JOURNAL OF FINANCIAL AND QUANTITATIVE ANALYSIS

doi:10.1017/S0022109022001508

https://doi.org/10.1017/S0022109022001508 Published online by Cambridge University Press

Getting Back to the Source: A New Approach to Measuring Ex Ante Litigation Risk Using Plaintiff-Lawyer Views of SEC Filings

Antonis Kartapanis
Texas A&M University Mays Business School
akartapanis@mays.tamu.edu (corresponding author)

Christopher G. Yust
Texas A&M University Mays Business School
cyust@mays.tamu.edu

Abstract

This study introduces a new measure of ex ante litigation risk using scrutiny of SEC filings by the *source* of securities litigation (plaintiffs' lawyers) to reduce measurement error, relative to existing measures. We show that plaintiff-lawyer views proxy for the largely unobservable factors that make firms more likely to face litigation risk. Lagged views precede the public bad news revelation that triggers litigation and predicts future realized litigation risk (i.e., securities class actions filings and plaintiff-lawyer investigations) and stock market outcomes. Finally, we provide new insights into the plaintiff-lawyer case selection process that otherwise cannot be observed.

I. Introduction

Securities litigation is a major source of risk that "has the ability to influence essentially every aspect of the firm's operations" (Arena and Ferris (2016), p. 12). For example, litigation risk affects IPO underpricing, stock returns, investment, cost of capital, executive and director turnover, disclosure, financial reporting, ownership structure, and financial market development (e.g., La Porta, Lopez-de-Silanes, and Shleifer (2006), Fich and Shivdasani (2007), Gande and Lewis (2009), Hanley and Hoberg (2012), Deng, Willis, and Xu (2014), Arena and Julio (2015), and Crane and Koch (2018)). A fundamental challenge for researchers in this area is that they are primarily interested in the effect of unobservable ex ante litigation risk (i.e., the threat of litigation) on firm behavior but can only observe litigation existence ex post. Thus, researchers must use proxies for litigation risk, which are

We thank Anwer Ahmed, Matteo Arena (the referee), Jane Barton, Brant Christensen, Dain Donelson, Rachel Flam, Jennifer Glenn, John Griffin, Nicholas Hallman, Jarrad Harford (the editor), Brad Hepfer, Shane Johnson, Elena Karahanna, Burch Kealey, Sarah Marriott, Gonzalo Maturana, John McInnis, Brian Monsen, Paul Ordyna, Arun Rai, Jaime Schmidt, John Schomburger, Sorin Sorescu, Edward Swanson, and Connie Weaver for helpful comments and advice. We thank Triza Nganga and Taylor Paskett for their research assistance. We gratefully acknowledge the research support provided by the Mays Business School. All errors are our own.

https://doi.org/10.1017/S0022109022001508 Published online by Cambridge University Press

generally based on litigation filings. However, such approaches are confounded by the fact that firms may face significant litigation risk, despite not being sued. While around 2% of firms are sued in a given year (Kim and Skinner (2012)), significantly more face substantial litigation risk (Nelson and Pritchard (2016)). Compounding this problem, surprisingly little is known about how plaintiffs' lawyers, the ultimate source of litigation, allocate their limited time and resources to select firms to scrutinize and sue, which adds more measurement error into litigation risk estimates based on filings (Chalmers, Dann, and Harford (2002), Arena and Ferris (2016)). Simply put, more insight is needed on which firms face litigation risk and why, which we address in this article.

Most researchers proxy for securities litigation risk by estimating firm-specific litigation probability using models calibrated on filings or using indicator variables for industries with high litigation incidence. The most common model in recent research is from Kim and Skinner (2012) and uses lagged firm characteristics. In contrast, we introduce a measure focused on the *origins* of litigation risk by identifying firms whose SEC filings face scrutiny from plaintiffs' lawyers, regardless of whether these firms are later sued. Plaintiffs' lawyers are largely responsible for identifying potential litigation targets (Choi and Thompson (2006), Segal (2018)) and can be easily classified because litigation is concentrated among relatively few plaintiffs' law firms (Erichson (2007), Greene (2017)). We expect the use of plaintiff-lawyer scrutiny will significantly improve litigation risk estimates as it more holistically identifies the relatively persistent and unobservable quality of firms that make them i) more likely to commit future misconduct and ii) relatively good litigation targets. Lagged plaintiff-lawyer views predate the class period end for virtually all cases, so they are not mechanically driven by litigation-in-process and related bad news events.

Our tests exploit the plaintiff-lawyer need for public information to monitor firms. The Private Securities Litigation Reform Act (PSLRA) in 1995 created several procedural hurdles to discourage securities litigation. In particular, the discovery process, which allows plaintiffs' lawyers to obtain confidential firm information, was delayed until after the litigation survives a motion to dismiss. Consequently, plaintiffs' lawyers use public information, such as SEC filings (e.g., 10-Ks and 8-Ks), in their ongoing monitoring and case-building processes (Coffee (2006), Choi (2007)). Validating the importance of SEC filings, over 90% of filings in our sample explicitly indicate SEC filings are used to build the cases.

We start by identifying internet protocol addresses (IPs) registered to plaintiffs' law firms and use these IPs to identify viewed filings on the SEC Electronic Data Gathering, Analysis and Retrieval (EDGAR) system from 2012 to 2016. We then validate whether these EDGAR views are associated with plaintiff-lawyer scrutiny. Tests show a spike in views before and after the litigation filing date but only after the class period end, when the bad news that triggers litigation is first publicly revealed. We also show that plaintiff-lawyer views after the class period end, but before the filing date, predict future case outcomes. We allow the coefficient on views to vary between law firms with the top reputations and remaining firms, as the top lawyers have greater resources and expertise to identify and monitor firms with high litigation risk (Coe (2019)) and find results are driven by views from top lawyers.

After validating the measure, we test whether firm-year, plaintiff-lawyer views proxy for ex ante firm litigation risk. That is, can lagged scrutiny that predates bad news events predict firms with future realized litigation risk? The goal of these tests is not primarily to identify firms that will be sued but rather those who *expect* they may be sued because most research focuses on how litigation risk, rather than litigation itself, affects firm behavior.

To the best of our knowledge, prior research has only used filings to identify firms facing litigation risk. We instead more holistically identify firms with observable realized litigation risk, increasing our construct validity, by combining filings with plaintiff-lawyer investigations (i.e., *potential* filings). Corporations claim these investigations are associated with high litigation risk (Fisher (2015), Joyce (2019)). Validating this, tests show most sued firms also have an investigation in the same year. However, over half of the firm-years with investigations lack litigation filings. Thus, by combining investigations and filings, we more than double firm-years with observable realized litigation risk from 3.2% to 6.8%. This approach reduces false negatives and brings us closer to the expected litigation risk rate of around 10% (see Nelson and Pritchard (2016), resulting in better-specified and more powerful models.

Our tests benchmark the predictive ability of a model using only lagged plaintiff-lawyer views, firm size, and indicator variables for industries with the highest litigation risk (see Brochet and Srinivasan (2014)) against the model from Kim and Skinner (2012). We use the model from Kim and Skinner (2012) as an arbiter of whether plaintiff-lawyer views can significantly improve ex ante litigation risk estimates because it is the most common method in recent research (e.g., Hopkins, Maydew, and Venkatachalam (2015), Barzuza and Curtis (2017), Hong and Li (2019), An, Chen, Naiker, and Wang (2020), and Armstrong, Blackburne, and Quinn (2021)). However, we also find improved predictive ability relative to other approaches. We use model precision and sensitivity as our primary metrics to evaluate model predictive ability, rather than the commonly used area under the receiver operating characteristic (ROC) curve (AUC). While the AUC effectively summarizes model performance for balanced data sets, it is inappropriate for highly imbalanced data sets, such as ours (i.e., rare litigation incidence) because higher values can be due to performance in irrelevant regions for researchers (Saito and Rehmsmeier (2015), Brabec, Komárek, Franc, and Machlica (2020)).

Consistent with plaintiff-lawyer views being a proxy for persistent firm attributes that create litigation risk, they are relatively persistent, unlike other bad news events. Plaintiff-lawyer views are positively associated with future litigation filings. Using in-sample (out-of-sample) tests, our model improves precision and sensitivity, relative to that of Kim and Skinner (2012), by over 20% (30%). When predicting filings *or* plaintiff-lawyer investigations, our model improves precision and sensitivity by around 40% (50%). Our ability to better predict a more inclusive proxy for realized litigation risk further validates that plaintiff-lawyer scrutiny has greater construct validity for litigation risk. In sum, using scrutiny at the source of litigation better measures the (often unobservable) factors that make firms vulnerable to litigation *before* any misconduct is revealed. Meanwhile, our model imposes fewer data requirements, allowing researchers to examine the effects of litigation risk on samples nearly 50% larger, including understudied over-the-counter (OTC) firms.

https://doi.org/10.1017/S0022109022001508 Published online by Cambridge University Press

We conduct myriad tests to ensure plaintiff-lawyer views proxy for ex ante litigation risk, rather than contemporaneous litigation or related bad news events, such as using views from year $t-2$ or excluding firm years with other bad news events in year $t-1$. Results rule out these alternative explanations.

The use of plaintiff-lawyer views will lower measurement error and potential bias in common research designs, even when litigation risk is only used as a control variable (Roberts and Whited (2013). Moreover, because litigation risk can endogenously affect both firm actions (e.g., disclosure) and litigation, researchers often match samples of sued and nonsued firms on litigation risk (e.g., Atanasov, Ivanov, and Litvak (2012), Donelson, McInnis, Mergenthaler, and Yu (2012), and Brochet and Srinivasan (2014)). Despite attempts to hold litigation risk constant using existing methods, sued firms in these matched samples face significantly higher plaintiff-lawyer scrutiny in years before the litigation, which can bias research inferences (Meyer and Mittag (2017)).

We next test if plaintiff-lawyer views also predict more traditional finance topics. Because these views measure litigation risk that *predates* bad news at the class period end, which is associated with significant stock price drops (Karpoff, Lee, and Martin (2008), Gande and Lewis (2009)), we expect views may also predict future stock market outcomes. We find that quarterly plaintiff-lawyer views are negatively (positively) associated with next-quarter abnormal returns (return volatility). These tests are consistent with plaintiffs' lawyers possessing material, adverse information about firm fundamentals that are not yet recognized by the market. As a result, plaintiff-lawyer views likely have many applications for finance researchers.

We conduct numerous additional analyses, many of which are reported in the Supplementary Material for brevity. For example, the strong predictive ability of plaintiff-lawyer views raises the question of what factors affect such scrutiny, which otherwise is a black box. We thus examine their determinants. We note that, despite an extensive vector of adverse events and lagged firm characteristics, the model only explains a relatively small amount of the variation, demonstrating plaintiffs' lawyers integrate extensive unobservable or unexpected factors into their process. However, views are associated with several variables that provide noteworthy insights. For example, consistent with a disparity in law firm resources, there are differences in the determinants of scrutiny by the top and remaining plaintiffs' law firms. The latter use more relatively low-quality signals of potential case quality. Consistent with concerns that disclosure may attract plaintiff-lawyer scrutiny (Baginski, Hassell, and Kimbrough (2002), Rogers and Van Buskirk (2009)), both voluntary 8-Ks and earnings warnings are positively associated with plaintiff-lawyer views. Further, while firms with high litigation risk may cut dividends to maintain cash flexibility (Arena and Julio (2023)), this may exacerbate litigation risk because ceasing dividends is positively associated with plaintiff-lawyer views. We also create a measure of predicted views and show it predicts realized litigation risk better than existing predicted measures, both after and before our sample period, despite imposing fewer data requirements. Thus, researchers can use predicted views, similar to how many use the model from Kim and Skinner (2012), if actual views are unavailable.

This study contributes to the literature in five main ways. First, we create and share a new measure of ex ante litigation risk using plaintiff-lawyer scrutiny, which has improved the ability to predict realized litigation risk relative to other methods. Researchers can use it (or its predicted value when needed) as a measure of litigation risk to reduce i) measurement error and the risk of bias and ii) sample attrition due to data requirements. Second, we show plaintiff-lawyer views predict future stock market outcomes and may be useful as independent assessments of firm risk-taking and agency costs, similar to research using directors' and officers' (D&O) insurance data (e.g., Chalmers et al. (2002), Boyer and Stern (2014)). Third, we are the first to examine plaintiff-lawyer investigations, which are of interest in their own right. Fourth, we provide the first empirical evidence on the plaintiff-lawyer case selection process, including how it differs between the top and remaining law firms, and obtain new insights on litigation risk determinants. For example, we contribute to the literature on the relation between litigation risk and disclosure (e.g., Field, Lowry, and Shu (2005), Hanley and Hoberg (2012)) and corporate payout policy (Arena and Julio (2023)). Fifth, we reassess the validity of the AUC as a performance metric in highly imbalanced data sets and suggest alternative tests popularized by advances in machine learning. This methodology can improve prediction model assessment for other important rare events examined across interdisciplinary research (e.g., bankruptcy, fraud, and extreme stock returns).

II. Background and Prior Research

A. Securities Litigation Background

Firms criticized securities litigation for being costly and often nonmeritorious (Alexander (1991), House of Representatives (1995)), so Congress passed the PSLRA to make it harder for plaintiffs to build and sustain cases. While the law is widely regarded to have increased case merits (Johnson, Nelson, and Pritchard (2007), Choi, Nelson, and Pritchard (2009)), criticisms remain (Coffee (2006)). The PSLRA delayed the costly discovery process until a case survives a motion to dismiss and raised the pleading standard to require plaintiffs to identify specific misleading statements by the defendant that establish strong fraud inferences (Perino (2003)). As regulatory filings are the main public sources on which to build cases, most filings post-PSLRA include accounting allegations (Coffee (2006), Choi (2007)). The law also presumes the best lead plaintiff is the largest claimholder, so institutional investors are often the lead plaintiffs, resulting in close relationships with plaintiffs' lawyers (Cox, Thomas, and Kiku (2006), Donelson, Hopkins, and Yust (2018)).

The plaintiff-lawyer role is uniquely important. Unlike the traditional legal setting in which plaintiffs solicit lawyers to pursue cases, in securities litigation, the roles are largely reversed (Cox et al. (2006)). Most institutional investors rely on plaintiffs' lawyers to monitor their portfolio holdings on an ongoing basis and alert them of any misconduct. In return, the lawyers expect to be selected as class counsel if the institutional investors sue (Segal (2018)). Notably, the lawyers use contingency fee arrangements, so they incur all the costs if a case is not filed or is dismissed (Baker, Perino, and Silver (2015), Donelson et al. (2018)). This creates strong incentives for plaintiffs' lawyers to accurately identify and monitor firms with high litigation risks.

Plaintiffs' lawyers also increasingly announce investigations for *potential* litigation through newswire services or advertisements, such as those in Appendix B, to attract plaintiffs. This process has been criticized for undermining the PSLRA reforms (Berger and Gans (2003)) and increasing corporate litigation risk (Fisher (2015), Hudson and Cummins (2016), Joyce (2019)). Finally, while all plaintiffs' law firms create litigation risk, some present greater risks than others (Badawi and Webber (2015), BTI Consulting (2018)). For example, corporations state that the top law firms bring more sophisticated cases and "have endless teams of excellent people who all have the same killer instinct" (Coe (2019)).

B. Securities Litigation Research

Litigation imposes significant costs on firms (e.g., reputational damages, lost managerial time, and defense and settlement costs) and the economy (Kim and Skinner (2012), Arena and Ferris (2016). Zingales ((2007), p. 19) states the significant increase in litigation risk over time, which is apparent from total class action settlements increasing from \$150 million in 1997 to \$9.7 billion in 2005, is "the most likely explanation" for the decline in the U.S. share of global IPOs and the huge increase in 144A registrations. Litigation filings and settlement amounts remain near all-time highs (Cornerstone Research (2020)). Executives take actions to reduce litigation risk, such as changing disclosures (Skinner (1997), Field et al. (2005)), but findings are mixed on their effectiveness. Thus, while litigation has always been present in the corporate environment, it is now regarded as a major source of risk affecting virtually every facet of a firm's operations (Arena and Ferris (2016)). Unsurprisingly then, it has been shown to affect numerous topics of interest to researchers (Lowry and Shu (2002), Fich and Shivdasani (2007), Deng et al. (2014), Arena and Julio (2015), and Crane and Koch (2018)). Arena and Ferris ((2016), p. 1) conclude that "as litigation risk and costs have increased, the study of their effects on firm value and activities has become more critical than ever."

Nonetheless, little is known about how plaintiffs' lawyers identify and select cases. Early research mainly proxied for litigation risk with an indicator variable for high litigation industries (e.g., Francis, Philbrick, and Schipper (1994)), but Kim and Skinner (2012) show that a model using lagged firm characteristics better proxies for litigation risk. Most papers now use the estimated litigation probability from Kim and Skinner (2012) or similar methods (e.g., Brocher and Srinivasan (2014), Banerjee, Humphery-Jenner, Nanda, and Tham (2018), Hong and Li (2019), Bonaimé, Harford, and Moore (2020), Arena and Julio (2023), and Freund, Nguyen, and Phan (2023)), but some still use only an indicator variable for high litigation industries or other approaches (e.g., Callen and Fang (2015), Bird, Karolyi, and Ruchti (2019), Huang, Hui, and Li (2019), and Hutton, Shu, and Zheng (2022)). However, not all firms that face high litigation risk are sued due to plaintiff-lawyer financial constraints (Landeo and Nikitin (2018)) or an excess of viable cases (Donelson, Kartapanis, and Yust (2021b)). Thus, while the litigation rate observed by Kim and Skinner (2012) is only 2%, Nelson and Pritchard (2016) estimate 5 times as many face substantive litigation risk. As a result, there is significant measurement error in litigation risk measures based on filings (Chalmers et al. (2002)). For this reason, Arena and Ferris ((2016), p. 6) note a superior proxy for litigation risk is D&O insurance data because "it is able to capture litigation risk for firms that successfully avoid litigation even though they might be highly exposed to the risk of a lawsuit." However, while such data is not publicly available for U.S. firms, plaintiff-lawyer views are.

III. Construction and Data Validation

A. Overview

We use three primary data sources to construct our plaintiff-lawyer scrutiny measure for 2012 to 2016. While plaintiff-lawyer views are available in earlier years, to maintain a constant sample period, we start in 2012 when plaintiff-lawyer investigation data, discussed in Section IV.A.2, becomes widely populated. The sample ends in 2016 due to current data availability. Each source is discussed below.

B. Plaintiff-Lawyer IP Addresses

We first obtain ARIN's Bulk Whois data to identify IP address owners and the registration date. Given the volume of law firms and difficulty matching, we focus on matching the most active plaintiffs' law firms using keywords based on their names with manual confirmation. We discuss this in detail and report the most active plaintiffs' law firms in our period in Table A.1 in the Supplementary Material. Focusing on the most active law firms captures most plaintiff-lawyer scrutiny, owing to the concentrated nature of the plaintiffs' bar (Greene (2017)). We identify 118 registered IP addresses (or blocks) related to 40 unique law firms.

C. Securities Class Actions

Securities class actions filed in our sample are collected from Stanford's Securities Class Action Clearinghouse. We exclude 31 cases that cannot be linked to Compustat (via name and ticker) or for which Compustat lacks CIK information that we use to link firms to EDGAR data, resulting in 932 initial cases. However, as our objective in the data validation section is to show that the identified IPs belong to plaintiffs' lawyers and are used in building a case, these analyses exclude 212 cases for which there is no valid IP for the plaintiffs' lawyers. Panel A of Table A.2 in the Supplementary Material summarizes the sample construction.

D. Views of SEC Filings on EDGAR

We obtain daily views of SEC filings from EDGAR log files provided by the SEC. The log files provide the visitor's IP address (masking the fourth octet, i.e., the last 3 digits), the filer's Central Index Key (CIK), and the filing viewed.[1] Rather than examining only cases in which plaintiffs' lawyers register a full block of IPs, we follow Chen, Cohen, Gurun, Lou, and Malloy (2020) to unmask the fourth octet

[1]Tests exclude views of index pages, web crawlers, and 3 days of views for a single IP address with unusual activity (i.e., more than 3,000 views on a single day).

and match IP addresses to those identified for plaintiffs' lawyers. While plaintiffs' lawyers undoubtedly also acquire information from other sources, SEC filings contain significant detailed firm-specific information, which is critical for ongoing monitoring and to identify misleading statements that provide strong inference of fraud (see Choi (2007)). Further, as noted by Bernard, Blackburne, and Thomas ((2020), p. 762), "few other sources are free, easily accessible, complete, and offer relatively anonymous access." To validate our measure, we examine whether the views follow patterns similar to what one would expect around key litigation dates.

