)							(2.65)
IO_{t-1}								-0.169**	-0.169**	-0.167**	-0.167**	-0.164**	-0.166**	-0.168**
								(-2.44)	(-2.44)	(-2.41)	(-2.41)	(-2.36)	(-2.40)	(-2.43)
$Price_{t-1}$								0.059***	0.060***	0.059***	0.059***	0.059***	0.059***	0.059***
								(3.01)	(3.02)	(2.99)	(3.00)	(3.01)	(2.96)	(3.00)
$Size_{t-1}$								0.063***	0.063***	0.063***	0.063***	0.062***	0.057***	0.063***
								(4.21)	(4.23)	(4.24)	(4.22)	(4.17)	(3.82)	(4.23)
BM_{t-1}								0.065***	0.064***	0.063***	0.063***	0.064***	0.057***	0.063***
								(3.08)	(3.09)	(3.10)	(3.06)	(3.07)	(2.85)	(3.07)
Ret_{t-1}								-0.265***	-0.264***	-0.263***	-0.266***	-0.266***	-0.240***	-0.265***
								(-4.92)	(-4.91)	(-4.86)	(-4.92)	(-4.93)	(-4.58)	(-4.92)
Vol_{t-1}								0.551***	0.551***	0.553***	0.556***	0.552***	0.526***	0.553***
								(3.52)	(3.52)	(3.52)	(3.54)	(3.51)	(3.36)	(3.53)
$Delaware_t$								-0.234	-0.241	-0.249	-0.235	-0.245	-0.245	-0.236
								(-1.34)	(-1.38)	(-1.42)	(-1.34)	(-1.40)	(-1.42)	(-1.35)
$Log(LegalComplexity)_t$								-0.063**	-0.065**	-0.062**	-0.062**	-0.058**	-0.092***	-0.064**
								(-2.44)	(-2.50)	(-2.41)	(-2.41)	(-2.24)	(-3.22)	(-2.47)
$Log(FilingTime)_t$								0.102***	0.105***	0.105***	0.103***	0.104***	0.092***	0.104***
								(7.52)	(7.52)	(7.54)	(7.52)	(7.52)	(6.72)	(7.52)
N	3558	3558	3558	3558	3558	3558	3558	2847	2847	2847	2847	2847	2847	2847
PseudoR^2	0.175	0.173	0.173	0.174	0.174	0.176	0.174	0.207	0.206	0.206	0.207	0.206	0.208	0.207
Circuit FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table A.2: Case Status and Machine Readability Controlling for Case Filing Length. This table reports the panel probit regressions of the case outcome indicator *Settled* on readability measures of the security class action lawsuit filings and firm-level control characteristics for the period of 1996-2018. *Settled* equals one if the security class action lawsuit case status is settled, and zero if it is dismissed. The readability measures include ARI, SI, FRE, FKG, and GF. PC1 is the first principal component of the five readability measures.Variable definitions are given in Appendix A. We control for circuit, year-quarter, and industry fixed effects. t-statistics based on standard errors clustered at the year-quarter level are reported in parentheses. Statistical significance is denoted by *, **, and *** at the 10%, 5%, and 1% levels, respectively.

Variable	(1)	(2)	(3)	(4)	(5)	(6)
ARI_t	0.042***					
	(2.95)					
SI_t		0.072***				
		(2.79)				
FRE_t			-0.014***			
			(-2.74)			
FKG_t				0.049***		
				(2.77)		
GF_t					0.039**	
					(2.21)	
$PC1_t$						0.046***
						(2.80)
$Log(WordCount)_t$	0.121	0.139*	0.150**	0.126*	0.156**	0.134*
	(1.60)	(1.83)	(2.02)	(1.67)	(2.10)	(1.79)
IO_{t-1}	-0.124	-0.126	-0.122	-0.119	-0.115	-0.122
	(-1.12)	(-1.14)	(-1.10)	(-1.08)	(-1.05)	(-1.11)
$Price_{t-1}$	0.082**	0.082**	0.081**	0.082**	0.082**	0.082**
	(2.14)	(2.16)	(2.12)	(2.14)	(2.15)	(2.14)
$Size_{t-1}$	-0.068***	-0.067***	-0.066***	-0.067***	-0.070***	-0.067***
	(-2.81)	(-2.78)	(-2.73)	(-2.78)	(-2.90)	(-2.77)
BM_{t-1}	0.046	0.044	0.043	0.044	0.044	0.045
	(0.88)	(0.85)	(0.83)	(0.85)	(0.85)	(0.86)
Ret_{t-1}	-0.142	-0.141	-0.139	-0.143	-0.139	-0.143
	(-1.46)	(-1.46)	(-1.43)	(-1.46)	(-1.42)	(-1.46)
Vol_{t-1}	0.314	0.307	0.312	0.321	0.309	0.316
	(1.25)	(1.23)	(1.24)	(1.27)	(1.24)	(1.26)
$Delaware_t$	-0.400	-0.402	-0.400	-0.402	-0.417	-0.400
	(-1.53)	(-1.54)	(-1.54)	(-1.54)	(-1.59)	(-1.53)
$Log(LegalComplexity)_t$	0.022	0.015	0.017	0.023	0.024	0.021
	(0.61)	(0.41)	(0.45)	(0.62)	(0.66)	(0.57)
$Log(FilingTime)_t$	-0.036**	-0.032*	-0.033*	-0.035**	-0.036**	-0.034*
	(-2.04)	(-1.82)	(-1.88)	(-1.97)	(-2.05)	(-1.93)
N	2596	2596	2596	2596	2596	2596
PseudoR^2	0.206	0.205	0.206	0.206	0.204	0.206
Circuit FE	Yes	Yes	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes

Table A.3: Case Duration and Machine Readability Controlling for Case Filing Length. This table reports the panel tobit regressions of case duration on readability measures of the security class action lawsuit filings and firm-level control characteristics for the period of 1996-2018. Case duration is defined as the log of the number of days between the filing date and the status date. The readability measures include ARI, SI, FRE, FKG, and GF. PC1 is the first principal component of the five readability measures. Variable definitions are given in Appendix A. We control for circuit, year-quarter, and industry fixed effects. t-statistics based on standard errors clustered at the year-quarter level are reported in parentheses. Statistical significance is denoted by *, **, and *** at the 10%, 5%, and 1% levels, respectively.

Variable	(1)	(2)	(3)	(4)	(5)	(6)
ARI_t	0.017*					
	(1.77)					
SI_t		0.023				
		(1.36)				
FRE_t			-0.004			
			(-1.30)			
FKG_t				0.021*		
				(1.78)		
GF_t					0.018	
					(1.49)	
$PC1_t$						0.017
						(1.60)
$Log(WordCount)_t$	0.178***	0.189***	0.194***	0.180***	0.192***	0.185***
	(2.77)	(2.99)	(3.09)	(2.81)	(3.08)	(2.92)
IO_{t-1}	-0.170**	-0.170**	-0.169**	-0.168**	-0.167**	-0.169**
	(-2.48)	(-2.48)	(-2.47)	(-2.45)	(-2.42)	(-2.47)
$Price_{t-1}$	0.059***	0.059***	0.059***	0.059***	0.059***	0.059***
	(2.98)	(2.98)	(2.97)	(2.98)	(2.98)	(2.98)
$Size_{t-1}$	0.058***	0.058***	0.058***	0.059***	0.058***	0.058***
	(3.97)	(3.97)	(3.97)	(3.98)	(3.91)	(3.98)
BM_{t-1}	0.059***	0.058***	0.057***	0.058***	0.058***	0.058***
	(2.83)	(2.83)	(2.83)	(2.82)	(2.82)	(2.82)
Ret_{t-1}	-0.246***	-0.244***	-0.243***	-0.246***	-0.245***	-0.245***
	(-4.66)	(-4.63)	(-4.60)	(-4.66)	(-4.66)	(-4.65)
Vol_{t-1}	0.532***	0.531***	0.532***	0.535***	0.532***	0.533***
	(3.38)	(3.38)	(3.38)	(3.40)	(3.38)	(3.39)
$Delaware_t$	-0.225	-0.231	-0.234	-0.226	-0.230	-0.228
	(-1.29)	(-1.32)	(-1.34)	(-1.29)	(-1.33)	(-1.30)
$Log(LegalComplexity)_t$	-0.092***	-0.094***	-0.094***	-0.091***	-0.090***	-0.092***
	(-3.19)	(-3.29)	(-3.27)	(-3.17)	(-3.12)	(-3.21)
$Log(FilingTime)_t$	0.092***	0.094***	0.093***	0.093***	0.092***	0.093***
	(6.78)	(6.72)	(6.77)	(6.78)	(6.77)	(6.77)
N	2847	2847	2847	2847	2847	2847
PseudoR^2	0.209	0.209	0.209	0.209	0.209	0.209
Circuit FE	Yes	Yes	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes

Table A.4: Two-Sample Tests of Price Reaction on Machine Readability and Institutional Ownership under Equal Weighting. This table presents the comparison of means of the filing date price reaction for two subsamples by medians of the principal component of the six machine readability measures PC1 and of institutional ownership levels by type. Post-announcement drift is defined as the cumulative return in excess of the equal-weight market return over the event window of (2, N). Post-announcement variance ratio is defined as the variance of stock return in excess of the equal-weighted market return over the event window of (2, N), scaled by the variance of stock return in excess of the equal-weighted market return over the estimation window of (-120, -21). N ranges from 11, 61, 101, and 151. All variables are defined in Appendix A. The table reports the means of price reaction of the two subsamples, the t-values, and the p-values of the differences in the two subsample means.

Readability (PC1)

	Drift				Abnormal Return Variance			
			Difference in Mean				Difference in Mean	
	More Readable	Less Readable	t-value	p-value	More Readable	Less Readable	t-value	p-value
(2, 11)	-0.019	-0.012	-1.59	0.0555	1.803	1.452	3.50	0.0002
(2, 61)	-0.018	-0.010	-0.68	0.2479	1.890	1.618	2.97	0.0015
(2, 101)	-0.033	-0.012	-1.47	0.0712	1.883	1.649	2.59	0.0048
(2, 151)	-0.050	-0.031	-1.09	0.1368	1.812	1.683	1.52	0.0644

LIO

	Drift				Abnormal Return Variance			
			Difference in Mean				Difference in Mean	
	Below Median	Above Median	t-value	p-value	Below Median	Above Median	t-value	p-value
(2, 11)	-0.028	-0.003	-5.13	0.0000	1.664	1.538	1.32	0.0934
(2, 61)	-0.055	0.006	-5.15	0.0000	1.646	1.699	-0.65	0.2583
(2, 101)	-0.062	-0.005	-3.83	0.0001	1.581	1.751	-2.16	0.0154
(2, 151)	-0.100	-0.007	-5.19	0.0000	1.526	1.721	-2.67	0.0038

DEDIO

	Drift				Abnormal Return Variance			
			Difference in Mean				Difference in Mean	
	Below Median	Above Median	t-value	p-value	Below Median	Above Median	t-value	p-value
(2, 11)	-0.024	-0.005	-3.84	0.0001	1.601	1.599	0.02	0.4914
(2, 61)	-0.037	-0.011	-2.14	0.0162	1.629	1.727	-1.19	0.1178
(2, 101)	-0.040	-0.029	-0.75	0.2272	1.559	1.796	-2.98	0.0014
(2, 151)	-0.069	-0.040	-1.64	0.0508	1.511	1.759	-3.36	0.0004