E. Validating Plaintiff-Lawyer Views

Using plaintiff-lawyer EDGAR views to proxy for litigation risk rests on our supposition that plaintiffs' lawyers generally view SEC filings through EDGAR and use this information for ongoing monitoring. We believe this is reasonable because EDGAR is a centralized way to access filings for all companies and is publicly anonymous in real-time, in contrast to firm investor relations websites, where firms monitor web traffic (Hodge and Pronk (2006)).

Nonetheless, to test this supposition, we test if plaintiffs' lawyers reference SEC filings in the initial litigation complaints. EDGAR is the primary source for these filings, so such references imply plaintiffs' lawyers likely accessed it before the filing date to build their case. We first search for the paragraph in which the plaintiffs (and their lawyers) state sources used to build the case and see if they reference SEC filings. From the 634 complaints for which we can identify a sources paragraph, 545 (86%) reference SEC filings. For complaints that do not reference SEC filings or lack a sources paragraph, we search the entire complaint for SEC filing references and find them in 132 of these cases (75%).[2] Thus 94% of the complaints either explicitly state SEC filings are used to build the case or reference an SEC filing.

Panel A of Table 1 presents descriptive statistics on plaintiff-lawyer views in the 20-day window leading up to, but not including, the filing date. There is at least one view from plaintiffs' lawyers in 51% of filings (VIEWS $>$ 0_DUMMY). There are 10.10 average views (TOT_VIEWS), which increases to 19.97 when examining cases with at least one view (TOT_VIEWS_IF $>$ 0). These variables are highly skewed, however, as the interquartile range for TOT_VIEWS_IF $>$ 0 is from 3 to 23 views with a median of 8. Thus, we use the natural log of plaintiff-lawyer views in subsequent analyses, similar to prior research using EDGAR views (e.g., Drake, Johnson, Roulstone, and Thornock (2020), Gibbons, Iliev, and Kalodimos (2021), and Iliev, Kalodimos, and Lowry (2021)). Another benefit of this logarithmic transformation is that additional plaintiff-lawyer views should matter more if a firm has little existing plaintiff-lawyer scrutiny than if it already has extensive scrutiny, similar to the logic motivating the use of logged analyst coverage in Hong, Lim, and Stein (2000).

[2]We identify source paragraphs referencing SEC filings by searching for "Securities and Exchange Commission," "SEC," "regulatory filings," and "public filings." We identify references to SEC filings outside source paragraphs by searching for "10-K," "10-Q," "8-K," "SEC filing," "Form 20-F," and "filed with the SEC."

TABLE 1

Descriptive Statistics of Plaintiff-Lawyer EDGAR Views

Table 1 presents descriptive statistics of plaintiff-lawyer EDGAR views over the 20-day window leading up to the litigation filing day (i.e., $t-20$ to $t-1$) from 2012 to 2016 for the case-level sample shown in Panel A of Table A.2 in the Supplementary Material. Panel A examines all filings, Panel B splits the securities class actions into those that eventually settle versus those that are dismissed, and Panel C splits the securities class actions into those that contain accounting allegations alleging Rule 10b-5 violations versus those that do not. In Panels B and C, *t*-tests are used to test for significant differences between the means of different groups. Refer to Appendix A for variable definitions.

Panel A. All Securities Class Actions

	N	Mean	P25	P50	P75	Std. Dev.
VIEWS > 0_DUMMY	720	0.51	0.0	1.0	1.0	0.50
TOT_VIEWS	720	10.10	0.0	1.0	8.0	25.00
TOT_UNQ_VIEW_FIRMS	720	0.77	0.0	1.0	1.0	1.04
TOT_VIEWS_IF > 0	364	19.97	3.0	8.0	23.0	32.25
TOT_UNQ_VIEW_FIRMS_IF_VIEWS > 0	364	1.53	1.0	1.0	2.0	1.00

Panel B. Settled Versus Dismissed Securities Class Actions

	Settled						Dismissed						Diff.	*t*-Stat.
	N	Mean	P25	P50	P75	Std. Dev.	*N*	Mean	P25	P50	P75	Std. Dev.		
VIEWS > 0_DUMMY	307	0.51	0.00	1.00	1.00	0.50	385	0.50	0.00	1.00	1.00	0.50	0.01	0.20
TOT_VIEWS	307	10.95	0.00	1.00	11.00	21.69	385	9.69	0.00	1.00	7.00	27.65	1.26	0.66
TOT_UNQ_VIEW_FIRMS	307	0.88	0.00	1.00	1.00	1.24	385	0.69	0.00	1.00	1.00	0.85	0.19	2.37
TOT_VIEWS_IF > 0	157	21.41	3.00	11.00	30.00	26.40	194	19.22	2.00	7.00	21.00	36.57	2.19	0.63
TOT_UNQ_VIEW_FIRMS_ IF_VIEWS > 0	157	1.72	1.00	1.00	2.00	1.25	194	1.37	1.00	1.00	2.00	0.70	0.35	3.29

Panel C. Accounting 10b-5 Versus Remaining Securities Class Actions

	Accounting 10b-5						Nonaccounting 10b-5						Diff.	*t*-Stat.
	N	Mean	P25	P50	P75	Std. Dev.	*N*	Mean	P25	P50	P75	Std. Dev.		
VIEWS > 0_DUMMY	225	0.48	0.00	0.00	1.00	0.50	495	0.52	0.00	1.00	1.00	0.50	−0.04	−0.76
TOT_VIEWS	225	12.56	0.00	0.00	8.00	33.31	495	8.98	0.00	1.00	8.00	20.06	3.58	1.79
TOT_UNQ_VIEW_FIRMS	225	0.80	0.00	0.00	1.00	1.15	495	0.76	0.00	1.00	1.00	0.99	0.04	0.46
TOT_VIEWS_IF > 0	109	25.94	2.00	9.00	32.00	44.18	255	17.42	3.00	8.00	21.00	25.19	8.52	2.32
TOT_UNQ_VIEW_FIRMS_ IF_VIEWS > 0	109	1.65	1.00	1.00	2.00	1.14	255	1.48	1.00	1.00	2.00	0.93	0.17	1.52

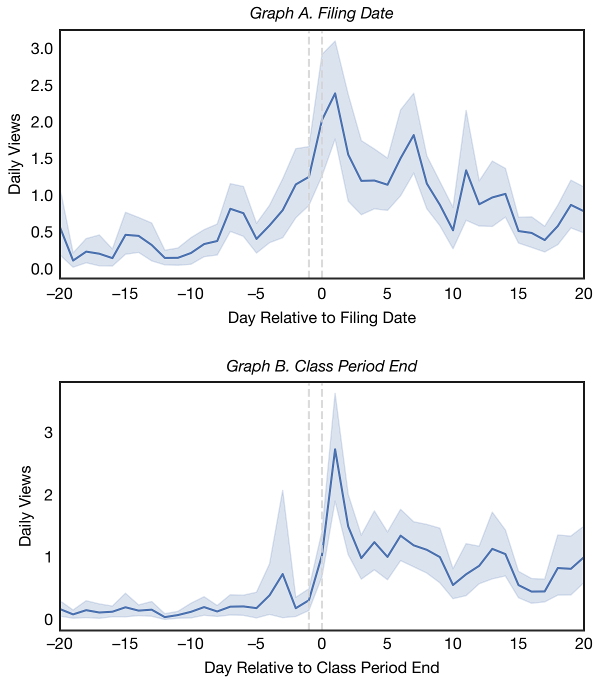
Given the criticism that plaintiffs' lawyers bring cases with little regard to their merits (e.g., *Wall Street Journal* (2014)), we also provide initial evidence on whether their scrutiny differs by case merits to provide insight into their case selection. Greater use of SEC filings in the case-building process by plaintiffs' lawyers likely increases the probability that they can identify specific misleading statements to meet the revised pleading standards and allege accounting violations, which are associated with case merits (Pritchard and Sale (2005), Johnson et al. (2007)). Thus we expect higher plaintiff-lawyer views for more meritorious cases. Accordingly, we partition our sample based on i) whether the case settles (see Dyck, Morse, and Zingales (2010), Donelson, Kartapanis, McInnis, and Yust (2021a)) and ii) whether it alleges accounting fraud.

Panel B of Table 1 shows descriptive statistics after splitting the sample into cases that settle versus are dismissed, so ongoing cases as of June 2022 are omitted from this analysis. Panel C splits the sample into accounting-related cases alleging misstatements or material omissions of information with fraudulent intent (i.e., Rule 10b-5 violations) versus remaining cases. Total plaintiff-lawyer views for settled cases are insignificantly higher, but when we examine the number of unique law firms viewing, we find more law firms view SEC filings for settled cases

FIGURE 1

Daily Views Around Relevant Litigation Dates

Graphs A and B of Figure 1 present the average number of daily views in the 20 days before and after the filing date (Graph A) and class period end (Graph B) as per the first identified securities complaint (FIC). The vertical gray lines are at days −1 and 0. The shaded areas present a 95% confidence interval calculated using bootstrap resampling. Refer to Panel A of Table A.2 in the Supplementary Material for the case-level sample composition.



($p < 0.01$). Further, there are more plaintiff-lawyer views for accounting fraud cases ($p < 0.1$). Thus, tests find some evidence that plaintiff-lawyer views in this window are associated with case merits, which is explored more formally in the following section.

We next examine daily view patterns. Figure 1 shows daily views for the 20 days before and after the i) filing date and ii) class period end. Consistent with plaintiffs' lawyers using SEC filings to build the case, Graph A shows a gradual increase in views in the days *before* the filing, which spike on the filing date and the following day and decrease thereafter. In contrast, before the class period end (i.e., before misconduct is publicly revealed), there is less public firm-specific news on which to build a case. Consistent with this, Graph B shows limited views in the days preceding the announcement. However, upon the initial release of the bad news, there is generally a significant stock price correction from its revelation (Karpoff et al. (2008), Gande and Lewis (2009), and Dyck et al. (2010)). On this date, we observe a spike in views that decreases over the next 20 days. This increase in views is consistent with plaintiffs' lawyers evaluating the bad news to decide whether to sue.[3]

[3]We further examine the data in Section I of the Supplementary Material to provide more insight, such as distinguishing between views from plaintiffs' lawyers that do and do not participate in the filings,

https://doi.org/10.1017/S0022109022001508 Published online by Cambridge University Press

https://doi.org/10.1017/S0022109022001508 Published online by Cambridge University Press

F. Predictive Ability for Case Outcomes

The prior analyses show our measure identifies plaintiff-lawyer scrutiny, but a potential concern with using it for predictive analyses is that we cannot obtain IPs for all plaintiffs' law firms. We believe this will not significantly harm its predictive ability as sued firms often receive scrutiny from multiple law firms. Thus, unless case selection differs systematically for the plaintiffs' lawyers whose IPs we can obtain, the measure should capture *general* plaintiff-lawyer scrutiny. However, before examining the general predictive ability of plaintiff-lawyer views, we first validate their predictive ability at the case-level over the relatively short period from the class period end to just before the filing date. The starting population for this test is the 932 cases from the start of Table A.2 in the Supplementary Material, before requiring data for control variables and excluding ongoing cases, rather than the 720 cases used in prior analyses.

In this analysis, we exploit the ability for plaintiff-lawyer views *after* the class period end to measure their appraisal of whether litigation is warranted and preparing the filing itself. Building on Table 1 and the related Figures A.2 and A.3 in the Supplementary Material, which find more evidence meritorious cases have higher plaintiff-lawyer views, we expect cases with greater views from the class period end to the filing date to be more meritorious (i.e., more likely to settle and settle for larger amounts). Such evidence also contributes to the extensive legal literature on developing better models to predict settlement outcomes (e.g., Cox et al. (2006), Choi (2007), and Donelson, Hopkins, and Yust (2015)). We estimate the following regression:

$$
\begin{aligned}
\text{(1)}\quad \text{SETTLED}/\ln(\text{SETTLEM}) = {} & \beta_0 + \beta_1 \ln\left(\text{VIEWS}_{[\text{Class End,Filing–1}]}\right) \\
& + \beta_2 \ln(\text{DAMAGES}) + \beta_3 \ln(\text{MVE}) + \beta_4 \text{ROA} \\
& + \beta_5 \text{TOBINS\_Q} + \beta_6 \text{INSTIT\_OWN}. \\
& + \text{Filing Year FE} + \text{Circuit FE} + \varepsilon,
\end{aligned}
$$

where $\text{SETTLED}_{i,t}$ is an indicator variable set to 1 if the case settles and $\ln(\text{SETTLEM}_{i,t})$ is the natural log of the settlement amount. The variable of interest, $\ln(\text{VIEWS}_{i,[\text{Class End,Filing-1}]})$, is the natural log of one plus plaintiff-lawyer views over this period. If these views predict case outcomes associated with merits, β_1 should be positive. We control for variables associated with merits and case outcomes: maximum damages ($\ln(\text{DAMAGES}_{i,t})$), firm size ($\ln(\text{MVE}_{i,t})$), profitability ($\text{ROA}_{i,t}$), Tobin's Q ($\text{TOBINS\_Q}_{i,t}$), and institutional ownership ($\text{INSTIT\_OWN}_{i,t}$) before the litigation (Cox et al. (2006), Donelson et al. (2015)). The equation includes filing year and circuit fixed effects due to potential time trends and different legal standards across federal circuits (Pritchard and Sale (2005)). See Appendix A for detailed variable definitions.

Column 1 (2) of Table 2 presents results estimating $\text{SETTLED}_{i,t}$ in equation (1) using a logistic (ordinary least squares (OLS)) regression. We use both logistic and OLS models due to potential bias issues with small samples

---

separately examining views using the case partitions from Table 1, and disaggregating views based on the filing type.

TABLE 2

Can Plaintiff-Lawyer Views Predict Case Outcomes?

Table 2 presents results from estimating equation (1) examining whether plaintiff-lawyer EDGAR views starting on the class period end date and ending the day prior to the filing date can predict outcomes associated with case merits (i.e., the case settles in columns 1–2 and has larger settlement amounts in column 3) using the full population of securities class actions with available data from 2012 to 2016. That is, the starting population for this test is the 932 cases shown in Panel A of Table A.2 in the Supplementary Material before requiring data for control variables, excluding ongoing cases given the examination of case outcomes. Column 1 (2–3) uses a logistic (OLS) regression. Standard errors appear in parentheses and are adjusted for clustering at the company level. ***, **, * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. To minimize the influence of outliers, all continuous variables are winsorized at the 1% and 99% levels. Standard errors appear in parentheses and are adjusted for clustering at the company level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. Refer to Appendix A for variable definitions.

	$SETTLED_{i,t}$	$SETTLED_{i,t}$	$\ln(SETTLEM_{i,t})$
	1	2	3
$\ln(VIEWS_{i,[\text{Class End,Filing-1}]})$	0.13**	0.03***	0.17***
	(0.05)	(0.01)	(0.04)
$\ln(DAMAGES_{i,t})$	0.06*	0.01**	0.13***
	(0.03)	(0.01)	(0.03)
$\ln(MVE_{i,t})$	−0.07	−0.02	0.41***
	(0.05)	(0.01)	(0.06)
$ROA_{i,t}$	0.00	−0.00	−0.52***
	(0.30)	(0.07)	(0.19)
$TOBINS\_Q_{i,t}$	0.00	0.00	−0.00
	(0.03)	(0.01)	(0.02)
$INSTIT\_OWN_{i,t}$	−0.26	−0.06	0.39*
	(0.24)	(0.06)	(0.22)
INTERCEPT	−0.71	0.36	9.86***
	(1.29)	(0.30)	(0.58)
Filing year FE	Yes	Yes	Yes
Circuit FE	Yes	Yes	Yes
No. of obs.	768	768	317
AUC	0.656		
Pseudo R^2	0.051		
R^2		0.066	0.527

and fixed effects (see Greene (2004), Angrist and Pischke (2009)). The coefficient of $\ln(VIEWS_{i,[\text{Class End,Filing-1}]})$ is positive and significant under each specification ($p < 0.05$). Column 3 presents the results of estimating $\ln(SETTLEM_{i,t})$ in equation (1) using OLS in the subsample of settled cases for which we can obtain settlement amounts, and $\ln(VIEWS_{i,[\text{Class End,Filing-1}]})$ is positive and significant ($p < 0.01$). Collectively, these results validate the predictive ability of plaintiff-lawyer views, particularly because the average case requires over 3 years to settle (Cornerstone Research (2021)).

The prior analyses treat views from *all* plaintiffs' lawyers as having the same importance. However, because the top plaintiffs' law firms pose greater litigation risks for defendants, we expect they may have differential predictive ability. We identify top plaintiffs' law firms using Chambers and Partners (2021), an independent research company that ranks firms based on tens of thousands of interviews a year with plaintiff-lawyer clients and third-party experts and separate views into those from the top ($TOP\_PLF\_LN\_VIEWS_{i,[\text{Class End,Filing-1}]}$) and remaining ($REM\_PLF\_LN\_VIEWS_{i,[\text{Class End,Filing-1}]}$) law firms and repeat the prior analyses. We use rankings from June 2011 to avoid a mechanical relation between their classification and case success in our sample. The point of this analysis is not whether top plaintiffs' lawyers cause better case outcomes but whether market participants can predict stronger future case outcomes based on the nature of this scrutiny around the case.

https://doi.org/10.1017/S0022109022001508 Published online by Cambridge University Press

We expect views from top plaintiffs' lawyers to be more positively associated with the probability a case settles. First, top lawyers are relatively more judicious in selecting high-quality cases, so increased scrutiny of a firm by such lawyers may indicate a case that is stronger and more likely to settle (Badawi and Webber (2015)). Relatedly, institutional investors as lead plaintiffs are associated with meritorious cases and may be more likely to select top lawyers (Cox et al. (2006)). Second, top lawyers have more experience and resources, so they can build stronger cases. Thus, higher scrutiny by top lawyers may indicate case quality has been enhanced through their involvement (Abramowicz (2004), Badawi and Webber (2015)). However, conditional on a case surviving a motion to dismiss, views by all plaintiffs' lawyers are likely positively associated with settlement amounts. Greater use of SEC filings increases the probability lawyers can identify specific factual misstatements, such as accounting violations, which are associated with settlement amounts (Donelson et al. (2015)). That said, top plaintiff-lawyer views may also have a stronger positive relation with settlement amounts because their expertise may allow them to use information from SEC filings to extract more value from defendants in settlement negotiations.

Table 3 shows the results. Only views from top plaintiffs' lawyers predict settlement outcomes ($p < 0.1$) in columns 1 and 2, consistent with these lawyers having a stronger association with case merits. Similarly, only views from top plaintiffs' lawyers predict settlement amounts in column 3 ($p < 0.01$). The coefficients on $\text{TOP\_PLF\_LN\_VIEWS}_{i,[\text{Class End,Filing-1}]}$ have larger magnitudes as expected but are insignificantly different from that of $\text{REM\_PLF\_LN\_VIEWS}_{i,[\text{Class End,Filing-1}]}$ ($p = 0.66$ in columns 1 and 2 and $p = 0.18$ in column 3, untabulated), potentially due to the small sample. However, given this evidence of differential predictive ability, we distinguish between views from top and remaining law firms in most subsequent tests.

IV. Results

A. Proxying for Ex Ante Litigation Risk

1. Sample Construction

Given the data validation, we test whether plaintiff-lawyer views at the firm-year level can proxy for ex ante litigation risk using firms with available data from 2012 to 2016. That is, can plaintiff-lawyer views that predate the litigation process and related bad news events nonetheless predict which firms face future litigation? We use this sample period for the reason noted in Section III.A but find similar results using the 2008–2016 time period in Section II of the Supplementary Material. We exclude firm-years missing data required by Kim and Skinner (2012) and our later determinants tests and missing CIKs to link firm-years to EDGAR. Our final sample contains 17,179 firm-year observations as shown in Panel B of Table A.2 in the Supplementary Material.

In addition to Compustat, CRSP, and the data sources described in Section III, we obtain data from Audit Analytics, directEDGAR, EDGAR daily indexes, IBES,

TABLE 3

Are Plaintiff-Lawyer Views Differentially Associated with Predicting Case Outcomes for Top Versus Remaining Plaintiffs' Lawyers?

Table 3 presents results from estimating equation (1), after replacing ln(VIEWS$_{i,\text{[ClassEnd,Filing-1]}}$) with TOP_PLF_LN_VIEWS$_{i,\text{[ClassEnd,Filing-1]}}$ and REM_PLF_LN_VIEWS$_{i,\text{[ClassEnd,Filing-1]}}$, examining whether EDGAR views from the top and remaining plaintiffs' law firms starting on the class period end date and ending the day prior to the filing date can predict outcomes associated with case merits (i.e., the case settles in columns 1–2 and has larger settlement amounts in column 3) using the full population of securities class actions with available data from 2012 to 2016. That is, the starting population for this test is the 932 cases shown in Panel A of Table A.2 in the Supplementary Material before requiring data for control variables, excluding ongoing cases given the examination of case outcomes. Columns 1 (2–3) use a logistic (OLS) regression. To minimize the influence of outliers, all continuous variables are winsorized at the 1% and 99% levels. Standard errors appear in parentheses and are adjusted for clustering at the company level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. Refer to Appendix A for variable definitions.

	SETTLED$_{i,t}$	SETTLED$_{i,t}$	ln(SETTLEM$_{i,t}$)
	1	2	3
TOP_PLF_LN_VIEWS$_{i,\text{[ClassEnd,Filing-1]}}$	0.11*	0.03*	0.17***
	(0.06)	(0.01)	(0.05)
REM_PLF_LN_VIEWS$_{i,\text{[ClassEnd,Filing-1]}}$	0.07	0.02	0.08
	(0.06)	(0.01)	(0.05)
ln(DAMAGES$_{i,t}$)	0.06*	0.01**	0.13***
	(0.03)	(0.01)	(0.03)
ln(MVE$_{i,t}$)	−0.07	−0.02	0.41***
	(0.05)	(0.01)	(0.06)
ROA$_{i,t}$	−0.01	−0.00	−0.52***
	(0.30)	(0.07)	(0.19)
TOBINS_Q$_{i,t}$	0.00	0.00	−0.00
	(0.03)	(0.01)	(0.02)
INSTIT_OWN$_{i,t}$	−0.24	−0.06	0.41*
	(0.24)	(0.06)	(0.22)
INTERCEPT	−0.56	0.40	10.35***
	(1.36)	(0.32)	(0.65)
Filing year FE	Yes	Yes	Yes
Circuit FE	Yes	Yes	Yes
No. of obs.	768	768	317
AUC	0.656		
Pseudo R^2	0.050		
R^2		0.065	0.531

Thomson Reuters, and non-GAAP adjustments from Bentley, Christensen, Gee, and Whipple (2018). We obtain historic headquarters and incorporation states from Bill McDonald's website and use Compustat for firm years missing this data. We use Factiva to identify plaintiff-lawyer investigation announcements.

Table 4 reports descriptive statistics (Panel A) and correlations (Panel B) for this sample. Around 3% of the firm-years are sued (SUED$_{i,t}$), similar to prior research, but over twice as many (6.8%) are either sued or face a plaintiff-lawyer investigation (SUED_INV$_{i,t}$). This percentage is closer to the estimated 10% of firms that face substantive litigation risk in Nelson and Pritchard (2016) and shows research that only uses filings to model litigation risk is misclassifying as low litigation risk over half the possible relevant sample. As this misclassification is unlikely to be random, it can bias coefficients toward or away from 0, depending on their correlation with the variable of interest (Meyer and Mittag (2017)). Around 44% of the firm years have at least one view by plaintiffs' lawyers (VIEWS $>$ 0_IND$_{i,t}$). Views from top plaintiffs' lawyers (TOP_PLF_VIEWS $>$ 0_IND$_{i,t}$) appear more targeted, as these lawyers scrutinize about 15% of the firm years versus 39% from their remaining peers (REM_PLF_VIEWS $>$ 0_IND$_{i,t}$). There

https://doi.org/10.1017/S0022109022001508 Published online by Cambridge University Press

TABLE 4

Descriptive Statistics and Correlations for Firm-Year Sample

Panel A of Table 4 reports descriptive statistics for the main variables used in the firm-year sample analysis, comprised of 17,179 firm-years from 2012 to 2016 as shown in Panel B of Table A.2 in the Supplementary Material. Panel B presents Pearson correlations. To minimize the influence of outliers, all continuous variables are winsorized at the 1% and 99% levels. * indicates statistical significance at the 10% level. Refer to Appendix A for variable definitions.

Panel A. Descriptive Statistics

	Mean	P25	P50	P75	Std. Dev.
SUED$_{i,t}$	0.032	0.000	0.000	0.000	0.177
SUED_INV$_{i,t}$	0.068	0.000	0.000	0.000	0.251
FPS$_{i,t}$	0.246	0.000	0.000	0.000	0.431
ln(ASSETS$_{i,t-1}$)	6.746	5.194	6.699	8.226	2.176
SALES_GR$_{i,t-1}$	0.066	−0.016	0.033	0.125	0.223
CAR$_{i,t-1}$	−0.025	−0.234	−0.011	0.190	0.424
RETURN_SKEW$_{i,t-1}$	0.237	−0.302	0.204	0.731	0.805
RETURN_VOL$_{i,t-1}$	0.114	0.067	0.097	0.143	0.068
SHARE_TURN$_{i,t-1}$	2.284	0.891	1.624	2.783	2.389
VIEWS > 0_IND$_{i,t}$	0.440	0.000	0.000	1.000	0.496
VIEWS > 0_IND$_{i,t-1}$	0.401	0.000	0.000	1.000	0.490
TOP_PLF_VIEWS > 0_IND$_{i,t}$	0.152	0.000	0.000	0.000	0.359
TOP_PLF_LN_VIEWS$_{i,t}$	0.274	0.000	0.000	0.000	0.791
TOP_PLF_LN_VIEWS$_{i,t-1}$	0.220	0.000	0.000	0.000	0.711
REM_PLF_VIEWS > 0_IND$_{i,t}$	0.385	0.000	0.000	1.000	0.487
REM_PLF_LN_VIEWS$_{i,t}$	0.677	0.000	0.000	1.099	1.135
REM_PLF_LN_VIEWS$_{i,t-1}$	0.644	0.000	0.000	0.693	1.126
BIOTECH$_{i,t}$	0.046	0.000	0.000	0.000	0.209
COMP_HARDWARE$_{i,t}$	0.015	0.000	0.000	0.000	0.120
ELECTRONICS$_{i,t}$	0.069	0.000	0.000	0.000	0.254
RETAIL$_{i,t}$	0.049	0.000	0.000	0.000	0.216
COMP_SOFTWARE$_{i,t}$	0.052	0.000	0.000	0.000	0.221

Panel B. Correlations

	1	2	3	4	5	6	7	8	9	10	11
(1) SUED$_{i,t}$	–										
(2) SUED_INV$_{i,t}$	0.68*	–									
(3) FPS$_{i,t}$	0.02*	0.03*	–								
(4) ln(ASSETS$_{i,t-1}$)	0.03*	0.04*	−0.15*	–							
(5) SALES_GR$_{i,t-1}$	0.04*	0.04*	0.01	−0.03*	–						
(6) CAR$_{i,t-1}$	−0.00	0.01*	0.04*	0.04*	0.19*	–					
(7) RETURN_SKEW$_{i,t-1}$	−0.01	−0.00	0.03*	−0.18*	0.01*	0.21*	–				
(8) RETURN_VOL$_{i,t-1}$	0.05*	0.06*	0.09*	−0.44*	0.04*	0.10*	0.35*	–			
(9) SHARE_TURN$_{i,t-1}$	0.10*	0.10*	0.10*	0.13*	0.09*	0.03*	0.05*	0.33*	–		
(10) VIEWS > 0_IND$_{i,t}$	0.19*	0.23*	0.04*	0.28*	0.03*	0.02*	−0.08*	−0.06*	0.14*	–	
(11) VIEWS > 0_IND$_{i,t-1}$	0.06*	0.09*	0.04*	0.28*	−0.02*	−0.01	−0.08*	−0.06*	0.15*	0.27*	–
(12) TOP_PLF_VIEWS > 0_IND$_{i,t}$	0.28*	0.29*	0.03*	0.19*	0.02*	0.00	−0.05*	−0.00	0.14*	0.48*	0.21*
(13) TOP_PLF_LN_VIEWS$_{i,t}$	0.40*	0.38*	0.03*	0.17*	0.03*	−0.01	−0.05*	0.01	0.14*	0.39*	0.21*
(14) TOP_PLF_LN_VIEWS$_{i,t-1}$	0.09*	0.12*	0.03*	0.16*	−0.01*	−0.08*	−0.07*	0.05*	0.17*	0.20*	0.38*
(15) REM_PLF_VIEWS > 0_IND$_{i,t}$	0.21*	0.23*	0.04*	0.27*	0.03*	0.03*	−0.08*	−0.07*	0.13*	0.89*	0.24*
(16) REM_PLF_LN_VIEWS$_{i,t}$	0.39*	0.40*	0.05*	0.27*	0.03*	0.01	−0.06*	−0.03*	0.16*	0.67*	0.28*
(17) REM_PLF_LN_VIEWS$_{i,t-1}$	0.08*	0.11*	0.05*	0.27*	−0.02*	−0.03*	−0.07*	−0.02*	0.18*	0.31*	0.70*
(18) BIOTECH$_{i,t}$	0.04*	0.05*	0.38*	−0.14*	−0.01	0.03*	0.03*	0.14*	0.06*	0.03*	0.02*
(19) COMP_HARDWARE$_{i,t}$	0.02*	0.02*	0.21*	−0.03*	−0.00	−0.00	0.01	0.02*	0.01	0.01	0.01
(20) ELECTRONICS$_{i,t}$	−0.02*	−0.01*	0.48*	−0.05*	−0.05*	−0.01	0.02*	0.04*	0.06*	−0.00	−0.00
(21) RETAIL$_{i,t}$	−0.00	0.00	0.40*	0.03*	0.05*	0.03*	−0.01*	−0.05*	0.04*	0.05*	0.06*
(22) COMP_SOFTWARE$_{i,t}$	0.01	0.00	0.36*	−0.09*	0.03*	0.01*	0.00	0.03*	0.03*	−0.00	−0.00

	12	13	14	15	16	17	18	19	20	21
(12) TOP_PLF_VIEWS > 0_IND$_{i,t}$	–									
(13) TOP_PLF_LN_VIEWS$_{i,t}$	0.82*	–								
(14) TOP_PLF_LN_VIEWS$_{i,t-1}$	0.29*	0.35*	–							
(15) REM_PLF_VIEWS > 0_IND$_{i,t}$	0.22*	0.23*	0.17*	–						
(16) REM_PLF_LN_VIEWS$_{i,t}$	0.33*	0.38*	0.24*	0.75*	–					
(17) REM_PLF_LN_VIEWS$_{i,t-1}$	0.25*	0.26*	0.37*	0.31*	0.42*	–				
(18) BIOTECH$_{i,t}$	0.03*	0.03*	0.01*	0.03*	0.06*	0.06*	–			
(19) COMP_HARDWARE$_{i,t}$	0.01	0.02*	0.02*	0.01*	0.03*	0.02*	−0.03*	–		
(20) ELECTRONICS$_{i,t}$	−0.02*	−0.02*	−0.01*	−0.00	−0.01	−0.01	−0.06*	−0.03*	–	
(21) RETAIL$_{i,t}$	0.05*	0.04*	0.04*	0.05*	0.04*	0.04*	−0.05*	−0.03*	−0.06*	–
(22) COMP_SOFTWARE$_{i,t}$	0.00	0.02*	0.02*	0.00	0.00	0.01	−0.05*	−0.03*	−0.06*	−0.05*

is significant variation in the number of views (TOP_PLF_LN_VIEWS$_{i,t}$ and REM_PLF_LN_VIEWS$_{i,t}$), which our tests exploit to identify firms with high levels of plaintiff-lawyer scrutiny. For example, the interquartile range for firm years with views is 11 (10) for top (remaining) lawyers (untabulated).

Panel B of Table 4 presents univariate evidence that plaintiff-lawyer views over the prior year ($\text{TOP\_PLF\_LN\_VIEWS}_{i,t-1}$ and $\text{REM\_PLF\_LN\_VIEWS}_{i,t-1}$) are positively associated with realized litigation ($\text{SUED}_{i,t}$ and $\text{SUED\_INV}_{i,t}$) in the following year ($p < 0.01$). We use lagged views to measure ex ante litigation risk to avoid a mechanical positive relation with litigation because views increase around key litigation dates, as shown in Section III.E.

2. Investigation Announcements

We posit that firms targeted with plaintiff-lawyer investigations face high litigation risk, so investigations ($\text{INVESTIG\_ANNCT.}_{i,t}$) are also a proxy for realized litigation risk. Securities lawyers work on a contingency fee basis, incurring all costs for unsuccessful litigation (Baker et al. (2015)). Their willingness to incur investigation costs strongly signals the targeted firm faces high litigation risk, so targeted firms should face litigation at higher rates than other firms. That is, if investigations and litigation filings both proxy for the same construct of interest (high litigation risk), we expect them to have a high overlap.

Table 5 presents the results. Over 80% of the firm years with filings also have investigations, more than 85% of which are published before or within a few days of the filing date (untabulated). Further, over 40% of the firm years with investigations have filings in the same year, which is over 13 times larger than the unconditional probability of a firm year having litigation (i.e., 43% vs. 3%). About 25% of firm years with an investigation but without a filing are sued in a different firm year in our sample (e.g., over 16% of them have filings in the next year, untabulated). This analysis validates that firms with plaintiff-lawyer investigations have high litigation risk. Thus, because most research is focused on the effect of litigation risk on firm actions, rather than litigation itself, researchers attempting to model predicted litigation risk should use the ex post revelation of *both* filings and plaintiff-lawyer investigations. However, to allow comparison to prior research and demonstrate this improved construct validity, we report results using both filings only and this combined measure.

3. Predicting Realized Litigation Risk

We conjecture that lagged plaintiff-lawyer views proxy for ex ante litigation risk. Firms differ on myriad unobservable dimensions that make them relatively more or less likely to face litigation *before* any bad news events that trigger such litigation are publicly revealed (firm culture, executive characteristics, etc.).

TABLE 5

Do Plaintiff-Lawyer Investigations Imply High Litigation Risk?

Table 5 presents descriptive statistics examining the supposition that firms targeted in investigation announcements by plaintiffs' lawyers face high litigation risk in the firm-year sample from 2012 to 2016 as shown in Panel B of Table A.2 in the Supplementary Material. We note that overall, 51.4% of all unique firms that face an investigation announcement during the sample period also face litigation at some point during the sample period. For the 458 firm years that include both a litigation filing and an investigation announcement, 326 investigation announcements were made on or prior to the filing date; another 68 were made within 4 days following the litigation filing date. Refer to Appendix A for variable definitions.

Category	# Firm-Years	% of $\text{SUED}_{i,t} = 1$	% of INVESTIG_ ANNCT.$_{i,t} = 1$
Firm-years with $\text{SUED\_INV}_{i,t} = 1$	1,163		
Firm-years with $\text{SUED}_{i,t} = 1$ and $\text{INVESTIG\_ANNCT.}_{i,t} = 1$	458	82.7	42.9
Firm-years with $\text{SUED}_{i,t} = 1$ and $\text{INVESTIG\_ANNCT.}_{i,t} = 0$	96	17.3	
Firm-years with $\text{SUED}_{i,t} = 0$ and $\text{INVESTIG\_ANNCT.}_{i,t} = 1$	609		57.1

For example, some firms are relatively "good" ("bad") firms, which make them less (more) likely to engage in actions that can later be alleged as misconduct. Furthermore, some firms make better or worse targets for plaintiffs' lawyers. For example, larger and less distressed firms have "deep pockets," making them more profitable litigation targets (Field et al. (2005), Coffee (2006), Baker and Griffith (2010), and Donelson et al. (2021a)). Collectively, these characteristics are relatively persistent and create ex ante litigation risk, but many are largely unobservable and thus cannot be directly measured. However, we can measure plaintiff-lawyer views. Similar to Iliev et al. (2021), who use investors' views of proxy filings on EDGAR as a measure of governance research, we assert that plaintiff-lawyer views measure firms' ongoing monitoring by plaintiffs' lawyers. Notably, plaintiffs' lawyers may increase their research following bad news events, including litigation filings that constitute realized litigation risk. Thus, for plaintiff-lawyer views to represent ex ante litigation risk, they must predate the litigation and the triggering events and instead proxy for the underlying firm quality that determines litigation risk.[4]

To assess the ability of plaintiff-lawyer scrutiny to proxy for litigation risk, we test whether it predicts realized litigation better than existing methods. As our benchmark for existing methods, we reestimate the recommended ex ante litigation risk model from Kim and Skinner (2012) and compare its ability to predict filings or a combination of filings and plaintiff-lawyer investigations to a simpler model that uses only *lagged* plaintiff-lawyer views, firm size, and indicator variables for high-litigation industries following Brochet and Srinivasan (2014). However, we show our model or variations thereof also outperforms alternative litigation risk measures and modifications to the Kim and Skinner (2012) model in the Supplementary Material. We include size and industry indicator variables as covariates because they impose virtually no data requirements, beyond a firm being in Compustat, and are two of the strongest litigation predictors Kim and Skinner (2012)). We limit our approach to this sparse model as it is important to balance improved explanatory power with parsimony to avoid significantly reducing sample sizes and potentially introducing bias (e.g., only examining large firms) when creating prediction models. We estimate the following logistic regressions:

$$\text{(2)}\quad \text{SUED}_{i,t}/\text{SUED\_INV}_{i,t} = \beta_0 + \beta_1\text{FPS}_{i,t} + \beta_2 \ln(\text{ASSETS}_{i,t-1}) \\ + \beta_3\text{SALES\_GR}_{i,t-1} + \beta_4\text{CAR}_{i,t-1} \\ + \beta_5\text{RETURN\_SKEW}_{i,t-1} + \beta_6\text{RETURN\_VOL}_{i,t-1} \\ + \beta_7\text{SHARE\_TURN}_{i,t-1} + \varepsilon_{i,t},$$

$$\text{(3)}\quad \text{SUED}_{i,t}/\text{SUED\_INV}_{i,t} = \beta_0 + \beta_1\text{TOP\_PLF\_LN\_VIEWS}_{i,t-1} \\ + \beta_2\text{REM\_PLF\_LN\_VIEWS}_{i,t-1} + \beta_3\text{BIOTECH}_{i,t} \\ + \beta_4\text{COMP\_HARDWARE}_{i,t} + \beta_5\text{ELECTRONICS}_{i,t} \\ + \beta_6\text{RETAIL}_{i,t} + \beta_7\text{COMP\_SOFTWARE}_{i,t} \\ + \beta_8 \ln(\text{ASSETS})_{i,t-1} + \varepsilon_{i,t},$$

[4]We confirmed with a senior plaintiffs' lawyer that most of their scrutiny was proactive and focused on ongoing monitoring of a relatively constant set of high litigation risk firms, rather than reactive, such as responding to bad news events (e.g., restatements) across a wide set of firms.

where $\text{SUED}_{i,t}$ ($\text{SUED\_INV}_{i,t}$) is an indicator variable set to 1 if firm i is sued (sued or faces an investigation) in year t. $\ln(\text{ASSETS}_{i,t-1})$ is the natural log of assets in year $t-1$. $\text{SALES\_GR}_{i,t-1}$ is the sales growth in year $t-1$. $\text{CAR}_{i,t-1}$ ($\text{SHARE\_TURN}_{i,t-1}$) is the cumulative abnormal return (share turnover) over year $t-1$, and $\text{RETURN\_SKEW}_{i,t-1}$ ($\text{RETURN\_VOL}_{i,t-1}$) is the return skewness (standard deviation) over year $t-1$. $\text{FPS}_{i,t}$ is an indicator variable set to 1 for firms operating in one of several high-litigation industries. $\text{BIOTECH}_{i,t}$, $\text{COMP\_HARDWARE}_{i,t}$, $\text{ELECTRONICS}_{i,t}$, $\text{RETAIL}_{i,t}$, and $\text{COMP\_SOFTWARE}_{i,t}$ are indicator variables set to 1 if the firm operates in the respective industry. See Appendix A for detailed definitions.

Our measures of plaintiff-lawyer scrutiny are the natural log of one plus total EDGAR views by top ($\text{TOP\_PLF\_LN\_VIEWS}_{i,t-1}$) and remaining ($\text{REM\_PLF\_LN\_VIEWS}_{i,t-1}$) lawyers in year $t-1$. Higher scrutiny implies higher litigation risk, so we expect a positive association between plaintiff-lawyer views and realized litigation risk. It is unclear whether views from the top or remaining plaintiffs' lawyers will have a stronger association, so we separately estimate these variables to allow the coefficient to vary. For example, views from top lawyers may have a stronger association because their more selective nature may imply a high proportion of scrutinized firms will face realized litigation risk. Alternatively, they may have a weaker association if they screen multiple firms to pursue only the most profitable cases.

Critically, the median length of time between the class period end and filing date in our sample is only 13 days, similar to the analysis of this period by NERA (2017). This is because there is a race-to-the-courthouse for most filings due to competition among plaintiffs' law firms (Perino (2003), Weiss (2008), and Cornerstone Research (2013)). Thus, for around 85% of the securities class actions in our sample, both the class period end and filing occur in year t. Accordingly, when using year $t-1$ plaintiff-lawyer view, by construction, virtually all views cannot include the spike in views around the class period end or related bad news events nor be due to plaintiffs' lawyers initiating the litigation process. Rather, these views identify firms facing general scrutiny due to inherent firm characteristics that make litigation more likely. As such, lagged plaintiff-lawyer views predate, and are relatively exogenous to, firms' litigation-related bad news events and thus a proxy for ex ante litigation risk.

To assess whether our proposed litigation risk measure improves the prediction of realized litigation risk, relative to other approaches, we need to compare how effectively the different models predict our outcome variables. Prediction models often involve trade-offs between the detected proportion of true positives (i.e., sensitivity) and true negatives (i.e., specificity) (Hosmer and Lemeshow (2000)). For example, a lower threshold to identify future sued firms increases the probability we correctly identify sued firms (i.e., true positives) and incorrectly classify nonsued firms (i.e., false positives). The most common model evaluation metric for binary dependent variables balances this trade-off by estimating the ROC curve (e.g., Bharath and Dittmar (2010), Dimmock and Gerken (2012), Kim and Skinner (2012), Acharya, Amihud, and Bharath (2013), and Nagel and Purnanandam (2020)). The ROC curve plots the probability of detecting a true positive and a false positive (i.e., 1 – specificity) "for the entire range of possible cutpoints"

https://doi.org/10.1017/S0022109022001508 Published online by Cambridge University Press

https://doi.org/10.1017/S0022109022001508 Published online by Cambridge University Press

(Hosmer and Lemeshow (2000), 160, emphasis added), and researchers often report the area under this curve (AUC) as the model performance metric.

However, while not recognized in finance, the AUC is an unreliable performance metric for imbalanced data sets (i.e., few observations set to 1) because much of its performance can be driven by irrelevant regions for researchers (see Fernandéz, Garcıa, Galar, Prati, Krawczyk, and Herrera (2018)). As summarized by Brabec and Machlica ((2018), pp. 3–4): "In the case of imbalanced data sets, the regions of no interest may represent most of the area under the curve, having a dominant influence on the value of AUC."[5] Given the rare incidence of litigation, AUC is an inappropriate metric in this setting and for other rare events, such as bankruptcy, fraud, and extreme returns (e.g., Dimmock and Gerken (2012)). In Section V of the Supplementary Material, we discuss this issue in detail and show why the AUC provides misleading insights in our setting.

Moreover, AUC treats false positives and negatives as equally costly, but false positives are arguably costlier in our setting, similar to Dimmock and Gerken (2012) and Bao, Ke, Li, Yu, and Zhang (2020). For example, if investors avoid firms they believe may be sued by using a model to predict litigation, false positives create opportunity costs by erroneously reducing their investment opportunity set and likely increasing the cost of capital for the suspect firms (Dimmock and Gerken (2012)). Further, corporate monitors have finite resources, so false positives divert scarce monitoring resources from firms that actually have high litigation risk. Accordingly, reducing false positives increases the value proposition of litigation risk prediction models (Beneish and Vorst (2022)). Finally, companies that falsely identify themselves as facing high litigation risk may overinvest in precautionary measures, such as liability insurance or disclosure changes (Harvard and Law Review (2019)). That said, false negatives still impose costs on the market (e.g., investors not scrutinizing high litigation risk firms), so we ensure that decreases in false positives are not achieved only by increasing false negatives.

Accordingly, our analysis focuses on the relevant region and compares models' ability to correctly identify firms that face future realized litigation risk (i.e., true positives) by examining precision and sensitivity (see Bao et al. (2020)). Intuitively, precision measures how many firms that our tests classify as having future realized litigation risk *actually* face realized litigation risk; sensitivity measures how many firms with future realized litigation risk are correctly identified. Notably, these metrics produce similar results when one's expectation of an event's likelihood is close to the true rate. Higher sensitivity also indicates a lower false negative rate (i.e., how many firms with future realized litigation risk are incorrectly identified), as the false negative rate can be calculated as 1 – sensitivity. For completeness and comparison with prior research, we also report specificity (i.e., how many firms that do not have future realized litigation risk are correctly identified) and the AUC. We depict the mathematical calculation of these performance metrics in Table A.3 in the Supplementary Material.

[5]For example, in network traffic intrusion-detection data sets, which are highly imbalanced, "If AUC was computed in the usual way over the complete ROC curve then 99.99% of the area would be irrelevant and would represent only noise in the final outcome" (Brabec and Machlica (2018), pp. 3–4).

https://doi.org/10.1017/S0022109022001508 Published online by Cambridge University Press

To calculate these metrics, we must select a probability threshold to classify firms using their predicted values (Hosmer and Lemeshow (2000)), so we classify observations in the top 3% of predicted litigation risk as observations with realized litigation risk, similar to the methodology of Bao et al. (2020). We use 3% for $\text{SUED}_{i,t}$ because it is both the expected and observed average litigation rate (Table 4). To be consistent, we also use a 3% threshold when predicting $\text{SUED\_INV}_{i,t}$ but find similar inferences using a 7% (see the average rate in Table 4) or 10% rate (see Nelson and Pritchard (2016)), both of which are reported in the Supplementary Material. Given concerns of overfitting prediction models (see Avramov (2002), Dangl and Halling (2012), and Géron (2017)), we also examine out-of-sample statistics. We use a K-fold analysis with K set equal to 10 (Dimmock and Gerken (2012), Kim and Skinner (2012), and Larcker and Zakolyukina (2012).

Columns 1–3 in Panel A of Table 6 present results with $\text{SUED}_{i,t}$ as the dependent variable. Column 1 shows results from estimating equation (2) and resembles findings in Kim and Skinner (2012), other than $\text{CAR}_{i,t-1}$, which loads negatively ($p < 0.05$). Column 2 shows results from estimating equation (3), and both $\text{TOP\_PLF\_LN\_VIEWS}_{i,t-1}$ and $\text{REM\_PLF\_LN\_VIEWS}_{i,t-1}$ are significantly positive ($p < 0.01$). Similar to Table 4, the coefficient of $\text{TOP\_PLF\_LN\_VIEWS}_{i,t-1}$ is larger than $\text{REM\_PLF\_LN\_VIEWS}_{i,t-1}$, but the difference is marginally insignificant ($p = 0.16$, untabulated). The effect of plaintiff-lawyer views is economically significant. The overall marginal effect for $\text{TOP\_PLF\_LN\_VIEWS}_{i,t-1}$ ($\text{REM\_PLF\_LN\_VIEWS}_{i,t-1}$) increases the litigation probability by 0.85 (0.53) percentage points. As the unconditional probability of litigation is only 3.2%, this corresponds to a 27% (17%) increase overall (untabulated). Column 3 presents results from combining the 2 equations excluding $\text{FPS}_{i,t}$; $\text{TOP\_PLF\_LN\_VIEWS}_{i,t-1}$ and $\text{REM\_PLF\_LN\_VIEWS}_{i,t-1}$ remain significantly positive ($p < 0.01$).

Notably, despite imposing fewer data constraints, the simplified model in column 2 has higher precision and sensitivity. Although the increases are small in absolute terms, in relative terms, they improve in-sample (out-of-sample) precision and sensitivity by over 20% (30%). The combined approach (column 3) performs identical in-sample but worse out-of-sample, so the additional variables from column 1 appear to provide limited improvement. Specificity is virtually unchanged across models. Notably, focusing on the observations of interest tells a different story than the AUC, demonstrating its unsuitability for imbalanced data sets.

Columns 4–6 report results examining $\text{SUED\_INV}_{i,t}$. The coefficient of $\text{TOP\_PLF\_LN\_VIEWS}_{i,t-1}$ is larger than $\text{REM\_PLF\_LN\_VIEWS}_{i,t-1}$ ($p < 0.1$, untabulated). Performance metrics show our model performs significantly better with this broader dependent variable. Specifically, column 5 reports improved precision and sensitivity, relative to the model in column 4, by around 40% (50%) in-sample (out-of-sample). Using the combined model in column 6, there are larger (smaller) performance improvements in-sample (out-of-sample). Thus, these additional variables may overfit estimates of litigation risk. Notably, there is over an 85% (94%) in-sample (out-of-sample) increase in precision with our model when using our broader measure of realized litigation risk (column 5 vs. 2). The precision for Kim and Skinner (2012) similarly increases (column 4 vs. 1),

TABLE 6

Predicting Realized Litigation Risk

Table 6 presents results examining proxies for ex ante litigation risk using realized litigation filings (columns 1–3) and realized litigation filings, supplemented with plaintiff-lawyer investigation announcements (columns 4–6) in the firm-year sample from 2012 to 2016 as shown in Panel B of Table A.2 in the Supplementary Material. Panel A uses all observations during the sample period; Panel B excludes firm-years with litigation filings where the class period end occurs during the prior fiscal year; Panel C presents results where we lag all variables (i.e., using plaintiff-lawyer views and market-related variables from year $t-2$). Columns 1 and 4 present results based on estimating equation (2) using variables recommended by Kim and Skinner (2012) (model 3 of their Table 7). Columns 2 and 5 present results based on estimating equation (3) using plaintiff-lawyer EDGAR views. Columns 3 and 6 present results based on combining the 2 groups of variables other than $FPS_{i,t}$ and estimating the equation. All equations use logistic estimation. Both in-sample and out-of-sample model performance metrics are reported. The out-of-sample statistics are calculated using the "*K*-fold" cross-validation procedure similar to Dimmock and Gerken (2012), Kim and Skinner (2012), and Larcker and Zakolyukina (2012) with a 10-fold cross-validation procedure. Precision is calculated as the percentage of predicted positive cases that are true positives. Sensitivity is calculated as the percent age of true positive cases correctly identified. Specificity is calculated as the percentage of true negative cases correctly identified. We classify cases as being predicted positive if the predicted probability is in the top 3% of all probabilities. To minimize the influence of outliers, all continuous variables are winsorized at the 1% and 99% levels. Standard errors appear in parentheses and are adjusted for clustering at the company level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. Refer to Appendix A for variable definitions.

Panel A. Overall

	$SUED_{i,t}$			$SUED\_INV_{i,t}$		
	1	2	3	4	5	6
$TOP\_PLF\_LN\_VIEWS_{i,t-1}$		0.28***	0.21***		0.28***	0.24***
		(0.05)	(0.05)		(0.03)	(0.04)
$REM\_PLF\_LN\_VIEWS_{i,t-1}$		0.17***	0.13***		0.19***	0.17***
		(0.04)	(0.04)		(0.03)	(0.03)
$FPS_{i,t}$	0.19*			0.26***		
	(0.10)			(0.08)		
$\ln(ASSETS_{i,t-1})$	0.15***	0.04**	0.11***	0.14***	0.03**	0.09***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)
$SALES\_GR_{i,t-1}$	0.88***		0.91***	0.62***		0.66***
	(0.18)		(0.18)	(0.13)		(0.13)
$CAR_{i,t-1}$	−0.21**		−0.13	−0.03		0.06
	(0.10)		(0.10)	(0.07)		(0.07)
$RETURN\_SKEW_{i,t-1}$	−0.13**		−0.09	−0.11***		−0.07
	(0.06)		(0.06)	(0.04)		(0.04)
$RETURN\_VOL_{i,t-1}$	4.63***		3.50***	4.40***		3.23***
	(0.71)		(0.74)	(0.55)		(0.56)
$SHARE\_TURN_{i,t-1}$	0.09***		0.08***	0.06***		0.05***
	(0.01)		(0.01)	(0.01)		(0.01)
$BIOTECH_{i,t}$		0.81***	0.73***		0.75***	0.66***
		(0.15)	(0.16)		(0.12)	(0.12)
$COMP\_HARDWARE_{i,t}$		0.55**	0.53**		0.62***	0.59***
		(0.25)	(0.25)		(0.18)	(0.18)
$ELECTRONICS_{i,t}$		−0.31	−0.37*		−0.13	−0.17
		(0.19)	(0.19)		(0.13)	(0.13)
$RETAIL_{i,t}$		−0.18	−0.21		−0.01	−0.03
		(0.21)	(0.21)		(0.16)	(0.16)
$COMP\_SOFTWARE_{i,t}$		0.22	0.15		0.09	0.05
		(0.18)	(0.18)		(0.14)	(0.14)
INTERCEPT	−5.35***	−3.98***	−5.08***	−4.39***	−3.17***	−4.04***
	(0.23)	(0.15)	(0.22)	(0.17)	(0.11)	(0.16)
No. of obs.	17,179	17,179	17,179	17,179	17,179	17,179
Pseudo R^2	0.038	0.031	0.055	0.029	0.035	0.050
In-sample						
Pred. posit.	516	516	516	516	516	516
Correct pred. Posit.	49	59	59	79	111	117
Precision	0.095	0.114	0.114	0.153	0.215	0.227
Sensitivity	0.088	0.106	0.106	0.068	0.095	0.101
False negative rate	0.912	0.894	0.894	0.932	0.905	0.899
Specificity	0.972	0.973	0.973	0.973	0.975	0.975
AUC	0.674	0.635	0.707	0.638	0.638	0.681

(continued on next page)

https://doi.org/10.1017/S0022109022001508 Published online by Cambridge University Press

TABLE 6 (continued)

Predicting Realized Litigation Risk

Panel A. Overall (continued)

	$SUED_{i,t}$			$SUED\_INV_{i,t}$		
	1	2	3	4	5	6
Out-of-sample						
Precision	0.085	0.112	0.100	0.146	0.217	0.208
Sensitivity	0.080	0.105	0.094	0.065	0.097	0.093
False negative rate	0.920	0.895	0.906	0.935	0.903	0.907
Specificity	0.971	0.972	0.972	0.972	0.975	0.974
AUC	0.669	0.626	0.696	0.635	0.634	0.670

Panel B. Excluding Litigation Filings with Class Period Ends in Year t − 1

	$SUED_{i,t}$			$SUED\_INV_{i,t}$		
	1	2	3	4	5	6
$TOP\_PLF\_LN\_VIEWS_{i,t-1}$		0.18***	0.13**		0.24***	0.20***
		(0.05)	(0.06)		(0.04)	(0.04)
$REM\_PLF\_LN\_VIEWS_{i,t-1}$		0.18***	0.14***		0.20***	0.17***
		(0.04)	(0.04)		(0.03)	(0.03)
$FPS_{i,t}$	0.19*			0.27***		
	(0.11)			(0.08)		
$\ln(ASSETS_{i,t-1})$	0.14***	0.05**	0.10***	0.14***	0.04**	0.08***
	(0.03)	(0.02)	(0.03)	(0.02)	(0.02)	(0.02)
$SALES\_GR_{i,t-1}$	0.90***		0.94***	0.61***		0.66***
	(0.19)		(0.19)	(0.14)		(0.14)
$CAR_{i,t-1}$	0.03		0.08	0.09		0.17**
	(0.11)		(0.11)	(0.07)		(0.08)
$RETURN\_SKEW_{i,t-1}$	−0.08		−0.04	−0.09**		−0.05
	(0.07)		(0.06)	(0.04)		(0.04)
$RETURN\_VOL_{i,t-1}$	3.96***		2.94***	4.06***		2.94***
	(0.75)		(0.78)	(0.56)		(0.57)
$SHARE\_TURN_{i,t-1}$	0.08***		0.07***	0.05***		0.04***
	(0.01)		(0.02)	(0.01)		(0.01)
$BIOTECH_{i,t}$		0.91***	0.81***		0.79***	0.69***
		(0.16)	(0.16)		(0.12)	(0.12)
$COMP\_HARDWARE_{i,t}$		0.55*	0.51*		0.62***	0.59***
		(0.29)	(0.28)		(0.20)	(0.19)
$ELECTRONICS_{i,t}$		−0.25	−0.32		−0.10	−0.14
		(0.20)	(0.20)		(0.13)	(0.13)
$RETAIL_{i,t}$		−0.11	−0.16		0.03	−0.00
		(0.22)	(0.22)		(0.16)	(0.16)
$COMP\_SOFTWARE_{i,t}$		−0.01	−0.10		−0.03	−0.08
		(0.23)	(0.23)		(0.16)	(0.15)
INTERCEPT	−5.31***	−4.10***	−5.08***	−4.36***	−3.22***	−4.02***
	(0.24)	(0.16)	(0.24)	(0.17)	(0.11)	(0.17)
No. of obs.	17,110	17,110	17,110	17,110	17,110	17,110
Pseudo R^2	0.032	0.024	0.046	0.026	0.030	0.045
In-sample						
Pred. posit.	514	514	514	514	514	514
Correct pred. posit.	43	51	52	74	101	106
Precision	0.084	0.099	0.101	0.144	0.196	0.206
Sensitivity	0.089	0.105	0.107	0.068	0.092	0.097
False negative rate	0.911	0.895	0.893	0.932	0.908	0.903
Specificity	0.972	0.972	0.972	0.973	0.974	0.975
AUC	0.663	0.616	0.689	0.632	0.629	0.673
Out-of-sample						
Precision	0.063	0.098	0.073	0.131	0.198	0.188
Sensitivity	0.068	0.105	0.078	0.062	0.094	0.090
False negative rate	0.932	0.895	0.922	0.938	0.906	0.910
Specificity	0.971	0.972	0.971	0.972	0.974	0.974
AUC	0.656	0.606	0.674	0.629	0.624	0.662

(continued on next page)

TABLE 6 (continued)

Predicting Realized Litigation Risk

Panel C. Using Plaintiff-Lawyer Views from Year t − 2

	$SUED_{i,t}$			$SUED\_INV_{i,t}$		
	1	2	3	4	5	6
$TOP\_PLF\_LN\_VIEWS_{i,t-2}$		0.13**	0.10		0.13***	0.11**
		(0.06)	(0.06)		(0.04)	(0.04)
$REM\_PLF\_LN\_VIEWS_{i,t-2}$		0.13***	0.11***		0.20***	0.19***
		(0.04)	(0.04)		(0.03)	(0.03)
$FPS_{i,t-1}$	0.30***			0.32***		
	(0.11)			(0.08)		
$\ln(ASSETS_{i,t-2})$	0.12***	0.07***	0.10***	0.11***	0.05***	0.06***
	(0.03)	(0.02)	(0.03)	(0.02)	(0.02)	(0.02)
$SALES\_GR_{i,t-2}$	0.73***		0.81***	0.48***		0.57***
	(0.17)		(0.17)	(0.13)		(0.13)
$CAR_{i,t-2}$	0.30**		0.34***	0.22***		0.29***
	(0.12)		(0.12)	(0.08)		(0.08)
$RETURN\_SKEW_{i,t-2}$	−0.12*		−0.10	−0.12**		−0.10**
	(0.06)		(0.06)	(0.05)		(0.05)
$RETURN\_VOL_{i,t-2}$	2.65***		1.93**	2.10***		1.22*
	(0.82)		(0.84)	(0.65)		(0.65)
$SHARE\_TURN_{i,t-2}$	0.06***		0.06***	0.05***		0.05***
	(0.01)		(0.02)	(0.01)		(0.01)
$BIOTECH_{i,t-1}$		1.06***	0.99***		0.83***	0.78***
		(0.16)	(0.16)		(0.13)	(0.13)
$COMP\_HARDWARE_{i,t-1}$		0.75***	0.71**		0.71***	0.69***
		(0.28)	(0.28)		(0.19)	(0.19)
$ELECTRONICS_{i,t-1}$		−0.22	−0.31		−0.08	−0.14
		(0.20)	(0.20)		(0.14)	(0.14)
$RETAIL_{i,t-1}$		0.00	−0.08		0.09	0.03
		(0.22)	(0.22)		(0.17)	(0.17)
$COMP\_SOFTWARE_{i,t-1}$		0.37*	0.29		0.07	0.02
		(0.20)	(0.20)		(0.16)	(0.16)
INTERCEPT	−4.93***	−4.15***	−4.80***	−3.91***	−3.26***	−3.63***
	(0.24)	(0.17)	(0.24)	(0.18)	(0.12)	(0.18)
No. of obs.	15,999	15,999	15,999	15,999	15,999	15,999
Pseudo R^2	0.026	0.020	0.038	0.019	0.023	0.033
In-sample						
Pred. posit.	480	480	480	480	480	480
Correct pred. posit.	31	49	53	55	81	88
Precision	0.065	0.102	0.110	0.115	0.169	0.183
Sensitivity	0.064	0.101	0.109	0.053	0.078	0.085
False negative rate	0.936	0.899	0.891	0.947	0.922	0.915
Specificity	0.971	0.972	0.972	0.972	0.973	0.974
AUC	0.640	0.604	0.667	0.611	0.614	0.644
Out-of-sample						
Precision	0.062	0.094	0.092	0.115	0.171	0.175
Sensitivity	0.062	0.093	0.091	0.053	0.079	0.081
False negative rate	0.938	0.907	0.909	0.947	0.921	0.919
Specificity	0.971	0.972	0.972	0.972	0.973	0.974
AUC	0.632	0.598	0.656	0.606	0.610	0.638

consistent with many of the apparent false positives in their model corresponding to firms being investigated. Although sensitivity decreases by around 8%–10% using $SUED\_INV_{i,t}$, it is clear the broader measure of realized litigation risk significantly reduces the population of high litigation risk false negatives.[6]

[6]The decrease in sensitivity is mechanical because the threshold to classify an observation as high litigation risk remains the top 3%, even though our ex ante expectation is around 10%. Using a 10% threshold, sensitivity (precision) increases by over 127% (40%) for column 5 relative to column 2 (Table A.35 in the Supplementary Material).

https://doi.org/10.1017/S0022109022001508 Published online by Cambridge University Press

As previously discussed, lagged plaintiff-lawyer views cannot be caused by the litigation process or related bad news events for virtually all sample cases due to the relatively short period between the class period end and the filing date. However, there is a small percentage of cases for which the class period end date occurs in year $t-1$. While we do not exclude these firm years in Panel A of Table 6 to avoid look-ahead bias, to ensure these cases do not drive results, Panel B (C) presents results after dropping firm years where the class period end is in year $t-1$ (using plaintiff-lawyer views and other variables from year $t-2$) to ensure the predictive ability of views is not due to the litigation process or related bad news events. We find similar results. We also conduct additional analyses in Section III of the Supplementary Material to further show that plaintiff-lawyer views proxy for the relatively persistent characteristics that make firms more likely to face realized litigation risk. For example, we show plaintiff-lawyer views are relatively persistent, whereas bad news events such as restatements and litigation filings are not. The takeaways from these analyses are clear: The predictive ability of plaintiff-lawyer views is due to ongoing plaintiff-lawyer scrutiny, rather than any alternative explanation. Thus, it seems clear that lagged plaintiff-lawyer views are superior proxies for ex ante litigation risk.

4. Matched Samples for Sued Firms

The ability of plaintiff-lawyer views to better predict realized litigation risk implies that it will reduce measurement error when used as a proxy for litigation risk. Even when used only as a control variable, measurement error in litigation risk proxies can result in inconsistent estimates on all regression parameters (Roberts and Whited (2013)). However, to demonstrate another way this measurement error can result in bias, we reexamine common methods of matching sued to nonsued firms. Prior research often matches sued and nonsued firms on litigation risk to examine the effect of firm choices on litigation or vice versa (e.g., Atanasov et al. (2012), Donelson et al. (2012), Brochet and Srinivasan (2014), Billings and Cedergren (2015), Pukthuanthong, Turtle, Walker, and Wang (2017)). Common ways to obtain a matched sample are using similar-sized firms in the same industry year with the closest predicted litigation risk from Kim and Skinner (2012) or using propensity scores, based on the covariates from their model. We examine if these matched firms *actually* have similar litigation risks.

The reason to use this research design is litigation risk may be correlated with both the dependent and independent variable of interest, so matching theoretically holds it constant across the sample. Otherwise, such correlations result in omitted variable bias (Roberts and Whited (2013), Meyer and Mittag (2017)). However, if plaintiff-lawyer views differ significantly between the 2 samples, this casts doubt on the effectiveness of these designs.

Table 7 presents the results. As shown in Panel A, where sued firms are matched to nonsued firms in the same industry, year, asset quintile, and the closest $\text{EX\_ANTE\_LIT\_RISK}_{i,t}$, sued firms face higher plaintiff-lawyer scrutiny from year $t-1$ to t ($p<0.01$), with year t being the litigation filing year. Further casting doubt that litigation risk is held constant, sued firms are nearly 12 times more likely than their matches to have a plaintiff-lawyer investigation in year t. In fact, matched firms have a similar rate of investigations as the average firm in our sample,

TABLE 7

Is Plaintiff-Lawyer Scrutiny Significantly Different Between Sued and Nonsued Firms Using Existing Methods to Match on Litigation Risk?

Table 7 presents results examining whether plaintiffs' law firm scrutiny is an omitted variable when matching sued to nonsued firms based on existing measures of litigation risk in the firm-year sample from 2012 to 2016 as shown in Panel B of Table A.2 in the Supplementary Material. Specifically, we examine whether yearly views by plaintiffs' lawyers significantly differ between firm-years in which a securities lawsuit has been filed ($SUED_{i,t}$ = 1) and a matched sample ($SUED_{i,t}$ = 0) of firm-years. Panel A identifies control firms for 552 sued firms using exact matching on year, industry (SIC2), total assets quintile (calculated at the population level on a yearly basis), and the closest match on $EX\_ANTE\_LIT\_RISK\_PROB_{i,t}$ as specified by KS2012. Panel B identifies control firms for 554 sued firms in the same year and industry (SIC2) based on propensity-score matching, with replacement, using the variables from model 3 of Table 7 of Kim and Skinner (2012). Model 3 of Table 7 is estimated at the population level using logistic estimation to follow the spirit of the KS2012 methodology. Panel A has two observations less than Panel B due to one SIC2 industry for one fiscal year not having any sued and nonsued firms in the same asset quintile. To minimize the influence of outliers, all continuous variables are winsorized at the 1% and 99% levels. *t*-tests are used to test for significant differences between the means of sued and matched firms. Refer to Appendix A for variable definitions.

Variable	*N*	SUED = 0	SUED = 1	Diff.	*t*-Stat.
Panel A. Matching on Year, Industry, Size, and Closest Litigation Risk					
$\ln(ASSETS_{i,t})$	552	7.10	7.15	0.05	0.36
$EX\_ANTE\_LIT\_RISK_{i,t}$	552	−0.59	−0.49	0.10	1.05
$TOP\_PLF\_LN\_VIEWS_{i,t}$	552	0.33	2.02	1.69	21.58
$TOP\_PLF\_LN\_VIEWS_{i,t-1}$	552	0.36	0.55	0.19	3.11
$TOP\_PLF\_LN\_VIEWS_{i,t-2}$	552	0.22	0.28	0.06	1.25
$REM\_PLF\_LN\_VIEWS_{i,t}$	552	0.71	3.08	2.37	30.46
$REM\_PLF\_LN\_VIEWS_{i,t-1}$	552	0.80	1.13	0.33	3.98
$REM\_PLF\_LN\_VIEWS_{i,t-2}$	552	0.75	0.83	0.08	1.00
$INVESTIG\_ANNCT._{i,t}$	552	0.07	0.83	0.76	38.69
Panel B. Matching on Year, Industry, and Propensity Score Using Litigation Determinants					
$FPS_{i,t}$	554	0.26	0.29	0.03	0.94
$\ln(ASSETS_{i,t-1})$	554	6.98	7.12	0.14	0.99
$SALES\_GR_{i,t-1}$	554	0.13	0.12	−0.01	−0.35
$CAR_{i,t-1}$	554	−0.02	−0.03	−0.01	−0.33
$RETURN\_SKEW_{i,t-1}$	554	0.24	0.21	−0.03	−0.58
$RETURN\_VOL_{i,t-1}$	554	0.14	0.13	−0.01	−1.31
$SHARE\_TURN_{i,t-1}$	554	3.23	3.56	0.33	1.77
$TOP\_PLF\_LN\_VIEWS_{i,t}$	554	0.44	2.02	1.58	19.28
$TOP\_PLF\_LN\_VIEWS_{i,t-1}$	554	0.37	0.55	0.18	2.85
$TOP\_PLF\_LN\_VIEWS_{i,t-2}$	554	0.20	0.28	0.08	1.79
$REM\_PLF\_LN\_VIEWS_{i,t}$	554	0.76	3.08	2.32	29.71
$REM\_PLF\_LN\_VIEWS_{i,t-1}$	554	0.82	1.12	0.30	3.65
$REM\_PLF\_LN\_VIEWS_{i,t-2}$	554	0.66	0.82	0.16	2.20
$INVESTIG\_ANNCT._{i,t}$	554	0.05	0.83	0.78	42.01

inconsistent with them having high litigation risk similar to that of the sued firms. As shown in Panel B, when using propensity-score matching to estimate litigation probability using covariates from Kim and Skinner (2012) and then matching within industry-year, we find similar inferences with year $t-2$ also being significantly different. The persistent and significant difference over multiple years between sued and matched firms demonstrates deficiencies in existing research designs and validates that our measure captures persistent plaintiff-lawyer scrutiny. Thus, absent matching on plaintiff-lawyer views, it is harder to ascertain whether findings are attributable to the variable of interest or litigation risk.

B. Predicting Market Outcomes

Because plaintiff-lawyer views predict future realized litigation risk, we expect they may also predict future market outcomes and other variables of interest to finance researchers. We focus on the ability to predict stock returns for two main reasons. First, plaintiff-lawyer views likely have important predictive ability in this

https://doi.org/10.1017/S0022109022001508 Published online by Cambridge University Press

setting because they predate the public bad news revelation at the class period end and litigation filings, which are associated with significant stock price drops (Karpoff et al. (2008), Gande and Lewis (2009)). More broadly, plaintiffs' lawyers seem to possess material adverse firm information that is not widely known, so we would not expect such information to be fully impounded into stock prices. Second, stock performance is one of the most researched areas in finance, and returns affect numerous firm outcomes that are also important research topics. Thus, if plaintiff-lawyer views predict future returns, this demonstrates they have myriad potential applications for future research.

For these tests, we switch from examining firm years with fiscal years 2012–2016 to examining firm quarters as we believe it is more likely returns will impound plaintiff-lawyer private information in the next quarter, rather than the next year. Our sample contains 91,136 firm-quarter observations as shown in Panel C of Table A.2 in the Supplementary Material. We examine future abnormal returns by estimating the following OLS regression:

$$
\begin{aligned}
(4) \qquad \mathrm{BHAR}_{i,t+1} = {} & \beta_0 + \beta_1 \ln(\mathrm{VIEWS}_{i,t}) + \beta_2 \ln(\mathrm{MVE}_{i,t}) \\
& + \beta_3 \ln(\mathrm{BOOK\_TO\_MARKET}_{i,t}) + \beta_4 \ln(\mathrm{TURNOVER}_{i,t}) \\
& + \beta_5 \mathrm{ALPHA}_{i,t} + \beta_6 \mathrm{INSTIT\_OWN}_{i,t} + \beta_7 \mathrm{NASDAQ}_{i,t} \\
& + \mathrm{FF48\,FE} + \mathrm{Fiscal\ Year} - \mathrm{Quarter\,FE} + \varepsilon_{i,t+1},
\end{aligned}
$$

where $\mathrm{BHAR}_{i,t+1}$ is buy-and-hold abnormal returns for quarter $t+1$ after using a Fama–French 4-factor model to estimate expected returns. We control for size ($\ln(\mathrm{MVE}_{i,t})$), book-to-market ($\ln(\mathrm{BOOK\_TO\_MARKET}_{i,t})$), share turnover ($\ln(\mathrm{TURNOVER}_{i,t})$), alpha ($\mathrm{ALPHA}_{i,t}$), institutional ownership ($\mathrm{INSTIT\_OWN}_{i,t}$), and an indicator variable for whether a firm trades on NASDAQ ($\mathrm{NASDAQ}_{i,t}$) (Loughran and McDonald (2011)), as well as industry and fiscal year-quarter fixed effects. Alternatively, we also estimate equation (4) after disaggregating plaintiff-lawyer views as has been done in prior analyses (i.e., $\mathrm{TOP\_PLF\_LN\_VIEWS}_{i,t}$ and $\mathrm{REM\_PLF\_LN\_VIEWS}_{i,t}$).

Panel A of Table 8 presents the results. Columns 1 and 2 show results for all quarterly observations in our sample period. The coefficient of $\ln(\mathrm{VIEWS}_{i,t})$ ($\mathrm{TOP\_PLF\_LN\_VIEWS}_{i,t}$) is negative and significant in column 1 (2) ($p < 0.05$), so higher quarterly plaintiff-lawyer views are associated with more negative returns in the next quarter. To ensure the findings are not restricted to the subset of firms that are sued or driven by bad news being released in the current quarter that could drive both current quarter views and next quarter returns, in columns 3–4 (5–6) we reestimate the model after eliminating quarters in which litigation is filed (class period end occurs for future litigation). We obtain similar inferences.

We next adjust equation (4) to examine the relation between plaintiff-lawyer views in quarter t and return volatility in quarter $t+1$ ($\mathrm{RETURN\_VOL}_{i,t+1}$). Litigation often includes allegations of excessively optimistic guidance or overstated earnings (Karpoff et al. (2008), Donelson et al. (2021a)), so we believe that higher plaintiff-lawyer views likely imply increased uncertainty about firms' fundamentals. We also expect that the "attention-grabbing" bad news events associated with plaintiff-lawyer views may induce higher trading, particularly by retail

TABLE 8

Predicting Future Market Outcomes

Table 8 presents results examining whether the current quarter's plaintiff-lawyer views are associated with future stock market returns by estimating equation (4) in the firm-quarter sample from 2012 to 2016 as shown in Panel C of Table A.2 in the Supplementary Material. Panel A examines next quarter's buy-and-hold abnormal returns using a Fama–French 4-factor model to calculate expected returns, and Panel B examines next quarter's daily return volatility. Columns 1, 3, and 5 examine total plaintiff-lawyer views ($\ln(\text{VIEWS}_{i,t})$); columns 2, 4, and 6 examine disaggregated plaintiff-lawyer views ($\text{TOP\_PLF\_LN\_VIEWS}_{i,t}$ and $\text{REM\_PLF\_LN\_VIEWS}_{i,t}$). Columns 1–2 examine all firm quarters. Columns 3–4 (5–6) exclude quarters in which litigation is filed (class period end occurs for subsequent litigation). For readability, we scale plaintiff-lawyer views by 100. To minimize the influence of outliers, all continuous variables are winsorized at the 1% and 99% levels. Refer to Appendix A for variable definitions.

Panel A. Future Returns

	$\text{BHAR}_{i,t+1}$					
	1	2	3	4	5	6
$\ln(\text{VIEWS}(/100)_{i,t})$	−0.29***		−0.33***		−0.33***	
	(0.10)		(0.10)		(0.10)	
$\text{TOP\_PLF\_LN\_VIEWS}\ (/100)_{i,t}$		−0.42**		−0.48**		−0.50**
		(0.21)		(0.21)		(0.21)
$\text{REM\_PLF\_LN\_VIEWS}\ (/100)_{i,t}$		−0.14		−0.17		−0.17
		(0.11)		(0.11)		(0.11)
$\ln(\text{MVE}_{i,t})$	0.00***	0.00***	0.00***	0.00***	0.00***	0.00***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
$\ln(\text{BOOK\_TO\_MARKET}_{i,t})$	−0.02***	−0.02***	−0.02***	−0.02***	−0.02***	−0.02***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
$\ln(\text{TURNOVER}_{i,t})$	−0.02***	−0.02***	−0.02***	−0.02***	−0.02***	−0.02***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
$\text{ALPHA}_{i,t}$	−57.88***	−57.89***	−57.96***	−57.97***	−57.93***	−57.94***
	(0.62)	(0.62)	(0.62)	(0.62)	(0.62)	(0.62)
$\text{INSTIT\_OWN}_{i,t}$	0.04***	0.04***	0.04***	0.04***	0.04***	0.04***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
$\text{NASDAQ}_{i,t}$	0.01***	0.01***	0.01***	0.01***	0.01***	0.01***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
INTERCEPT	−0.07***	−0.07***	−0.08***	−0.08***	−0.08***	−0.08***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
FF48 FE	Yes	Yes	Yes	Yes	Yes	Yes
Fyear × Qtr FE	Yes	Yes	Yes	Yes	Yes	Yes
No. of obs.	91,136	91,136	90,426	90,426	90,614	90,614
R^2	0.195	0.195	0.195	0.195	0.195	0.195

Panel B. Future Return Volatility

	$\text{RETURN\_VOL}_{i,t+1}$					
	1	2	3	4	5	6
$\ln(\text{VIEWS}\ (/100)_{i,t})$	0.08***		0.08***		0.08***	
	(0.01)		(0.01)		(0.01)	
$\text{TOP\_PLF\_LN\_VIEWS}\ (/100)_{i,t}$		0.13***		0.14***		0.13***
		(0.02)		(0.02)		(0.02)
$\text{REM\_PLF\_LN\_VIEWS}\ (/100)_{i,t}$		0.05***		0.05***		0.05***
		(0.01)		(0.01)		(0.01)
$\ln(\text{MVE}_{i,t})$	−0.00***	−0.00***	−0.00***	−0.00***	−0.00***	−0.00***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
$\ln(\text{BOOK\_TO\_MARKET}_{i,t})$	0.00***	0.00***	0.00***	0.00***	0.00***	0.00***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
$\ln(\text{TURNOVER}_{i,t})$	0.00***	0.00***	0.00***	0.00***	0.00***	0.00***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
$\text{ALPHA}_{i,t}$	−1.23***	−1.23***	−1.23***	−1.22***	−1.23***	−1.22***
	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
$\text{INSTIT\_OWN}_{i,t}$	−0.01***	−0.01***	−0.01***	−0.01***	−0.01***	−0.01***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
$\text{NASDAQ}_{i,t}$	−0.00***	−0.00***	−0.00***	−0.00***	−0.00***	−0.00***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
INTERCEPT	0.06***	0.06***	0.06***	0.06***	0.06***	0.06***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
FF48 FE	Yes	Yes	Yes	Yes	Yes	Yes
Fyear × Qtr. FE	Yes	Yes	Yes	Yes	Yes	Yes
No. of obs.	91,136	91,136	90,426	90,426	90,614	90,614
R^2	0.439	0.439	0.438	0.438	0.438	0.438

investors. Thus, building on prior research (e.g., Brandt, Brav, Graham, and Kumar (2010), Zhang (2010), and Chen, Huang, and Jha (2012)), we expect views to be positively associated with future volatility.

Panel B of Table 8 presents the results. We find both types of views are positively associated with future volatility ($p < 0.01$). Collectively, the results in Table 8 show plaintiff-lawyer views help predict future market outcomes. Moreover, we believe they may also be useful in nonlitigation-focused research as a proxy for otherwise unobservable firm risk-taking and agency costs, similar to prior research using D&O insurance data (e.g., Core (1997), Core (2000), Chalmers et al. (2002), Boyer and Stern (2014), and Cao and Narayanamoorthy (2014)).

V. Additional Analyses

A. Determinants of Plaintiff-Lawyer Scrutiny

The strong predictive ability of plaintiff-lawyer views raises the question of *what* determines which firms are scrutinized. Thus, we examine firm events and characteristics that may increase or decrease *contemporaneous* plaintiff-lawyer attention. This evidence provides insight into not only the types of firms that plaintiffs' lawyers may persistently scrutinize but also which discretionary factors managers may change to reduce scrutiny. It also provides insight into the relative mix of observable and unobservable information that plaintiffs' lawyers use to best allocate their monitoring resources. We examine the determinants in a multiple regression setting by estimating the following OLS regression:

$$\begin{aligned}(5)\quad \text{TOP\_PLF\_LN\_VIEWS/REM\_PLF\_LN\_VIEWS} = {} & \beta_0 + \beta_{1-4}\text{Accounting Events} \\ & + \beta_{5-6}\text{Personnel Events} + \beta_{7-8}\text{Disclosure} \\ & + \beta_{9-10}\text{Earnings Characteristics} \\ & + \beta_{11-12}\text{Visibility} + \beta_{13-16}\text{Complexity} \\ & + \beta_{17-18}\text{External Monitors} \\ & + \beta_{19-22}\text{Market Turmoil} \\ & + \beta_{23-27}\text{High Risk Industries} + \text{Year FE} + \varepsilon.\end{aligned}$$

We separately estimate equation (5) for the top and remaining plaintiffs' lawyers because their case selection process may differ. While we expect many of these determinants to have a similar effect on both types of lawyers, we expect that remaining plaintiffs' lawyers are more likely to use relatively low-quality external signals, due to a lack of internal resources and expertise (Badawi and Webber (2015)). We also expect top plaintiffs' lawyers may rely more on unobservable information obtained through their superior resources and expertise.

Accounting Events is a vector of indicator variables set to 1 when announcements of the following adverse accounting-related events occur in year t: auditor changes (AUDITOR_CHANGE_ANNCT$_{i,t}$), major restatement (MAJOR_RESTATE_ANNCT$_{i,t}$), nontimely SEC filings (NONTIMELY_FILING_ANNCT$_{i,t}$), and internal control weaknesses (ICWs, ICW_ANNCT$_{i,t}$). For adverse personnel-related events (PERSONNEL_EVENTS), we include indicator variables set to 1 for CEO and CFO turnover in the year (CEO_TURNOVER$_{i,t}$ and CFO_TURNOVER$_{i,t}$, respectively). We expect these events are positively associated with scrutiny from both types of law firms.

Given the extensive literature on disclosure and litigation (e.g., Francis et al. (1994), Field et al. (2005), and Rogers and Van Buskirk (2009)), DISCLOSURE includes the number of voluntary 8-K filings in the year (Bourveau, Lou, and Wang (2018), He and Plumlee (2020) (ln(VOLUNT_8-Ks$_{i,t}$)), and an indicator variable set to 1 if the firm provides earnings warnings (EARN_WARN_ANNCT$_{i,t}$). Because of mixed prior findings, we do not form predictions.

For most remaining variables, we use prior year values. We use lagged values from financial statements because they are publicly issued in year *t*. In other words, the timing of our independent variables is largely set to match the contemporaneous plaintiff-lawyer views that they may affect. We also continue to use lagged stock market values to match Table 6. However, we show in Tables A.38 and A.39 in the Supplementary Material that we obtain largely similar results if we instead use year t or $t-1$ values for all independent variables.

Earnings Characteristics is composed of indicator variables set to 1 if the firm discloses quarterly non-GAAP EPS with positive adjustments in the year (POSITIVE_NON-GAAP_ADJ$_{i,t}$) and if prior year discretionary accruals are positive (POSITIVE_DISC_ACCR$_{i,t-1}$). We expect a positive relation with plaintiff-lawyer scrutiny and positive non-GAAP adjustments, as they could indicate attempts to mislead. However, it is less clear how POSITIVE_DISC_ACCR$_{i,t-1}$ will affect scrutiny. Companies manipulate earnings to maximize stock prices, which may increase future litigation (DuCharme, Malatesta, and Sefcik (2004)). However, any significant relation would require lawyers to compare accruals across clients to calculate discretionary accruals, and discretionary accruals are not associated with accounting-related cases (Donelson et al. (2021a)).

Visibility is composed of size (ln(ASSETS$_{i,t-1}$)) and firm age (ln(AGE$_{i,t-1}$)). We expect ln(ASSETS$_{i,t-1}$) to be positively associated with all plaintiff-lawyer scrutiny, as it is positively correlated with both potential damages and the size of the affected class of plaintiffs. The prediction is less clear for ln(AGE$_{i,t-1}$) because both young and mature firms receive extensive media coverage and scrutiny. *Complexity* includes indicator variables set to 1 for firms with multiple segments (MULTISEGMENTS$_{i,t-1}$) and multinational operations (MULTINATIONAL$_{i,t-1}$), and should be positively correlated with plaintiff-lawyer scrutiny, as those firms are more prone to misreporting (Peterson (2012)). We also include indicator variables set to 1 for firm years with losses (LOSS$_{i,t-1}$) and no dividends (NO_DIVIDEND_PAID$_{i,t-1}$). Both variables may be positively associated with plaintiff-lawyer views because they increase information asymmetry (Khang and King (2006)) and Arena and Julio (2023) argue firms with high litigation risk decrease dividends for cash flexibility. However, the indirect nature of LOSS$_{i,t-1}$ in particular, relative to even the appearance of misconduct, may make it less useful to top lawyers.

External Monitors includes institutional ownership (INSTIT_OWN$_{i,t-1}$) and an indicator variable for having a top external auditor (BIG4$_{i,t-1}$). The presence of a top auditor may reduce plaintiff-lawyer scrutiny if viewed as substitute monitor that reduces misreporting, consistent with the SEC's approach (Holzman, Marshall, and Schmidt (2020)). However, because Donelson et al. (2021a) report that top-plaintiff lawyers do not consider firms' external auditor, this relation may be

https://doi.org/10.1017/S0022109022001508 Published online by Cambridge University Press

concentrated among remaining law firms. Alternatively, $\text{INSTIT\_OWN}_{i,t-1}$ may relate positively to all plaintiff-lawyer scrutiny because, the more institutional investors, the more plaintiffs' lawyers may monitor such firms for potential misconduct (Cheng, Huang, Li, and Lobo (2010)). Finally, we include *Market Turmoil* using the relevant Kim and Skinner (2012) variables (i.e., $\text{CAR}_{i,t-1}$, $\text{RETURN\_VOL}_{i,t-1}$, $\text{RETURN\_SKEW}_{i,t-1}$, and $\text{SHARE\_TURN}_{i,t-1}$) and indicator variables for *High-Risk Industries*, which should be associated with plaintiff-lawyer scrutiny in similar ways to their association with litigation filings. We include year-fixed effects to control for macroeconomic factors. See Appendix A for variable definitions.

Column 1 of Table 9 shows results from estimating equation (5) examining $\text{TOP\_PLF\_LN\_VIEWS}_{i,t-1}$. Despite the extensive vector of bad news events and firm characteristics, the R^2 indicates it only explains 9.2% of the variation in $\text{TOP\_PLF\_LN\_VIEWS}_{i,t-1}$. Coupled with the persistence of plaintiff-lawyer views, this is consistent with views being largely driven by relatively persistent unobservable factors. Most variables are statistically significant with signs consistent with our predictions. However, some relations or lack thereof merit discussion.

Both $\ln(\text{VOLUNT\_8-Ks}_{i,t})$ and $\text{EARN\_WARN\_ANNCT}_{i,t}$ are positively associated with plaintiff-lawyer views ($p < 0.05$), providing new insights into the literature on the effect of disclosure on litigation. Recent research generally finds timely disclosure reduces litigation (e.g., Field et al. (2005), Donelson et al. (2012), and Billings and Cedergren (2015)), but our results indicate concerns that disclosure triggers plaintiff-lawyer scrutiny remain valid (see Skinner (1997)). Although firms concerned with facing litigation may cease dividends to preserve cash flexibility (Arena and Julio (2023)), the positive relation with $\text{NO\_DIVIDEND\_PAID}_{i,t-1}$ indicates such actions may exacerbate litigation risk and ultimately constrain cash holdings ($p < 0.01$). We find $\text{ICW\_ANNCT}_{i,t}$ does not affect scrutiny, potentially because ICW allegations are not associated with large settlements (Cornerstone Research (2021)), nor does $\text{LOSS}_{i,t-1}$, consistent with scrutiny not being driven by negative earnings. Similarly, $\text{BIG4}_{i,t-1}$ does not affect scrutiny, consistent with Donelson et al. (2021a). Finally, we find that $\text{CAR}_{i,t-1}$ and $\text{RETURN\_SKEW}_{i,t-1}$ ($\text{RETURN\_VOL}_{i,t-1}$ and $\text{SHARE\_TURN}_{i,t-1}$) are negatively (positively) associated with future plaintiff-lawyer scrutiny ($p < 0.01$). Thus, just as plaintiff-lawyer views predict future market outcomes in Table 8, consistent with the market eventually responding to such scrutiny, unusual market outcomes also predict scrutiny.[7]

Column 2 presents the results from estimating equation (5) examining $\text{REM\_PLF\_LN\_VIEWS}_{i,t-1}$. In most aspects, the results resemble those from column 1, with a few notable exceptions. Consistent with expectations, the R^2 indicates the model can explain 17.5% of the variation in $\text{REM\_PLF\_LN\_VIEWS}_{i,t-1}$, nearly double that in column 1. Thus, other plaintiffs' lawyers appear to rely *relatively* more on observable information, although most of their views are also driven by unobservable factors. Other plaintiffs' lawyers appear to scrutinize

[7]We use annual market variables in this analysis to correspond with Table 6 but find similar inferences if we use quarterly market variables in a specification similar to Table A.37 in the Supplementary Material.

TABLE 9

Determinants of Plaintiff-Lawyer Views

Column 1 (2) in Table 9 presents results based on estimating equation (5) using OLS examining determinants of scrutiny by top (remaining) plaintiffs' lawyers in the firm-year sample from 2012 to 2016 as shown in Panel B of Table A.2 in the Supplementary Material. To minimize the influence of outliers, all continuous variables are winsorized at the 1% and 99% levels. Standard errors appear in parentheses and are adjusted for clustering at the company level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. Refer to Appendix A for variable definitions.

	$TOP\_PLF\_LN\_VIEWS_{i,t}$	$REM\_PLF\_LN\_VIEWS_{i,t}$
	1	2
Accounting Events		
$AUDITOR\_CHANGE\_ANNCT_{i,t}$	0.01 (0.03)	0.03 (0.04)
$MAJOR\_RESTATE\_ANNCT_{i,t}$	0.44*** (0.08)	0.46*** (0.09)
$NON\text{-}TIMELY\_FILING\_ANNCT_{i,t}$	0.12*** (0.03)	0.21*** (0.05)
$ICW\_ANNCT_{i,t}$	0.04 (0.03)	0.10** (0.04)
Personnel Events		
$CEO\_TURNOVER_{i,t}$	0.10*** (0.02)	0.15*** (0.03)
$CFO\_TURNOVER_{i,t}$	0.06*** (0.02)	0.08*** (0.03)
Disclosure		
$\ln(VOLUNT\_8\text{-}Ks_{i,t})$	0.06*** (0.01)	0.21*** (0.01)
$EARN\_WARN\_ANNCT_{i,t}$	0.05** (0.02)	0.06** (0.03)
Earnings Characteristics		
$POSITIVE\_NON\text{-}GAAP\_ADJ_{i,t}$	0.02 (0.02)	0.03 (0.02)
$POSITIVE\_DISC\_ACCR_{i,t-1}$	−0.02 (0.01)	−0.04** (0.02)
Visibility		
$\ln(ASSETS_{i,t-1})$	0.08*** (0.01)	0.18*** (0.01)
$\ln(AGE_{i,t-1})$	−0.05*** (0.01)	0.01 (0.01)
Complexity		
$MULTISEGMENTS_{i,t-1}$	−0.03 (0.02)	−0.04* (0.02)
$MULTINATIONAL_{i,t-1}$	0.05*** (0.02)	0.12*** (0.02)
$LOSS_{i,t-1}$	0.03 (0.02)	0.08*** (0.02)
$NO\_DIVIDEND\_PAID_{i,t-1}$	0.07*** (0.02)	0.08*** (0.02)
External Monitors		
$BIG4_{i,t-1}$	−0.02 (0.02)	−0.13*** (0.03)
$INSTIT\_OWN_{i,t-1}$	0.03 (0.03)	−0.03 (0.04)
Market Turmoil		
$CAR_{i,t-1}$	−0.05*** (0.02)	−0.04** (0.02)
$RETURN\_VOL_{i,t-1}$	0.64*** (0.13)	0.87*** (0.18)
$RETURN\_SKEW_{i,t-1}$	−0.02*** (0.01)	−0.03*** (0.01)
$SHARE\_TURN_{i,t-1}$	0.03*** (0.00)	0.04*** (0.01)

(continued on next page)

TABLE 9 (continued)

Determinants of Plaintiff-Lawyer Views

	TOP_PLF_LN_VIEWS$_{i,t}$	REM_PLF_LN_VIEWS$_{i,t}$
	1	2
High-Risk Industries		
BIOTECH$_{i,t}$	0.15***	0.46***
	(0.04)	(0.07)
COMP_HARDWARE$_{i,t}$	0.12	0.25***
	(0.08)	(0.10)
ELECTRONICS$_{i,t}$	−0.03	0.01
	(0.03)	(0.04)
RETAIL$_{i,t}$	0.15***	0.14***
	(0.04)	(0.05)
COMP_SOFTWARE$_{i,t}$	0.11***	0.15***
	(0.04)	(0.05)
INTERCEPT	−0.65***	−1.19***
	(0.06)	(0.07)
Year FE	Yes	Yes
No. of obs.	17,179	17,179
R^2	0.092	0.175

more relatively low-quality signals. For example, both ICW_ANNCT$_{i,t}$ and LOSS$_{i,t-1}$ are associated with more scrutiny ($p < 0.05$). Further, BIG4$_{i,t-1}$ is associated with lower scrutiny ($p < 0.01$), similar to SEC scrutiny (Holzman et al. (2020)). Additionally, POSITIVE_DISC_ACCR$_{i,t-1}$ is associated with lower scrutiny ($p < 0.05$), suggesting firms may be able to manipulate earnings to avoid attracting scrutiny from other plaintiffs' lawyers. These differences are consistent with these law firms using these variables as screening mechanisms to winnow firms to those most likely to be risky due to fewer resources. Collectively, these findings provide insight into the factors attracting plaintiff-lawyer scrutiny.

B. Predicted Plaintiff-Lawyer EDGAR Views

Despite its many benefits, the use of plaintiff-lawyer views as a measure of litigation risk faces one main challenge – the SEC only periodically releases EDGAR log files, and the latest one is updated through June 2017.[8] Thus, if plaintiff-lawyer views are unavailable, we propose the use of predicted plaintiff-lawyer views (PRED_LN_VIEWS$_{i,t-1}$). In untabulated tests, the correlation between predicted top and remaining plaintiff-lawyer views exceeds 93%, so we use predicted *total* plaintiff-lawyer views to avoid multicollinearity. We estimate these using the determinants from equation (5) after excluding *Market Turmoil*, to reduce data requirements, and year-fixed effects. Appendix C shows the results.

We first calculate PRED_LN_VIEWS$_{i,t-1}$ in the period after our sample (2017–2019) because we can also identify plaintiff-lawyer investigations in these

[8]Another challenge is that the ARIN Whois Bulk data set only lists the current owner of an IP address, resulting in noisier measures the further back one looks. That is, if a law firm unregisters an IP, researchers cannot identify that an IP belongs to the firm in earlier periods. However, we can identify IP addresses for a plaintiffs' law firm involved in most securities class actions since 2001, so we do not expect this to be that problematic for classifying firms as having high or low litigation risk (see Figure A.9 in the Supplementary Material).

TABLE 10

Using Predicted Plaintiff-Lawyer Views to Predict Realized Litigation Risk

Table 10 presents results examining *predicted* plaintiff-lawyer views as a proxy for ex ante litigation risk. Columns 1–2 and 5–6 use realized litigation filings ($SUED_{i,t}$), whereas columns 3–4 use realized litigation filings supplemented with investigation announcements ($SUED\_INV_{i,t}$). We use a later (an earlier) time period in columns 1–4 (5–6) to assess the predictive ability of our model in periods where it was not estimated. Columns 1–4 use fiscal years 2017 up to those ending on Dec. 31, 2019; columns 5–6 use fiscal years 2007 to 2011. We do not use investigation announcements for the earlier period as those announcements were not common prior to 2012. Columns 1, 3, and 5 use predicted ex ante litigation risk from Kim and Skinner (2012). Columns 2, 4, and 6 use predicted plaintiff-lawyer scrutiny obtained using Appendix C. All models are estimated using logistic estimation. Both in-sample and out-of-sample model performance metrics are reported. The out-of-sample statistics are calculated using the "*K*-fold" cross validation procedure similar to Dimmock and Gerken (2012), Kim and Skinner (2012), and Larcker and Zakolyukina (2012) with a 10-fold cross validation procedure. Precision is calculated as the percentage of predicted positive cases that are true positives. Sensitivity is calculated as the percentage of true positive cases correctly identified. Specificity is calculated as the percentage of true negative cases correctly identified. We classify cases as being predicted positive if the predicted probability is in the top 3% of all probabilities. To minimize the influence of outliers, all continuous variables are winsorized at the 1% and 99% levels. Standard errors appear in parentheses and are adjusted for clustering at the company level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. Refer to Appendix A for variable definitions.

	$SUED_{i,t}$		$SUED\_INV_{i,t}$		$SUED_{i,t}$	
	1	2	3	4	5	6
$PRED\_LN\_VIEWS_{i,t-1}$		0.93***		1.08***		0.91***
		(0.14)		(0.12)		(0.14)
$EX\_ANTE\_LIT\_RISK_{i,t}$	0.18***		0.18***		0.11***	
	(0.02)		(0.02)		(0.02)	
$FPS_{i,t}$	0.29**		0.36***			
	(0.11)		(0.09)			
$\ln(ASSETS_{i,t-1})$	0.13***	0.00	0.15***	−0.00	0.23***	0.11***
	(0.02)	(0.03)	(0.02)	(0.03)	(0.02)	(0.03)
$BIOTECH_{i,t}$		0.43**		0.44***	1.06***	0.65***
		(0.20)		(0.16)	(0.17)	(0.19)
$COMP\_HARDWARE_{i,t}$		0.10		0.06	0.25	−0.09
		(0.39)		(0.32)	(0.39)	(0.39)
$ELECTRONICS_{i,t}$		0.11		−0.19	0.12	0.08
		(0.22)		(0.21)	(0.19)	(0.19)
$RETAIL_{i,t}$		−0.22		−0.20	0.04	−0.18
		(0.25)		(0.20)	(0.23)	(0.23)
$COMP\_SOFTWARE_{i,t}$		−0.17		0.01	0.46**	0.20
		(0.22)		(0.18)	(0.23)	(0.23)
INTERCEPT	−3.85***	−3.86***	−3.35***	−3.33***	−5.47***	−5.44***
	(0.18)	(0.19)	(0.14)	(0.15)	(0.19)	(0.20)
No. of obs.	9,133	9,133	9,133	9,133	17,065	17,065
Pseudo R^2	0.033	0.030	0.038	0.043	0.041	0.046
In-sample						
Pred. posit.	274	274	274	274	512	512
Correct pred. posit.	35	44	60	71	45	51
Precision	0.128	0.161	0.219	0.259	0.088	0.100
Sensitivity	0.076	0.095	0.073	0.086	0.116	0.132
False negative rate	0.924	0.905	0.927	0.914	0.884	0.868
Specificity	0.972	0.973	0.974	0.976	0.972	0.972
AUC	0.649	0.628	0.647	0.648	0.667	0.674
Out-of-sample						
Precision	0.114	0.146	0.221	0.264	0.085	0.098
Sensitivity	0.069	0.089	0.075	0.090	0.113	0.132
False negative rate	0.931	0.911	0.925	0.910	0.887	0.868
Specificity	0.971	0.972	0.974	0.975	0.971	0.972
AUC	0.645	0.621	0.645	0.644	0.656	0.666

years to test how well it performs *outside* the period in which the model was calibrated. We first predict $SUED_{i,t}$, and the results are shown in columns 1 and 2 of Table 10. Column 1 uses the predicted litigation risk using the coefficients in Kim and Skinner (2012) ($EX\_ANTE\_LIT\_RISK_{i,t}$), $\ln(ASSETS_{i,t-1})$, and $FPS_{i,t}$; column 2 uses a similar model but instead uses $PRED\_LN\_VIEWS_{i,t-1}$ as a measure of litigation risk and multiple high industry risk indicator variables. Our model has the highest precision and sensitivity with in-sample (out-of-sample) performance improvements of over 25% (20%) relative to the model using EX_ANTE_LIT_RISK$_{i,t}$. Columns 3 and 4 report results of predicting SUED_INV$_{i,t}$, and our model similarly obtains in-sample (out-of-sample) performance improvements of around 18% (25%). Additionally, we note that the precision of both our model and the model with predicted values from Kim and Skinner (2012) achieve increases in model precision of over 60% with minor changes in sensitivity due to the inclusion of plaintiff-lawyer investigations and the resulting decrease in false positives.

We also calculate PRED_LN_VIEWS$_{i,t-1}$ for the years 2007–2011 to demonstrate the predictive ability before the period the model was calculated and predict SUED$_{i,t}$ in columns 5 and 6. While these tests likely understate the performance of our model without data on plaintiff-lawyer investigations, we continue to obtain the highest precision and sensitivity using our model. Specifically, we obtain in-sample (out-of-sample) precision and sensitivity improvements of over 10% (25%), relative to the model using EX_ANTE_LIT_RISK$_{i,t}$.

Alternatively, we can also directly compare the precision and sensitivity of observations in the top 3% of EX_ANTE_LIT_RISK$_{i,t}$ versus those for PRED_LN_VIEWS$_{i,t-1}$ as standalone variables. PRED_LN_VIEWS$_{i,t-1}$ outperforms EX_ANTE_LIT_RISK$_{i,t}$ by even larger margins. There are minor performance decreases using only PRED_LN_VIEWS$_{i,t-1}$, relative to the model in Table 10, but large decreases using only EX_ANTE_LIT_RISK$_{i,t}$ (untabulated).

Thus, predicted plaintiff-lawyer views also improve predictive ability, relative to methods from prior research, despite similarly imposing fewer data constraints. Accordingly, while we advise researchers to use actual plaintiff-lawyer views when available, the use of predicted views will still reduce measurement error and bias in future research.

VI. Conclusion

"Litigation risk and lawsuits have significant long-lasting effects on the defendant firm, its executives and directors" (Arena and Ferris (2016), p. 8), but ex ante litigation risk is notoriously difficult to measure. We introduce a new measure of ex ante securities litigation risk based on plaintiff-lawyer views of firms' SEC filings. We validate that the measure captures plaintiff-lawyer scrutiny, and despite imposing fewer data requirements, lagged annual values improve predictions of litigation risk. Notably, this predictive ability is not due to plaintiff-lawyer views capturing the litigation process itself or related bad news events. We show our measure can reduce measurement error in common research designs and that its ex ante nature allows it to predict future stock returns and return volatility, demonstrating it likely has myriad uses for finance researchers. We also provide new economic insights into the determinants of plaintiff-lawyer scrutiny and demonstrate that predicted plaintiff-lawyer views can similarly improve predictions of litigation risk.

Collectively, our study has significant implications for the extensive cross-disciplinary research on securities litigation risk. Additionally, we introduce researchers to a new potential research topic: plaintiff-lawyer investigations. Finally, we demonstrate shortcomings in the use of the AUC as a performance metric in highly imbalanced data sets and suggest alternative and more diagnostic performance metrics, which can be used to better predict myriad rare events beyond litigation (e.g., bankruptcy, fraud, and extreme stock returns).

Appendix A. Variable Definitions

ALPHA$_{i,t}$: Alpha, based on a Fama–French 4-factor model. We estimate the model over days [−252, −1] relative to fiscal quarter's t end and require at least 60 observations of daily returns to estimate the model.

AUDITOR_CHANGE_ANNCT$_{i,t}$: An indicator variable set to 1 if there is an auditor change announcement during the fiscal year, and 0 otherwise.

BHAR$_{i,t+1}$: Buy-and-hold abnormal returns over the period [1, 60] days relative to fiscal quarter's t end. Expected returns are calculated using a Fama–French 4-factor model. We estimate the model over days [−252, −1] relative to fiscal quarter's t end and require at least 60 observations of daily returns to estimate the model.

BIG4$_{i,t}$: An indicator variable set to 1 if the firm is audited by one of the top 4 audit firms (i.e., the Big 4), and 0 otherwise.

BIOTECH$_{i,t}$: An indicator variable set to 1 if the firm's SIC code is between 2,833 and 2,836, and 0 otherwise.

CAR$_{i,t}$: Cumulative abnormal return during the fiscal year based on monthly returns.

CEO_TURNOVER$_{i,t}$: An indicator variable set to 1 if the firm announces CEO turnover during the fiscal year, and 0 otherwise. We use the Audit Analytics' Director and Officer Changes data set to maximize coverage. We exclude cases in which ACTION per Audit Analytics is set to "Appointed," "Retracted Resignation," "Reelected," "Change Misreported," "Nominated," "Returned to Position," or "Engaged."

CFO_TURNOVER$_{i,t}$: An indicator variable set to 1 if the firm announces CFO turnover during the fiscal year, and 0 otherwise. We use the Audit Analytics' Director and Officer Changes data set to maximize coverage. We exclude cases in which ACTION per Audit Analytics is set to "Appointed," "Retracted Resignation," "Reelected," "Change Misreported," "Nominated," "Returned to Position," or "Engaged."

COMP_HARDWARE$_{i,t}$: An indicator variable set to 1 if the firm's SIC code is between 3,570 and 3,577, and 0 otherwise.

COMP_SOFTWARE$_{i,t}$: An indicator variable set to 1 if the firm's SIC code is between 7,371 and 7,379, and 0 otherwise.

EARN_WARN_ANNCT$_{i,t}$: An indicator variable set to 1 if the firm provides earning warnings during the fiscal year per IBES guidance, and 0 otherwise. We rely on IBES guidance codes to identify earning warnings (i.e., cases in which the guidance code is equal to 1). The variable is set to 0 for firms missing IBES coverage.

ELECTRONICS$_{i,t}$: An indicator variable set to 1 if the firm's SIC code is between 3,600 and 3,674, and 0 otherwise.

EX_ANTE_LIT_RISK$_{i,t}$: Ex ante litigation risk per Kim and Skinner (2012) (i.e., model 3 of their Table 7). We use the log odds value (i.e., we do not convert the predicted value to probability).

$FPS_{i,t}$: Indicator variable set to 1 if the firm is in a high litigation risk industry as defined in Francis et al. (1994), and 0 otherwise. Specifically, it is set to 1 if the firm is in any of the following industries: biotechnology (SIC codes 2833–2836 and 8731–8734), computers (SIC codes 3570–3577 and 7370–7374), electronics (SIC codes 3600–3674), or retailing (SIC codes 5200–5961).

$ICW\_ANNCT_{i,t}$: An indicator variable set to 1 if the management (auditor) announces ineffective internal controls under Sarbanes–Oxley Act (SOX) Section 302 (404) during the fiscal year, and 0 otherwise.

$INSTIT\_OWN_{i,t}$: Proportion of institutional ownership as of fiscal year end. For Table 8, the variable is calculated as of fiscal quarter end. Missing values are set to 0.

$INVESTIG\_ANNCT._{i,t}$: An indicator variable set to 1 if an investigation by a plaintiffs' law firm is announced against the firm during the fiscal year, and 0 otherwise. To identify investigation announcements, we search for press release newswires on Factiva that contain: i) "announces investigation" or ii) "investigating" and "on behalf" and then match the targeted firms to our data set.

$\ln(AGE_{i,t})$: Natural logarithm of one plus the number of years since the firm first appeared on Compustat.

$\ln(ASSETS_{i,t})$: Natural logarithm of total assets, in millions, at the fiscal year end.

$\ln(BOOK\_TO\_MARKET_{i,t})$: Natural logarithm of the book-to-market ratio, calculated as the sum of quarter t total liabilities and market value of equity, scaled by total assets.

$\ln(DAMAGES_{i,t})$: Natural logarithm of the difference between the maximum market capitalization during the class period less the market capitalization the day following class period end. Market capitalization is calculated in actual dollar value.

$\ln(MVE_{i,t})$: Natural logarithm of the market value of equity, in millions, at the end of the fiscal year. For Table 8, the variable is calculated as of fiscal quarter end.

$\ln(SETTLEM_{i,t})$: Natural logarithm of the securities class action settlement amount.

$\ln(TURNOVER_{i,t})$: Natural logarithm of share turnover. Share turnover is defined as split-adjusted trading volume, scaled by shares outstanding during the first day of the fiscal quarter.

$\ln(VIEWS_{i,t})$: Natural logarithm of one plus the total number of EDGAR views by plaintiffs' law firms during the fiscal year. For Table 8, the variable is calculated as the total number of EDGAR views during the fiscal quarter. We exclude index and web crawler views.

$\ln(VIEWS_{i,[\text{Class End,Filing-1}]})$: Natural logarithm of one plus the total number of EDGAR views by plaintiffs' law firms from the class period end up to, but not including, the litigation filing day.

$\ln(VOLUNT\_8\text{-}Ks_{i,t})$: Natural logarithm of one plus the number of Form 8-Ks with Item Codes 2.02, 7.01, or 8.01 that the firm submitted on EDGAR during the fiscal year (see Bourveau et al. (2018), He and Plumlee (2020)). We also include amended filings.

$LOSS_{i,t}$: An indicator variable set to 1 if the firm's net income for the fiscal year is negative, and 0 otherwise.

$MAJOR\_RESTATE\_ANNCT_{i,t}$: An indicator variable set to 1 if a major restatement (i.e., disclosed via Item 4.02 in an 8-K) is announced during the fiscal year, and

0 otherwise. We use the filing date of the original restatement announcement, rather than the date of the 8-K filing with Item 4.02, which may occur on a subsequent date.

MULTISEGMENTS$_{i,t}$: An indicator variable set to 1 if the firm has more than one business segment in the fiscal year, and 0 otherwise. Missing values for the number of business segments are set to 1.

MULTINATIONAL$_{i,t}$: An indicator variable set to 1 if the firm has pretax foreign income in the fiscal year, and 0 otherwise.

NASDAQ$_{i,t}$: An indicator variable set to 1 if the firm is trading on NASDAQ as of fiscal quarter end; 0 otherwise.

NO_DIVIDEND_PAID$_{i,t}$: An indicator variable set to 1 if the firm did not pay dividends during the fiscal year; 0 otherwise.

NON-TIMELY_FILING_ANNCT$_{i,t}$: An indicator variable set to 1 if the firm submits a non-timely filing during the fiscal year; 0 otherwise.

POSITIVE_DISC_ACCR$_{i,t-1}$: An indicator variable set to 1 if the prior fiscal year's modified Jones discretionary accruals are positive as per Dechow, Sloan, and Sweeney (1995); 0 otherwise. We require at least 15 observations in a given year-SIC2 industry to calculate discretionary accruals. Accruals are calculated as net income minus cash flows from operations.

POSITIVE_NON-GAAP_ADJ$_{i,t}$: An indicator variable set to 1 if quarterly GAAP-reported EPS (epsfiq per Compustat) announced during the fiscal year is less than management-provided non-GAAP EPS as per Bentley et al. (2018); 0 otherwise. If the manager does not provide non-GAAP EPS, the variable is set to 0.

PRED_LN_VIEWS$_{i,t}$: Predicted number of EDGAR views by all plaintiffs' lawyers using the model from Appendix C.

RETAIL$_{i,t}$: An indicator variable set to 1 if the firm's SIC code is between 5200 and 5961; 0 otherwise.

REM_PLF_LN_VIEWS$_{i,t}$: Natural logarithm of one plus the total number of EDGAR views by plaintiffs' law firms not ranked among the top plaintiffs' law firms by Chambers and Partners as of June 2011 (see TOP_PLF_LN_VIEWS$_{i,t}$) during the fiscal year. For Table 8, the variable is calculated as the total number of EDGAR views during the fiscal quarter.

REM_PLF_LN_VIEWS$_{i,[\text{Class End,Filing-1}]}$: Natural logarithm of one plus the total number of EDGAR views by plaintiffs' law firms not ranked among the top plaintiffs' law firms by Chambers and Partners as of June 2011 (see TOP_PLF_LN_VIEWS$_{i,t}$) from the class period end up to, but not including, the litigation filing day.

REM_PLF_VIEWS > 0_IND$_{i,t}$: An indicator variable set to 1 if a plaintiffs' law firm not ranked among the top plaintiffs' law firms by Chambers and Partners as of June 2011 (see TOP_PLF_LN_VIEWS$_{i,t}$) accessed the firm's EDGAR filings at least once during the fiscal year; 0 otherwise.

RETURN_SKEW$_{i,t}$: Skewness of monthly raw returns during the fiscal year.

RETURN_VOL$_{i,t}$: Standard deviation of monthly returns during the fiscal year. For Table 8, the variable is calculated using daily returns over the fiscal quarter.

ROA$_{i,t}$: Net income scaled by total assets as of the end of the fiscal year.

https://doi.org/10.1017/S0022109022001508 Published online by Cambridge University Press

$SALES\_GR_{i,t-1}$: Sales growth is measured as the change in net sales from fiscal year t − 2 to t − 1, divided by total assets as of t − 2.

$SETTLED_{i,t}$: An indicator variable set to 1 if the firm is sued during the fiscal year and the lawsuit is eventually settled; 0 otherwise. We require the lawsuit to contain fraud allegations (i.e., alleging violations of Rule 10b-5).

$SHARE\_TURN_{i,t}$: Share turnover, defined as split-adjusted trading volume scaled by shares outstanding during the first month of the fiscal year.

$SUED_{i,t}$: An indicator variable set to 1 if a securities class action is filed against the firm during the fiscal year; 0 otherwise.

$SUED\_INV_{i,t}$: An indicator variable set to 1 if SUED $_{i,t}$ is equal to 1 or INVESTIG_ ANNCT. $_{i,t}$ is equal to 1 in the fiscal year; 0 otherwise.

$TOBINS\_Q_{i,t}$: The sum of the market value of common stock, preferred stock, and firm debt, scaled by total assets as of the end of the fiscal year. Preferred stock and debt are assumed to have a market value equal to book value. Missing values for debt are set equal to 0.

$TOP\_PLF\_LN\_VIEWS_{i,t}$: Natural logarithm of one plus the total number of EDGAR views by plaintiffs' law firms ranked among the top plaintiffs' law firms by Chambers and Partners as of June 2011 during the fiscal year. For Table 8, the variable is calculated as the total number of EDGAR views during the fiscal quarter. The list of top plaintiff' law firms includes: i) Bernstein, Litowitz, Berger & Grossmann LLP, ii) Grant and Eisenhofer, iii) Labaton Sucharow LLP, and iv) Robbins Geller Rudman & Dowd LLP.

$TOP\_PLF\_LN\_VIEWS_{i,[\text{Class End,Filing-1}]}$: Natural logarithm of one plus the total number of EDGAR views by plaintiffs' law firms ranked among the top plaintiffs' law firms by Chambers and Partners as of June 2011 (see $TOP\_PLF\_LN\_VIEWS_{i,t}$) from the class period end up to, but not including, the litigation filing day.

$TOP\_PLF\_VIEWS > 0\_IND_{i,t}$: An indicator variable set to 1 if any of the top plaintiffs' law firms (see $TOP\_PLF\_LN\_VIEWS_{i,t}$) by Chambers and Partners as of June 2011 accessed the firm's EDGAR filings at least once during the fiscal year; 0 otherwise.

TOT_UNQ_VIEW_FIRMS: Number of unique plaintiffs' law firms accessing EDGAR during the 20-day window preceding the litigation filing (i.e., t − 20 to t − 1).

TOT_UNQ_VIEW_FIRMS_IF_VIEWS > 0: Number of unique plaintiffs' law firms accessing EDGAR during the 20-day window preceding the litigation filing (i.e., t − 20 to t − 1), conditional on the presence of at least one view.

TOT_VIEWS: EDGAR views from plaintiffs' lawyers taking place in the 20-day window preceding the litigation filing (i.e., t − 20 to t − 1).

TOT_VIEWS_IF > 0: EDGAR views from plaintiffs' lawyers taking place in the 20-day window preceding the litigation filing (i.e., t − 20 to t − 1), conditional on the presence of at least one view.

VIEWS > 0_DUMMY: An indicator variable set to 1 if TOT_VIEWS is larger than 0; 0 otherwise.

$VIEWS > 0\_IND_{i,t}$: An indicator variable set to 1 if plaintiffs' lawyers accessed the firm's EDGAR filings at least once during the fiscal year; 0 otherwise.

Appendix B. Plaintiff-Lawyer Investigation Announcements

Appendix B presents excerpts from two examples of recent investigation announcements by plaintiffs' lawyers.

Pomerantz (2020)

NEW YORK, NY/ACCESSWIRE/September 1, 2020/Pomerantz LLP is investigating claims on behalf of investors of Galapagos NV ("Galapagos" or the "Company") (NASDAQ:GLPG). Such investors are advised to contact Robert S. Willoughby at newaction@pomlaw.com or 888-476-6529, ext. 7980.

Rosen Law Firm (2020)

Rosen Law Firm Announces Investigation of Securities Claims Against McDonald's Corporation – MCD

August 10, 2020 12:28 PM Eastern Daylight Time

NEW YORK--(BUSINESS WIRE)--Rosen Law Firm, a global investor rights law firm, announces an investigation of potential securities claims on behalf of shareholders of McDonald's Corp. (NYSE: MCD) resulting from allegations that McDonald's may have issued materially misleading business information to the investing public.

Appendix C. Determinants Model for Predicted Plaintiff-Lawyer Views

Appendix C presents results similar to Table 9 estimating equation (5) using OLS after i) excluding year-fixed effects and *Market Turmoil* and ii) replacing the dependent variable with total plaintiff-lawyer views (ln(VIEWS$_{i,t}$)). To minimize the influence of outliers, all variables are winsorized at the 1% and 99% levels. Standard errors appear in parentheses and are adjusted for clustering at the company level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. Refer to Appendix A for variable definitions.

	ln(VIEWS$_{i,t}$)
	1
Accounting Events	
AUDITOR_CHANGE_ANNCT$_{i,t}$	0.06 (0.04)
MAJOR_RESTATE_ANNCT$_{i,t}$	0.54*** (0.10)
NON-TIMELY_FILING_ANNCT$_{i,t}$	0.25*** (0.05)
ICW_ANNCT$_{i,t}$	0.13*** (0.04)
Personnel Events	
CEO_TURNOVER$_{i,t}$	0.18*** (0.04)
CFO_TURNOVER$_{i,t}$	0.10*** (0.03)
Disclosure	
ln(VOLUNT_8-Ks$_{i,t}$)	0.23*** (0.01)
EARN_WARN_ANNCT$_{i,t}$	0.10*** (0.03)
Earnings Characteristics	
POSITIVE_NON-GAAP_ADJ$_{i,t}$	0.04 (0.03)
POSITIVE_DISC_ACCR$_{i,t-1}$	−0.05** (0.02)

	$\ln(\text{VIEWS}_{i,t})$
	1
Visibility	
$\ln(\text{ASSETS}_{i,t-1})$	0.21***
	(0.01)
$\ln(\text{AGE}_{i,t-1})$	−0.05***
	(0.02)
Complexity	
$\text{MULTISEGMENTS}_{i,t-1}$	−0.07***
	(0.03)
$\text{MULTINATIONAL}_{i,t-1}$	0.13***
	(0.03)
$\text{LOSS}_{i,t-1}$	0.16***
	(0.03)
$\text{NO\_DIVIDEND\_PAID}_{i,t-1}$	0.18***
	(0.03)
External Monitors	
$\text{BIG4}_{i,t-1}$	−0.08***
	(0.03)
$\text{INSTIT\_OWN}_{i,t-1}$	−0.01
	(0.04)
High-Risk Industries	
$\text{BIOTECH}_{i,t}$	0.55***
	(0.07)
$\text{COMP\_HARDWARE}_{i,t}$	0.31***
	(0.11)
$\text{ELECTRONICS}_{i,t}$	0.06
	(0.05)
$\text{RETAIL}_{i,t}$	0.24***
	(0.06)
$\text{COMP\_SOFTWARE}_{i,t}$	0.21***
	(0.06)
INTERCEPT	−1.13***
	(0.07)
Year FE	No
N	17,179
R^2	0.162

Supplementary Material

To view supplementary material for this article, please visit http://doi.org/10.1017/S0022109022001508.

References

Abramowicz, M. "How Lawyers Compete." *Regulation*, 27 (2004), 38–39.

Acharya, V. V.; Y. Amihud; and S. T. Bharath. "Liquidity Risk of Corporate Bond Returns: Conditional Approach." *Journal of Financial Economics*, 110 (2013), 358–386.

Alexander, J. C. "Do the Merits Matter? A Study of Settlements in Securities Class Actions." *Stanford Law Review*, 43 (1991), 497–598.

An, Z.; C. Chen; V. Naiker; and J. Wang. "Does Media Coverage Deter Firms from Withholding Bad News? Evidence from Stock Price Crash Risk." *Journal of Corporate Finance*, 64 (2020), 1–25.

Angrist, J. D., and J.-S. Pischke. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Press (2009).

Arena, M., and S. Ferris. "A Survey of Litigation in Corporate Finance." *Managerial Finance*, 43 (2016), 4–18.

Arena, M., and B. Julio. "The Effects of Securities Class Action Litigation on Corporate Liquidity and Investment Policy." *Journal of Financial and Quantitative Analysis*, 50 (2015), 251–275.

Arena, M., and B. Julio. "Litigation Risk Management Through Corporate Payout Policy." *Journal of Financial and Quantitative Analysis*, 58 (2023), 148–174.

Armstrong, C.; T. Blackburne; and P. Quinn. "Are CEOs' Purchases More Profitable than they Appear?" *Journal of Accounting and Economics*, 71 (2021), 1–22.

Atanasov, V.; V. Ivanov; and K. Litvak. "Does Reputation Limit Opportunistic Behavior in the VC Industry? Evidence from Litigation Against VCs." *Journal of Finance*, 67 (2012), 2215–2246.

Avramov, D. "Stock Return Predictability and Model Uncertainty." *Journal of Financial Economics*, 64 (2002), 423–458.

Badawi, A. B., and D. H. Webber. "Does the Quality of the Plaintiffs' Law Firm Matter in Deal Litigation?" *Journal of Corporate Law*, 359 (2015), 101–133.

Baginski, S. P.; J. M. Hassell; and M. D. Kimbrough. "The Effect of Legal Environment on Voluntary Disclosure: Evidence from Management Earnings Forecasts Issued in U.S. and Canadian Markets." *Accounting Review*, 77 (2002), 25–50.

Baker, T., and S. J. Griffith. *Ensuring Corporate Misconduct: How Liability Insurance Undermines Shareholder Litigation*. Chicago, IL: University of Chicago Press (2010).

Baker, L. A.; M. A. Perino; and C. Silver. "Is the Price Right? An Empirical Study of Fee-Setting in Securities Class Actions." *Columbia Law Review*, 115 (2015), 1371–1452.

Banerjee, S.; M. Humphery-Jenner; V. Nanda; and M. Tham. "Executive Overconfidence and Securities Class Actions." *Journal of Financial and Quantitative Analysis*, 53 (2018), 2685–2719.

Bao, Y.; B. Ke; B. Li; Y. J. Yu; and J. Zhang. "Detecting Accounting Fraud in Publicly Traded U.S. Firms Using a Machine Learning Approach." *Journal of Accounting Research*, 58 (2020), 199–235.

Barzuza, M., and Q. Curtis. "Board Interlocks and Outside Directors' Protection." *Journal of Legal Studies*, 48 (2017), 129–160.

Beneish, M. D., and P. Vorst. "The Cost of Fraud Prediction Errors." *Accounting Review*, 97 (2022), 91–121.

Bentley, J. W.; T. E. Christensen; K. H. Gee; and B. C. Whipple. "Disentangling Managers' and Analysts' Non-GAAP Reporting." *Journal of Accounting Research*, 56 (2018), 1039–1081.

Berger, M. W., and R. S. Gans. "Mass Solicitations Dupe Institutional Investors." Bernstein Litowitz Berger & Grossman LLP, June 19. https://web.archive.org/web/20210225215837/; https://www.blbglaw.com/news/publications/2000-01-01-mass-solicitations-dupe-institutional-investors-by-max-w-berger/_res/id=File1/masssolic.pdf (2003).

Bernard, D.; T. Blackburne; and J. Thomas. "Information Flows Among Rivals and Corporate Investment." *Journal of Financial Economics*, 136 (2020), 760–779.

Bharath, S. T., and A. K. Dittmar. "Why Do Firms Use Private Equity to Opt Out of Public Markets?" *Review of Financial Studies*, 23 (2010), 1771–1818.

Billings, M. B., and M. C. Cedergren. "Strategic Silence, Insider Selling and Litigation Risk." *Journal of Accounting and Economics*, 59 (2015), 119–142.

Bird, A.; S. A. Karolyi; and T. G. Ruchti. "Information Sharing, Holdup, and External Finance: Evidence from Private Firms." *Review of Financial Studies*, 32 (2019), 3075–3104.

Bonaimé, A.; J. Harford; and D. Moore. "Payout Policy Trade-Offs and the Rise of 10b5-1 Preset Repurchase Plans." *Management Science*, 66 (2020), 2762–2786.

Bourveau, T.; Y. Lou; and R. Wang. "Shareholder Litigation and Corporate Disclosure: Evidence from Derivative Lawsuits." *Journal of Accounting Research*, 56 (2018), 797–842.

Boyer, M. M., and L. H. Stern. "D&O Insurance and IPO Performance: What Can We Learn from Insurers?" *Journal of Financial Intermediation*, 23 (2014), 504–540.

Brabec, J.; T. Komárek; V. Franc; and L. Machlica. "On Model Evaluation Under Non-constant Class Imbalance." *Computational Science – International Conference on Computational Science*, 12140 (2020), 74–87.

Brabec, J., and L. Machlica. "Bad Practices in Evaluation Methodology Relevant to Class-Imbalanced Problems." Critiquing and Correcting Trends in Machine Learning Workshop at NeurIPS 2018. https://arxiv.org/abs/1812.01388 (2018).

Brandt, M. W.; A. Brav; J. R. Graham; and A. Kumar. "The Idiosyncratic Volatility Puzzle: Time Trend or Speculative Episodes?" *Review of Financial Studies*, 23 (2010), 863–899.

Brochet, F., and S. Srinivasan. "Accountability of Independent Directors: Evidence from Firms Subject to Securities Litigation." *Journal of Financial Economics*, 111 (2014), 430–449.

BTI Consulting. "37 Law Firms Most Feared in Litigation." October 17. https://bticonsulting.com/themadclientist/37-law-firms-most-feared-in-litigation (2018).

Callen, J. L., and X. Fang. "Religion and Stock Price Crash Risk." *Journal of Financial and Quantitative Analysis*, 50 (2015), 169–195.

Cao, Z., and G. S. Narayanamoorthy. "Accounting and Litigation Risk: Evidence from Directors' and Officers' Insurance Pricing." *Review of Accounting Studies*, 19 (2014), 1–42.

Chalmers, J. M.; L. Y. Dann; and J. Harford. "Managerial Opportunism? Evidence from Directors' and Officers' Insurance Purchases." *Journal of Finance*, 57 (2002), 609–636.

Chambers and Partners. "About Us." https://chambers.com/about-us (2021).

Chen, H.; L. Cohen; U. Gurun; D. Lou; and C. Malloy. "IQ from IP: Simplifying Search in Portfolio Choice." *Journal of Financial Economics*, 138 (2020), 118–137.

Chen, C.; A. G. Huang; and R. Jha. "Idiosyncratic Return Volatility and the Information Quality Underlying Managerial Discretion." *Journal of Financial and Quantitative Analysis*, 47 (2012), 873–899.

Cheng, C. A.; H. H. Huang; Y. Li; and G. Lobo. "Institutional Monitoring through Shareholder Litigation." *Journal of Financial Economics*, 95 (2010), 356–383.

Choi, S. J. "Do the Merits Matter Less after the Private Securities Litigation Reform Act?" *Journal of Law, Economics, and Organization*, 23 (2007), 598–626.

Choi, S. J.; K. K. Nelson; and A. C. Pritchard. "The Screening Effect of the Private Securities Litigation Reform Act." *Journal of Empirical Legal Studies*, 6 (2009), 35–68.

Choi, S. J., and R. B. Thompson. "Securities Litigation and Its Lawyers: Changes during the First Decade after the PSLRA." *Columbia Law Review*, 106 (2006), 1489–1533.

Coe, A. "The 4 Firms That Scare General Counsel The Most." Law360, September 18. https://www.law360.com/articles/1199903/the-4-firms-that-scare-general-counsel-the-most (2019).

Coffee, J. C. "Reforming the Securities Class Action: An Essay on Deterrence and Its Implementation." *Columbia Law Review*, 106 (2006), 1534–1586.

Core, J. E. "On the Corporate Demand for Directors' and Officers' Insurance." *Journal of Risk and Insurance*, 64 (1997), 63–87.

Core, J. E. "The Directors' and Officers' Insurance Premium: An Outside Assessment of the Quality of Corporate Governance." *Journal of Law, Economics, & Organization*, 16 (2000), 449–477.

Cornerstone Research. "Securities Class Action Filings: 2012 Year in Review." https://securities.stanford.edu/research-reports/1996-2012/Cornerstone-Research-Securities-Class-Action-Filings-2012-YIR.pdf (2013).

Cornerstone Research. "Securities Class Action Settlements: 2019 Review and Analysis." https://securities.stanford.edu/research-reports/1996-2019/Securities-Class-Action-Settlements-2019-Review-and-Analysis.pdf (2020).

Cornerstone Research. "Accounting Class Action Filings and Settlements: 2020 Review and Analysis." https://securities.stanford.edu/research-reports/1996-2020/Accounting-Class-Action-Filings-and-Settlements-2020-Review.pdf (2021).

Cox, J. D.; R. S. Thomas; and D. Kiku. "Does the Plaintiff Matter? An Empirical Analysis of Lead Plaintiffs in Securities Class Actions." *Columbia Law Review*, 106 (2006), 1587–1640.

Crane, A. D., and A. Koch. "Shareholder Litigation and Ownership Structure: Evidence from a Natural Experiment." *Management Science*, 64 (2018), 5–23.

Dangl, T., and M. Halling. "Predictive Regressions with Time-Varying Coefficients." *Journal of Financial Economics*, 106 (2012), 157–181.

Dechow, P. M.; R. G. Sloan; and A. P. Sweeney. "Detecting Earnings Management." *Accounting Review*, 70 (1995), 193–225.

Deng, S.; R. H. Willis; and L. Xu. "Shareholder Litigation, Reputational Loss, and Bank Loan Contracting." *Journal of Financial and Quantitative Analysis*, 49 (2014), 1101–1132.

Dimmock, S. G., and W. C. Gerken. "Predicting Fraud by Investment Managers." *Journal of Financial Economics*, 105 (2012), 153–173.

Donelson, D. C.; J. J. Hopkins; and C. G. Yust. "The Role of Directors' and Officers' Insurance in Securities Fraud Class Action Settlements." *Journal of Law and Economics*, 58 (2015), 747–778.

Donelson, D. C.; J. J. Hopkins; and C. G. Yust. "The Cost of Disclosure Regulation: Evidence from D&O Insurance and Nonmeritorious Securities Litigation." *Review of Accounting Studies*, 23 (2018), 528–588.

Donelson, D. C.; A. Kartapanis; J. M. McInnis; and C. G. Yust. "Measuring Accounting Fraud and Irregularities Using Public and Private Enforcement." *Accounting Review*, 96 (2021a), 183–213.

Donelson, D. C.; A. Kartapanis; and C. G. Yust. "Does Media Coverage Cause Meritorious Shareholder Litigation? Evidence from the Stock Option Backdating Scandal." *Journal of Law and Economics*, 64 (2021b), 567–601.

Donelson, D. C.; J. M. McInnis; R. D. Mergenthaler; and Y. Yu. "The Timeliness of Bad Earnings News and Litigation Risk." *Accounting Review*, 87 (2012), 1967–1991.

Drake, M. S.; B. A. Johnson; D. T. Roulstone; and J. R. Thornock. "Is There Information Content in Information Acquisition?" *Accounting Review*, 95 (2020), 113–139.

DuCharme, L. L.; P. H. Malatesta; and S. E. Sefcik. "Earnings Management, Stock Issues, and Shareholder Lawsuits." *Journal of Financial Economics*, 71 (2004), 27–49.

Dyck, A.; A. Morse; and L. Zingales. "Who Blows the Whistle on Corporate Fraud?" *Journal of Finance*, 65 (2010), 2213–2253.

Erichson, H. M. "CAFA's Impact on Class Action Lawyers." *University of Pennsylvania Law Review*, 156 (2007), 1593–1627.

Fernandéz, A.; S. Garcıa; M. Galar; R. C. Prati; B. Krawczyk; and F. Herrera. *Learning from Imbalanced Data Sets*. Cham: Springer (2018).

Fich, E. M., and A. Shivdasani. "Financial Fraud, Director Reptuation, and Shareholder Wealth." *Journal of Financial Economics*, 86 (2007), 306–336.

Field, L.; M. Lowry; and S. Shu. "Does Disclosure Deter or Trigger Litigation?" *Journal of Accounting and Economics*, 39 (2005), 487–507.

Fisher, D. "Hedge Funds Pump Up Mass Torts With Loans, Advertising." Forbes, October 23. https://www.forbes.com/sites/danielfisher/2015/10/23/hedge-funds-finance-firms-pump-money-into-advertising-driven-litigation/?sh=65731e911fb5 (2015).

Francis, J.; D. Philbrick; and K. Schipper. "Shareholder Litigation and Corporate Disclosures." *Journal of Accounting Research*, 32 (1994), 137–164.

Freund, S.; N. Nguyen; and H. V. Phan. "Shareholder Litigation and Corporate Social Responsibility." *Journal of Financial and Quantitative Analysis*, 58 (2023), 512–542.

Gande, A., and C. M. Lewis. "Shareholder-Initiated Class Action Lawsuits: Shareholder Wealth Effects and Industry Spillovers." *Journal of Financial and Quantitative Analysis*, 44 (2009), 823–850.

Géron, A. *Hands-On Machine Learnings with Scikit-Learn & TensorFlow*. Sebastopol, CA: O'Reilly (2017).

Gibbons, B.; P. Iliev; and J. Kalodimos. "Analyst Information Acquisition via EDGAR." *Management Science*, 67 (2021), 769–793.

Greene, W. "The Behaviour of the Maximum Likelihood Estimator of Limited Dependent Variable Models in the Presence of Fixed Effects." *Econometrics Journal*, 7 (2004), 98–119.

Greene, D. "Who is Winning the Securities Class Action War—Plaintiffs or Defendants?" D&O Disclosure, May 1. https://www.dandodiscourse.com/2017/05/01/who-is-winning-the-securities-class-action-war-plaintiffs-or-defendants/ (2017).

Hanley, K. W., and G. Hoberg. "Litigation Risk, Strategic Disclosure and the Underpricing of Initial Public Offerings." *Journal of Financial Economics*, 103 (2012), 235–254.

Harvard Law Review. "Note: Congress, the Supreme Court, and the Rise of Securities-Fraud Class Actions." *Harvard Law Review*, 132 (2019), 1067–1088.

He, J., and M. A. Plumlee. "Measuring Disclosure Using 8-K Filings." *Review of Accounting Studies*, 25 (2020), 903–962.

Hodge, F., and M. Pronk. "The Impact of Expertise and Investment Familiarity on Investors' Use of Online Financial Report Information." *Journal of Accounting, Auditing & Finance*, 21 (2006), 267–292.

Holzman, E.; N. T. Marshall; and B. Schmidt. "Who's on the Hot Seat for an SEC Investigation?" Working Paper, available at https://papers.ssrn.com/sol3/papers.cfm?abstractid=3223815 (2020).

Hong, C. Y., and F. W. Li. "The Information Content of Sudden Insider Silence." *Journal of Financial and Quantitative Analysis*, 54 (2019), 1499–1538.

Hong, H.; T. Lim; and J. C. Stein. "Bad New Travels Slowly: Size, Analyst Coverage, and the Profitability of Momentum Strategies." *Journal of Finance*, 55 (2000), 265–295.

Hopkins, J. J.; E. L. Maydew; and M. Venkatachalam. "Corporate General Counsel and Financial Reporting Quality." *Management Science*, 61 (2015), 129–145.

Hosmer, D. W., and S. Lemeshow. "Assessing the Fit of the Model." In *Applied Logistic Regression*. Hoboken, NJ: John Wiley & Sons (2000).

House of Representatives. "Securities Litigation Reform. H.R. 1058." Conference Report 104–369 (1995).

Huang, A.; K. W. Hui; and R. Z. Li. "Federal Judge Ideology: A New Measure of Ex Ante Litigation Risk." *Journal of Accounting Research*, 57 (2019), 431–489.

Hudson, E. E., and K. Cummins. "United States: Plaintiff Attorney Advertising: Protected Or Prosecutable?" mondaq, July 26. https://www.mondaq.com/unitedstates/marketing/512980/plaintiff-attorney-advertisingprotected-or-prosecutable (2016).

Hutton, A.; S. Shu; and X. Zheng. "Regulatory Transparency and the Alignment of Private and Public Enforcement: Evidence from the Public Disclosure of SEC Comment Letters." *Journal of Financial Economics*, 145 (2022), 297–321.

Iliev, P.; J. Kalodimos; and M. Lowry. "Investors' Attention to Corporate Governance." *Review of Financial Studies*, 34 (2021), 5581–5628.

Johnson, M. F.; K. K. Nelson; and A. C. Pritchard. "Do the Merits Matter More? The Impact of the Private Securities Litigation Reform Act." *Journal of Law, Economics, and Organization*, 23 (2007), 627–652.

Joyce, T. "INSIGHT: Advertising by Plaintiffs' Firms Driving High Class Action Settlement Rate." Bloomberg Law, April 19. https://news.bloomberglaw.com/us-law-week/insight-advertising-by-plaintiffs-firms-driving-high-class-action-settlement-rate (2019).

Karpoff, J. M.; D. S. Lee; and G. S. Martin. "The Cost to Firms of Cooking the Books." *Journal of Financial and Quantitative Analysis*, 43 (2008), 581–611.
Khang, K., and T.-H. D. King. "Does Dividend Policy Relate to Cross-Sectional Variation in Information Asymmetry? Evidence from Returns to Insider Trades." *Financial Management*, 35 (2006), 71–94.
Kim, I., and D. J. Skinner. "Measuring Securities Litigation Risk." *Journal of Accounting and Economics*, 53 (2012), 290–310.
La Porta, R.; F. Lopez-de-Silanes; and A. Shleifer. "What Works in Securities Laws?" *Journal of Finance*, 61 (2006), 1–32.
Landeo, C. M., and M. Nikitin. "Financially-Constrained Lawyers: An Economic Theory of Legal Disputes." *Games and Economic Behavior*, 109 (2018), 625–647.
Larcker, D. F., and A. A. Zakolyukina. "Detecting Deceptive Discussion in Conference Calls." *Journal of Accounting Research*, 50 (2012), 495–540.
Loughran, T., and B. McDonald. "When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks." *Journal of Finance*, 66 (2011), 35–65.
Lowry, M., and S. Shu. "Litigation Risk and IPO Underpricing." *Journal of Financial Economics*, 65 (2002), 309–335.
Meyer, B. D., and N. Mittag. "Misclassification in Binary Choice Models." *Journal of Econometrics*, 200 (2017), 295–311.
Nagel, S., and A. Purnanandam. "Banks' Risk Dynamics and Distance to Default." *Review of Financial Studies*, 33 (2020), 2421–2467.
Nelson, K. K., and A. C. Pritchard. "Carrot or Stick? The Shift from Voluntary to Mandatory Disclosure of Risk Factors." *Journal of Empirical Legal Studies*, 13 (2016), 266–297.
NERA. "Recent Trends in Securities Class Action Litigation: 2016 Full-Year Review." NERA Economic Consulting, January. https://www.nera.com/content/dam/nera/publications/2017/PUB_2016_Securities_Year-End_Trends_Report_0117.pdf (2017).
Perino, M. A. "Did the Private Securities Litigation Reform Act Work?" *University of Illinois Law Review*, 2003 (2003), 913–978.
Peterson, K. "Accounting Complexity, Misreporting, and the Consequences of Misreporting." *Review of Accounting Studies*, 17 (2012), 72–95.
Pomerantz. "SHAREHOLDER ALERT: Pomerantz Law Firm Investigates Claims On Behalf of Investors of Galapagos NV - GLPG." yahoo! Money, September 1. https://money.yahoo.com/shareholder-alert-pomerantz-law-firm-031500090.html (2020).
Pritchard, A. C., and H. A. Sale. "What Counts as Fraud? An Empirical Study of Motions to Dismiss Under the Private Securities Litigation Reform Act." *Journal of Empirical Legal Studies*, 2 (2005), 125–149.
Pukthuanthong, K.; H. Turtle; T. Walker; and J. Wang. "Litigation Risk and Institutional Monitoring." *Journal of Corporate Finance*, 45 (2017), 342–359.
Roberts, M. R., and T. M. Whited. "Endogeneity in Empirical Corporate Finance." In *Handbook of the Economics of Finance*, Vol. 2, G. M. Constantinides, M. Harris, and R. M. Stulz, eds. Amsterdam: Elsevier (2013), 493–572.
Rogers, J. L., and A. Van Buskirk. "Shareholder Litigation and Changes in Disclosure Behavior." *Journal of Accounting and Economics*, 47 (2009), 136–156.
Rosen Law Firm. "Rosen Law Firm Announces Investigation of Securities Claims Against McDonald's Corporation – MCD." Businesswire, August 10. https://www.businesswire.com/news/home/20200810005580/en/Rosen-Law-Firm-Announces-Investigation-Securities-Claims (2020).
Saito, T., and M. Rehmsmeier. "The Precision-Recall Plot Is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets." *PLoS One*, 10 (2015), 1–21.
Segal, J. "The Asset Class No One Knows They Own." Institutional Investor, November 6. https://www.institutionalinvestor.com/article/b1bq4dlkt57shz/The-Asset-Class-No-One-Knows-They-Own (2018).
Skinner, D. J. "Earnings Disclosures and Stockholder Lawsuits." *Journal of Accounting and Economics*, 23 (1997), 249–282.
Wall Street Journal. "Trial Lawyers Party at the Supreme Court." June 24. https://www.wsj.com/articles/trial-lawyers-party-at-the-supreme-court-1403563169 (2014).
Weiss, E. J. "The Lead Plaintiff Provisions of the PSLRA After a Decade, or "Look What's Happened to My Baby."" *Vanderbilt Law Review*, 61 (2008), 543–577.
Zhang, C. "A Reexamination of the Causes of Time-Varying Stock Return Volatilities." *Journal of Financial and Quantitative Analysis*, 45 (2010), 663–684.
Zingales, L. "Is the U.S. Capital Market Losing its Competitive Edge?" Working Paper, University of Chicago (2007).

https://doi.org/10.1017/S0022109022001508 Published online by Cambridge University Press