Table A.5: Post-Announcement Drift, Machine Readability, and Institutional Ownership. This table reports the panel OLS regressions of Post-announcement drift on readability measures of the security class action lawsuit filings, institutional ownership measures, and firm-level control characteristics for the period of 1996-2018. Post-announcement drift is defined as the cumulative return in excess of the equal-weighted market return over the event window of (2, N). N ranges from 11, 61, 101, and 151. PC1 is the first principal component of the six readability measures, ARI, SI, FRE, FKG, GF, and Log(WordCount). All variables are defined in Appendix A. We control for circuit, year-quarter, and industry fixed effects. t-statistics based on standard errors clustered at the year-quarter level are reported in parentheses. Statistical significance is denoted by *, **, and *** at the 10%, 5%, and 1% levels, respectively.

	Panel A: (2, 11)		Panel B: (2, 61)		Panel C: (2, 101)		Panel D: (2, 151)	
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
LIO_{t-1}	0.001		0.263***		0.276**		0.357**	
	(0.02)		(2.78)		(2.38)		(2.41)	
LIO_{t-1}*$PC1_t$	0.007		0.014		0.027		0.031	
	(0.77)		(0.49)		(0.82)		(0.81)	
$DEDIO_{t-1}$		0.006		-0.051		-0.137		-0.312
		(0.15)		(-0.37)		(-0.81)		(-1.43)
$DEDIO_{t-1}$*$PC1_t$		-0.003		0.125**		0.144*		0.181**
		(-0.20)		(2.45)		(1.80)		(2.15)
$PC1_t$	-0.001	0.000	-0.006	-0.005	-0.008	-0.006	-0.008	-0.006
	(-0.42)	(0.17)	(-0.88)	(-1.41)	(-1.07)	(-1.23)	(-0.94)	(-1.12)
N	2761	2761	2403	2403	2279	2279	2165	2165
Adjusted R^2	0.033	0.033	0.039	0.037	0.063	0.062	0.065	0.064
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Circuit FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table A.6: Post-Announcement Variance Ratio, Machine Readability, and Institutional Ownership. This table reports the panel OLS regressions of Post-announcement variance ratio on readability measures of the security class action lawsuit filings, institutional ownership measures, and firm-level control characteristics for the period of 1996-2018. Post-announcement variance ratio is defined as the variance of stock return in excess of the equal-weighted market return over the event window of (2, N), scaled by the variance of stock return in excess of the equal-weighted market return over the estimation window of (-120, -21). N ranges from 11, 61, 101, and 151. PC1 is the first principal component of the six readability measures, ARI, SI, FRE, FKG, GF, and Log(WordCount). All variables are defined in Appendix A. We control for circuit, year-quarter, and industry fixed effects. t-statistics based on standard errors clustered at the year-quarter level are reported in parentheses. Statistical significance is denoted by *, **, and *** at the 10%, 5%, and 1% levels, respectively.

	Panel A: (2, 11)		Panel B: (2, 61)		Panel C: (2, 101)		Panel D: (2, 151)	
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
LIO_{t-1}	0.648		-0.328		-0.075		-0.045	
	(1.04)		(-0.43)		(-0.11)		(-0.07)	
$LIO_{t-1}*PC1_t$	0.411**		0.322**		0.402***		0.203	
	(2.03)		(2.15)		(3.18)		(1.49)	
$DEDIO_{t-1}$		-0.599		-0.034		0.082		0.092
		(-0.79)		(-0.04)		(0.10)		(0.10)
$DEDIO_{t-1}*PC1_t$		0.090		-0.000		0.031		-0.041
		(0.28)		(-0.00)		(0.08)		(-0.08)
$PC1_t$	0.007	0.063**	-0.012	0.033	-0.026	0.028	-0.004	0.024
	(0.18)	(2.42)	(-0.37)	(1.54)	(-0.91)	(1.52)	(-0.15)	(1.34)
N	2738	2738	2384	2384	2260	2260	2147	2147
Adjusted R^2	0.134	0.133	0.117	0.116	0.175	0.173	0.195	0.194
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Circuit FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes