Risk Fact or Fiction: The Information Content of Risk Factor Disclosures

Maclean Gaulin
University of Utah
mac.gaulin@utah.edu

January, 2019

Abstract

Inconsistent with concerns of uninformative boilerplate or 'copy and paste' disclosure, I find that managers time their identification of new risk factors and removal of previously identified ones to align with the expected occurrence of future adverse outcomes. By using individual risk factors as the unit of disclosure, I am able to provide novel evidence that managers remove stale disclosures on a timely basis. To shed light on what shapes the disclosure equilibrium, I study the managerial response to demand 'shocks' from public and private enforcement actions. The results show that firms respond to investor demand in a manner consistent with the litigation shield hypothesis, and that this effect persists for multiple years. Consistent with the regulatory cost-benefit function, public enforcement does not result in a net increase in disclosed risk factors, but does evoke more definitive disclosures through more specific language and an increased use of numbers.

JEL Classification: G32, G34, J33
Keywords: Disclosure, Risk Factors, Litigation Risk, Regulatory Enforcement

---

This paper is based on my dissertation of the same name. I am grateful to K. Ramesh for his guidance and feedback, and my other committee members Brian Rountree, Shiva Sivaramakrishnan, Robin Sickles, and Brian Akins. This paper has benefited from comments from Thomas Hemmer, Edwin Hu, Alberto Teguia, Gary Lind, Rafael Copat, the brown bag and workshop participants at Rice University, as well as participants of presentations at University of British Columbia, University of Chicago, University of Utah, University of Rochester, Chapman University, UCLA, Southern Methodist University, and the Division of Economic and Risk Analysis (SEC). All remaining errors are my own.

I. INTRODUCTION

Do managers disclose *risk factors* consistent with the regulatory requirement to warn of *future adverse outcomes*? Or, as practitioners and researchers alike worry,[1] are these disclosures uninformative, with firms including every conceivable risk in an effort to minimize potential litigation? Surprisingly, when studying the informativeness of risk factors, extant academic literature has focused on neither the *risk factor* (as the unit of disclosure) nor specific *future adverse outcomes*. Instead, the focus has been primarily on the relation between aggregate measure of risk disclosure (such as total aggregated word count) and market outcomes (e.g. beta, return volatility) in the cross section.[2] These studies provide important evidence that high-risk firms have longer risk disclosures, but they provide little evidence as to whether this association describes the antecedents and consequences of individual firm's disclosure behavior over time. Motivated in part by the ongoing SEC review of disclosure regulations,[3] this study seeks to address this gap through two channels. First, I test whether managers disclose risk factors in a timely fashion to warn of future adverse outcomes. Second, I examine how the demands for risk factors from various stakeholders shape the managerial disclosure decision.

I develop two novel approaches to address the question of whether managers warn of future adverse outcomes. First, I develop a set of measures that track individual risk factors over time, allowing for direct observation of the managerial decision to add, maintain, and remove distinct risk factors across reporting periods. Guided by regulatory language and its practitioner interpretation, I focus on the individual risk factor as the unit of measure to capture self-contained, unique risks. Second, I test the informativeness of my measures on distinct adverse outcomes rather than general market risk, to identify the relation between the time-series evolution of risk factor disclosures and the outcomes about which they are

[1] e.g. SEC Chair Mary Jo White (2013), Brown and Tucker (2011); Lukomnik (2016); Dyer, Lang, and Stice-Lawrence (2017).

[2] For example Kravet and Muslu (2013); Campbell, Chen, Dhaliwal, Lu, and Steele (2014); Nelson and Pritchard (2016); Bao and Datta (2014); Chiu, Guan, and Kim (2015), or Filzen, McBrayer, and Shannon (2016) among others.

[3] Concept release File No. S7-06-16, p.147: www.sec.gov/rules/concept/2016/33-10064.pdf

intended to warn. Together, these two approaches allow me to test directly whether managers, on average, add new, retain existing, and remove obsolete risk factor disclosures in advance of specific adverse events, including realized negative reporting outcomes (net loss and net operating loss) as well as adverse real outcomes (significant sales decline, general business lawsuits, and securities litigations). Inconsistent with concerns of uninformative boilerplate or 'copy and paste' disclosure, my results suggest that managers time their identification of new risks and the removal of previously identified risks to align with the likelihood of future adverse outcomes. These results support the hypothesis that managers disclose risk factors in accordance with the SEC regulation that requires distinct risk factors addressing specific risks, in essence, forecasting adverse outcomes.

The regulation —and court precedence regarding the use of risk factors as legal protection— also require risk factors be detailed and specific. To test whether managers use improved descriptive precision in their disclosures to warn of future adverse outcomes, I perform two sets of robustness tests. First, in an attempt to capture the 'definitiveness' of disclosures, I measure three proxies from the risk factors: specificity (effectively proper nouns, as used in Hope, Hu, and Lu 2016), numeric intensity (percentage of numbers used, as used in Bozanic, Dietrich, and Johnson 2017), and verbosity (average words per risk factor, to capture the ontology of Bloomfield 2008). I find that managers generally increase the level of definitiveness ahead of adverse outcomes, consistent with managers gaining and disclosing improved information as adverse events become more probable. Second, using a Latent Dirichlet Allocation (LDA) approach, I perform a topic analysis of the risk factors and relate specific topics to corresponding adverse outcomes. I identify two key topics that plausibly relate to macro and micro adverse outcomes: oil price declines and goodwill impairments respectively, and find that on average firms disclose more risk factors specific to these topics in advance of their occurrences. While an LDA labeling approach is inherently subject to researcher judgment, the evidence from this topic analysis lends credence to the conclusion that firms are disclosing 'specific' risk factors commensurate with their outcome probability.

Collectively, my measures are consistent with the principal outlined in Bloomfield, Nelson, and Soltes (2016), which suggests that researchers construct "variables [...] in a way that is most likely to capture the constructs specified in the theory they wish to test." Given that my paper studies the information content in mandatory disclosures, I attempt to construct measures that faithfully represents the regulatory intent of the disclosure requirements. The regulation requires firms to include a *discussion of the most significant factors that make the offering speculative or risky* and that the factors be 'organized logically,' with each risk factor summarized by a caption. Therefore, I focus on the individual risk factor itself as the unit of disclosure. By narrowing the unit of measure to individual risk factors, I observe specific risks being addressed through time, and capture managers' decision to add, remove, or continue to disclose about individual risks, thereby signaling whether they consider a given risk factor to be *significant*. By testing whether disclosing these individual risks predicts adverse outcomes, I find that the manager's decision to signal that a specific risk is *significant* is timely and informative.

My findings are important in light of the fact that risk factor disclosures have become an increasingly large portion of annual reports, growing from 11.3% to 16.1% of the length (by word count) over the past decade. This growth has raised questions among practitioners and academics alike as to whether the increase in disclosure is actually informative or merely boilerplate.[4] Previous literature addressing this concern has typically focused on the length of risk disclosure using a word count or "bag of words" approach. However, to address the topical issue of whether managers disclose risk factors per the regulation, these previously used measures may not fully capture the information managers attempt to convey.

If managers adhere to the regulation and convey the *most significant factors that make the offering speculative or risky*, they may be focusing on deciding *which* risks they believe to be significant, rather than considering the risk factor section in aggregate. Therefore, to the extent that the breadth of risks factors identified by managers improves information

[4]For example SEC Chair Mary Jo White (2013), Brown and Tucker (2011); Lukomnik (2016); Dyer, Lang, and Stice-Lawrence (2017).

'fineness' (e.g. Blackwell 1951; D'Souza, Ramesh, and Shen 2009; Cheng, Huang, and Li 2016), the length of disclosure could merely be a byproduct of increased informativeness. While finding that generally riskier firms have longer disclosures is an important first step, my study adds to the literature by studying a potential underlying mechanism through which these disclosures convey information. Notably, to the best of my knowledge, my study is the first to shed light on systematic and timely risk factor disclosure *reductions*: firms remove 2.5 risk factors on average every year, and 25% of the firm-years exhibit a net *reduction* in the total number of risk factors (see figure I). Further, by demonstrating that firms update their disclosures in an informative fashion over time, I hope to provide valuable evidence to the current regulatory initiative seeking to modernize the disclosure framework.

Given this evidence that managers provide risk factor disclosures that are informative about future adverse outcomes, I then study the demand for these disclosures from two salient sources: investors and regulators. This an important aspect of the disclosure equilibrium, one which is not strictly dictated by the normal disclosure cost considerations. I proxy for direct demand from investors via private enforcement through securities litigation, and demand from regulators via public enforcement through the SEC comment letter process. Consistent with predictions that disclosure of risk factors can act as a litigation shield (Skinner 1994; 1997; Robbins and Rothenberg 2005), I find that firms respond to private enforcement by increasing the number of new risk factors they identify, though not the definitiveness thereof. Also consistent with the large cost of securities litigation (Rose 2008), I find that this penchant for increased disclosure persists for multiple years. In response to public enforcement actions, however, I find that firms do not increase the total number of risk factors they identify, but do significantly increase the definitiveness of their risk factors.

Collectively, the demand-side results provide important insights regarding the economic determinants of the informativeness of risk factor disclosures. I find evidence suggesting that private investors' demand for risk factors is in line with the regulation requiring the disclosure of a plurality of distinct risk factors, but suggests that investors are informed by *which* risks managers deem 'most significant,' rather than through descriptions of those risks. In contrast, the demand for risk factor disclosures from regulators seems to be consistent with their refraining from dictating *which* risks managers identify, but requiring the already identified factors to be definitive. Thus, while private enforcement influences managerial choice of *which* risk factors to disclose, public enforcement shapes the *quality* of the disclosures for risk factors that managers choose to report. These results shed greater light on my earlier findings on the predictive value of risk factor disclosures and provide novel evidence on the complementary role of public versus private enforcement in sculpting the disclosure landscape. Using regulation-driven measures, my study provides compelling evidence not only on the consequences of risk factor disclosures, but also on the antecedent, demand-side incentives that help determine disclosure properties.

The prior literature examining the informativeness of risk disclosures largely focuses on the aggregate level, and draws mixed conclusions (see Section II). My study complements this literature by studying individual risk factor disclosures at the firm level, and suggests a potential reason for the disagreement may be that the information in risk factors is not reflected primarily through aggregate measures, but rather through the time-series evolution of individual risk factors. Additionally, by finding that the regulators and investors demand different dimensions of information from the disclosures, my results may suggest a reason for previous findings that regulatory impact weakens the association between disclosure length and underlying risk.

My study contributes to the literature in two ways. First, my paper contributes to the growing literature on the informativeness of risk factor disclosures. While previous work by Nelson and Pritchard (2016) and Beatty, Cheng, and Zhang (2018) conclude that regulatory intervention from the SEC leads to less informative disclosures, my findings suggest an alternative perspective. Specifically, I find that the SEC elicits more detailed disclosures, both in specificity and in length of descriptions. While one outcome of this regulatory focus may be that the length of disclosure is less correlated with underlying economic activity, my findings suggest that regulatory intervention may lead to more definitive or quantitative disclosures. Additionally, my results point to an interesting role that risk factor disclosures may play in the information milieu. Rogers and Van Buskirk (2009) find that firms provide less voluntary disclosure in response to securities litigation, and conclude that managers view *voluntary* disclosure as a legal liability.[5] My finding that risk factor disclosures are *expanded* after litigation suggests that these *mandated* disclosures may serve as a mitigating factor for litigation risk, potentially substituting for voluntary information when the cost of the latter increases.

Second, this study contributes to the developing literature on information overload and boilerplate disclosure. I demonstrate that managers update their disclosures in a timely fashion, suggesting that they are not merely 'copy and pasting.' Instead, my results suggest that firms are informatively choosing when to add, and potentially more importantly, when to remove information. In contrast to previous studies of boilerplate disclosure, which take the position that increasing similarity over time connotes uninformative disclosure (e.g. Brown and Tucker 2011; Nelson and Pritchard 2016; Dyer, Lang, and Stice-Lawrence 2017), my finding that the removal of disclosures is informative suggests that persistent disclosures may not be inherently uninformative. I also show that managers supply disclosure to meet information demand differentially based on the source of the demand. This is in line with the findings of Bird and Karolyi (2016) and Boone and White (2015) regarding the provision of voluntary information to meet institutional investor demand. Together my results suggest caution in interpreting the increasing length of disclosure as a sufficient signal of information overload. This study demonstrates that in some cases a more nuanced measure of information may be warranted, potentially one based on the underlying data generating process, as informed by regulation or managerial intent.

[5]While Naughton, Rusticus, Wang, and Yeung (2015) suggest this result may not be driven by lower litigation costs, my findings of increased *mandatory* risk factor disclosures are still in contrast to the reduction of *voluntary* disclosure.

II. HYPOTHESES

Risk Factor Supply

Previous literature has focused primarily on the decision usefulness of risk factor disclosures by studying their capital market effects, generally in the cross section. However the evidence in this extant research on the informativeness of risk factor disclosures is mixed.[6] Studies on voluntary risk disclosures (before the SEC mandated their inclusion in annual reports) have found mixed results as to which disclosures signal risk. While Kravet and Muslu (2013) find that industry-level (but not firm-idiosyncratic) risk disclosures are correlated with stock- and analyst-based risk measures, Nelson and Pritchard (2016) finds the association between risk disclosure and market risk measures only exists for firms facing high ex-ante litigation risk. Despite disagreeing on whether risk disclosures are informative generally or just conditionally so in some cross-sections, both studies do highlight that the length of these disclosures has increased significantly over time.

Motivated in part by this increase in voluntary disclosure, in 2005 the SEC mandated that firms include a risk factor section in their annual and quarterly reports.[7] Studies regarding the mandatory disclosure regime are also mixed on the efficacy of the regulation and the informativeness of the disclosures themselves. Nelson and Pritchard (2016) find that the association between the length of risk disclosure and both ex-ante and ex-post market risk vanishes after the regulation, and conclude that the mandate did not fully substitute for the value of previous voluntary risk disclosures. Campbell, Chen, Dhaliwal, Lu, and Steele (2014), on the other hand, find significant associations between risk factor disclosures and both ex-ante and ex-post measures of market risk (as well as ex-post information asymmetry) on average both for all risk disclosures, and for specific word-categories.

One reason for the mixed evidence in prior research may stem from how each study measures the risk disclosures themselves. For example, past studies have measured risk

[6]'Informativeness' in these studies is typically defined as 'being decision useful,' and is often demonstrated via aggregated risk disclosures' correlations with returns or other market risk measures.

[7]Securities Offering Reform, SEC File No. S7-38-04 p. 259: sec.gov/rules/final/33-8591.pdf

disclosure using: whether any disclosure is present (Filzen 2015), change in the number of risk-related sentences (Kravet and Muslu 2013), total or change in word count (Campbell, Chen, Dhaliwal, Lu, and Steele 2014; Brown, Tian, and Tucker 2018), key-word counts for multiple risk categories (Campbell, Chen, Dhaliwal, Lu, and Steele 2014), topic-based measures from a trained model (Huang and Li 2008; Bao and Datta 2014), and 'similarity' scores between disclosures (Nelson and Pritchard 2016; Brown, Tian, and Tucker 2018). The varied use of proxies in previous literature has potentially contributed to the lack of agreement of what information risk factors convey to which stakeholders. As seen above, many of these papers reach different decisions when choosing a proxy to match the construct under study.

This issue of matching proxy to construct is what Bloomfield, Nelson, and Soltes (2016) refer to as *distillation*. A key input to this distillation process is the construct the researcher is attempting to capture. Previous research on risk factors has primarily focused on aggregate *risk disclosure*, and selected proxies accordingly as listed above. This study instead focuses on risk factors as defined by the SEC regulation for two reasons: doing so provides a concrete foundation from which to derive measures, and because the regulation, enforced by the SEC, informs (and is consistent with) practitioner guidance and court precedence.

The requirement to disclose risk factors is set forth in Item 503(c) of Regulation S-K:

> **§229.503 (Item 503) (c) Risk factors**. Where appropriate, provide under the caption "Risk Factors" a discussion of the most significant factors that make the offering speculative or risky. This discussion must be concise and organized logically. Do not present risks that could apply to any issuer or any offering. Explain how the risk affects the issuer or the securities being offered. Set forth each risk factor under a subcaption that adequately describes the risk.

Given the regulation requires that managers disclose 'the most significant factors,' that do 'not [...] apply to any issuer' with each risk factor summarized by a caption, this suggests that risk factors be treated as a) distinct, and b) a signal of managers' belief about specific adverse outcomes. The SEC review process adheres firms to the verifiable portion of these requirements, or requires justification otherwise (Bozanic, Dietrich, and Johnson 2017). Thus, for formulating proxies to measure risk factor disclosure, the regulation arguably provides a defensible basis.

In addition to regulatory enforcement, investors could influence managers to disclose risk factors consistent with the regulation, because by doing so, managers can potentially protect against securities litigation. This 'litigation shield' hypothesis (Skinner 1994, 1997) posits that warning investors of negative outcomes in advance can reduce the expected cost of securities litigation via a reduction in the stock impact of the 'corrective disclosure' and shortening the class period.[8] Additionally, by disclosing information in a timely manner, firms can potentially refute a plaintiff's claim that the firm did not adequately provide investors with information (Robbins and Rothenberg 2005). However, to invoke the litigation shield, firms must disclose specific and detailed risk factors (consistent with the regulation). This is because the value of risk factors as legal protection stems from the 'bespeaks caution' doctrine and the subsequent Private Securities Litigation Reform Act (PSLRA), which provide a 'safe harbor' against legal liability when firms disclose forward-looking information, as long as they include specific cautionary language regarding the uncertainty of the forecasts.

Together, the demand from investors and regulatory enforcement suggests that managers have strong incentives to disclose risk factors consistent with the regulation. However, as these disclosures potentially mitigate both litigation and regulatory actions, firms may be incentivized to 'overload' stakeholders and disclose of every conceivable risk they face. This information overload concern is voiced by both legal opinions and regulators. Justice Marshall raised this concern in 1976, cautioning that "management's fear of exposing itself to substantial liability may cause it simply to bury the shareholders in an avalanche of trivial information – a result that is hardly conducive to informed decision-making"[9] SEC Chair Mary Jo White (2013) echoed this concern, questioning "whether investors need and are optimally served by the detailed and lengthy disclosures about all of the topics that

[8] To bring a securities litigation, plaintiffs must define both an initial misleading disclosure and a subsequent corrective disclosure (and demonstrate scienter, among other factors).

[9] TSC Industries, Inc. v. Northway, 426 U.S. 438, 448-449 (1976))

companies currently provide." As these suggest, it is unclear ex-ante whether disclosures will convey an informative signal, especially given the discretion allowed by the subjective materiality threshold of risk factor disclosures and the absence of guidance on risk factor detail or specificity.

To test whether managers do disclose risk factors consistent with the regulation and in a timely fashion, I take two novel approaches to capture how managers potentially convey information. The first is my focus on the individual risk factor as the unit of measure, which allows me to capture the evolution of each risk factor over time. To do so I label each risk factor a firm discloses in their annual report as belonging to one of three groups: new risk factors that were not present in the previous year, old risk factors that were present in the previous year and persist unchanged in the current year, and removed risk factors, that were present in the previous year but no longer included in the current year. This allows me to observe how many new risk factors firms are identifying, as well as how many they are removing, which is potentially indicative of those risks managers deem no longer apply to their firm. This ability to capture managers disclosing information through the *removal* of risk factors is unique to my approach.

After selecting a measure based on the regulation and practitioner guidance of *how* firms are conveying information, my second novel approach uses similar reasoning to define *what* firms are disclosing. As discussed above, previous literature has primarily studied the impact of risk disclosures on general stock market measures, to test whether risk factors provide information on firm risk in general.[10] These studies importantly demonstrate that managers provide novel information to markets, but do not directly address the channel through which this information manifests. As discussed above, the regulation suggests one such channel: signals of managers' belief about *specific adverse outcomes*. Therefore, instead of focusing on

[10]See, for example, Campbell, Chen, Dhaliwal, Lu, and Steele (2014), Nelson and Pritchard (2016), Huang and Li (2008), Hope, Hu, and Lu (2016), Kravet and Muslu (2013), Chiu, Guan, and Kim (2015). One recent contrary example is Campbell, Cecchini, Cianci, Ehinger, and Werner (2018), who study the relation between tax-related risk disclosures and taxes paid. However, they still study aggregate disclosure (through word counts) and largely focus on the relationship between firms when studying cash tax effects.

inferences from general market risk, I directly study whether risk factor disclosures warn of adverse outcomes themselves.

My first set of hypotheses posit that managers provide information in risk factor disclosures to signal future adverse outcomes. I predict that they do so in two ways: updating *which* risks they identify, and updating the *definitiveness* with which they disclose their risks. As an example of the former, if managers identify a new risk they deem among their 'most significant,' and intend to fulfill both regulatory requirement and investor demand, then they will add a risk factor outlining this risk. Additionally, as the probability of an adverse outcome increases, managers have incentives both to gather more information about the event (Heinle and Smith 2017), and to disclose their information (in order to mitigate disclosure costs associated with non-disclosure, e.g. litigation risk discussed above). If managers disclose consistent with their subjective probabilities of the adverse outcomes, this will translate into managers forecasting, or forewarning against, these adverse outcomes.[11] Conversely, when the probability of an adverse outcome decreases, if managers are continually updating their risk factors to reflect only those 'most significant factors,' then they will remove factors when adverse outcomes are less likely.

> **H1:** Firms disclose more and remove fewer risk factors in advance of adverse economic outcomes.

My second prediction for how managers provide information about future adverse outcomes is through the definitiveness, or level of detail provided about these events. Hope, Hu, and Lu (2016) find that more specific risk factor disclosures are associated with stronger market and analyst responses, suggesting that writing more definitive risk factor disclosures conveys more information to investors. More definitive risk factors being more informative to investors is also consistent with court precedence requiring risk factors to be detailed in order to benefit

[11] This presumes that managers' information sets reflect the objective probabilities of the associated outcomes, i.e. that managers can accurately estimate the risks their firms face. One confounding possibility is untruthful signaling, resulting in non-disclosure (it is unlikely that managers would falsely disclose adverse outcomes they do not face). The presence of such non-disclosures should bias against my finding any forecasting results, however.

from the legal protection of the 'bespeaks caution doctrine.'[12] As discussed above, managers may obtain more precise information about adverse outcomes as they become more probable. If managers disclose risk factors per the regulation and consistent with the legal requirements, together these forces will result in an increase in definitiveness of the risk factor disclosures in advance of adverse outcomes.

> **H2:** Firms provide more definitive risk factors in advance of adverse economic outcomes.

Risk Factor Demand

The motivation for basing my measures and tests on the regulation is that both investors and regulators provide incentive for managers to adhere to the regulation. To capture the validity of this assumption, I test whether these two forces (public regulators and private investors) actually demand timely and/or definitive risk factors. Observing demand directly is difficult, therefore I employ a methodology similar to that used in Rogers and Van Buskirk (2009) in which I measure manager's disclosure response to a demand 'shock.' By observing how managers respond, I am able to draw conclusions about managers' beliefs as to what disclosures their investors or regulators demand.

I first focus on the demand for disclosures from investors, to whom many benefits of disclosure accrue. One potentially significant benefit may be to avoid securities litigation, which can be significantly costly to a firm (Rose 2008). To avoid this, managers are likely to attempt to reduce their litigation risk by disclosing information they believe investors demand (Skinner 1994; White 2013). Anecdotally, firms do use risk factor disclosures to gain legal protection,[13] when the disclosures adhere to quality standards set forth by court precedent:

[12] e.g. In re Worlds of Wonder Securities Litigation, 35 F.3d 1407 (9th Cir. 1994), Inst. Investors Group v. Avaya, Inc., 564 F.3d 242, 256 (3d Cir. 2009)

[13] See e.g. In re Convergent Technologies Security Litigation, 948 F.2d 507,515(9th Cir. 1991), In re Worlds of Wonder Securities Litigation, 35 F.3d 1407 (9th Cir. 1994).

> Cautionary language must be extensive and specific. A vague or blanket (boilerplate) disclaimer which merely warns the reader that the investment has risks will ordinarily be inadequate to prevent misinformation. To suffice, the cautionary statements must be substantive and tailored to the specific future projections, estimates or opinions in the prospectus which the plaintiffs challenge. (Inst. Investors Group v. Avaya, Inc. 564 F.3d 242, 256; 3d Cir. 2009)

In Slayton vs American Express (604 F.3d 758, 762; 2d Cir. 2010) the court concluded that the above standards were not met, stating: "the defendants' [risk factor disclosure] verges on the mere boilerplate [...] Our conclusion is bolstered by the fact that the defendants' cautionary language remained the same even while the problem changed." This legal precedent is consistent with the regulatory requirement that risk factor disclosures be detailed and address specific outcomes. The latter court opinion also suggests that to gain legal protection, firms must update their risk factor disclosures in a timely fashion to reflect new information.

If firms believe risk factors provide a litigation shield, then a private enforcement event will signal deficient disclosures, and managers will increase their disclosure as a result. On the other hand, if managers believe disclosure causes litigation risk (Francis, Philbrick, and Schipper 1994; Rogers and Van Buskirk 2009), then managers will scale back their disclosures after a private enforcement event. However, given the legal practitioner guidance on the litigation value of risk factor disclosure (e.g. Robbins and Rothenberg 2005), I predict that unlike voluntary disclosures, firms will increase their risk factor disclosures after a private enforcement event. As discussed above, court precedence requires risks factor disclosures to be definitive and specific to the risk they address. Therefore, I predict that in response to a private enforcement event, firms will increase both the number of risk factors they identify and the definitiveness of their disclosure.

> **H3:** Firms respond to private enforcement by identifying more risk factors and increasing their definitiveness.

Regulatory oversight, which Cox, Thomas, and Kiku (2003) describe as the "first line of defense against ongoing disclosure violations," also potentially plays a role in disclosure choice through the comment letter process (see, e.g. Johnston and Petacchi 2017; Bozanic, Dietrich, and Johnson 2017; Brown, Tian, and Tucker 2018). The incentives of public enforcement differ from those of investors, however (Cox, Thomas, and Kiku 2003). The demand for information from the SEC is reflected in their stated goals:[14]

> All investors, whether large institutions or private individuals, should have access to certain basic facts about an investment prior to buying it, and so long as they hold it. To achieve this, the SEC requires public companies to disclose meaningful financial and other information to the public. [...] Only through the steady flow of timely, comprehensive, and accurate information can people make sound investment decisions.

The SEC is also resource and capacity constrained (Cox, Thomas, and Kiku 2003; Jackson and Roe 2009). The information available to the review process are largely the public disclosures provided by the firm.[15] In the absence of manager's private information, it is unclear that an SEC reviewer has the capacity to identify novel omitted, yet significant, risk factors on any systematic basis. This is consistent with the claims made by the SEC: "The Division's review process is not a guarantee that the disclosure is complete and accurate."[16] Therefore, I do not expect public enforcement to result in firms expanding the set of risk factors they identify. However, the SEC does have the ability to measure adherence to Item 503(c) in the existing disclosures they receive, specifically: *Do not present risks that could apply to any issuer or any offering. Explain how the risk affects the issuer or the securities being offered.* Because the SEC review process vis-à-vis risk factors is guided by this regulation, I predict the reviews will identify and request improvements in vague or generic disclosures. If this is the case, the SEC comment letters should result in improvements in definitiveness of subsequent disclosures.

> **H4:** Firms respond to public enforcement by increasing the definitiveness of risks they identify.

[14]From www.sec.gov/about/whatwedo.shtml

[15] See www.sec.gov/divisions/corpfin/cffilingreview.htm

[16]*Id., referring to the Division of Corporation Finance.*

III. DATA AND METHODOLOGY

Risk Factor disclosures have been required in annual and quarterly disclosures under Item 1A since the SEC regulation took effect in 2005. My sample starts with these earliest filings, and extends through fiscal year 2015. I start gathering the sample from the Compustat database between fiscal year end 2005 and 2015 (112,402 firm years), removing those firms with missing historical Central Index Key (CIK) identifiers (22,715 missing). I further require firm-years have non-missing total assets (10,654 missing) and a valid match to the CRSP securities database (27,784 missing), resulting in a sample of 51,249 firm years. With the 9,632 unique CIKs from this sample of firms, I extract the list of all matching 10-K and 10-KSB filings from the EDGAR index available from the SEC. This results in a sample of 67,648 annual reports filed for fiscal year 2005 or later, from which I then attempt to extract the Item 1A Risk Factor section.[17]

I extract the risk factor section following the methodology described in Campbell, Chen, Dhaliwal, Lu, and Steele (2014). Appendix B describes this process in detail, resulting in a final sample of 31,549 firm years with non-missing total assets, previous year returns from CRSP, and non-missing risk factors.[18] The individual risk factors are separated using contextual clues provided in the HTML document, such as bold, italic, underline, and paragraph demarcations. I require that each risk factor be defined by an accentuated heading (bold, underlined, or italic) which is isolated on its own line or located at the beginning of a paragraph. The practice of captioning risk factors is pervasive because it is required in the wording of the regulation, and is enforced by SEC reviewers through comment letters. I use these extracted risk factor heading to define a unique risk factor.

The methodology proposed in this paper tracks risk factor evolution over time, using three proxies to do so: the total number of risk factors, the number of new risk factors, and the

[17] The number of EDGAR filings I search is larger than the number of CRSP-Compustat merged firm years because I do not filter out firm-years with missing data. I do this to avoid unnecessarily dropping observations in the year-over-year comparison of risk factor statements, for example when a firm has a valid annual report in the previous year but no listed CRSP identifier.

[18] The sample becomes 26,547 firm years when previous year non-missing risk factor disclosure is required for the regressions.

number of dropped risk factors which were included in the previous filing but do not appear in the current filing. To determine whether a risk factor was included in a previous year, I compare the risk factor heading from the text with all the previous year's set of risk factors without replacement. To compare these headings, I use the Ratcliff and Metzener Gestalt Pattern Match algorithm (RMGPM).[19] The RMGPM is a flexible string match algorithm that allows for a parameterized amount of flexibility for determining if two strings are the same.

To match risk factors across years, I start with the set of risk factors comprising N text strings in year t, sorted from longest to shortest number of characters. I then iterate through the N strings, comparing each to the set of M risk factors from year t-1, also sorted by decreasing length. For each of the N risk factors, I search for an exact match, and failing that, I iteratively reduce the parameter restricting the 'exactness' of the match. If no match is found after allowing for a minimum of 50% character level match, the risk factor is counted as 'new.' When a match is found, the matching string from the set of t-1 risk factors is also removed. After searching for all N risk factors from year t, the remaining risk factors from year t-1 are counted as 'dropped.'

I measure the definitiveness of risk factors using three proxies: specificity, numeric intensity, and average words per risk factor. I calculate the specificity of risk factor disclosures using the Stanford Named Entity Recognition (NER) algorithm, similar to that employed in Hope, Hu, and Lu (2016). The NER algorithm extracts specific entities (proper nouns), and is intended to capture whether disclosure uses general language or names a specific entity or location (e.g. our competitor vs. Microsoft, our supplier vs. Foxconn). I use three of the seven entity categories from the pre-trained classifier provided with the algorithm: location, person, and organization.[20] Unlike Hope, Hu, and Lu (2016), I omit the remaining four categories

[19] Included in the *difflib* package in the Python standard library. It is the algorithm that is used in the Python language to compare source code for differences.

[20] Using the pre-trained, seven-class classifier (Location, Person, Organization, Money, Percent, Date, Time from *english.muc.7class.distsim.crf.ser.gz*) downloaded from nlp.stanford.edu/software/CRF-NER.shtml#Models

to avoid overlap with the other proxy of definitiveness, numeric intensity. Consistent with Brown and Tucker (2011), I measure numeric intensity as the percentage of words that are numeric. The numeric intensity is intended to capture the level of detail provided in disclosures, through instances of numbers, currencies, percentages, or dates. These two measures are included as changes to capture the aggregate difference in definitiveness of risk factors, including new risk factors but also changes in the definitiveness of the existing risk factor disclosures. To capture the brevity of the descriptions used in the risk factors, I include the change in average words per risk factor. This proxy is consistent with what Bloomfield (2008) describes as 'ontology,' whereby negative outcomes are difficult to describe and require longer, more detailed explanations. Additionally, as adverse events become more likely, managers potentially gather more information (Heinle and Smith 2017) which could translate into more information available to disclose.

To test whether managers warn of adverse outcomes directly, I focus on four outcomes: negative net income, negative operating income, sales decline (greater than 10% of previous year's sales),[21] and business, or non-securities, litigation. These outcomes are chosen to reflect significant adverse events which are ubiquitously negative.[22] The business lawsuit data are from the CapitalIQ Key Developments database. I define a business lawsuit as an event in a fiscal year with event code 25, corresponding to Litigation events. I form the variable *Lawsuit Intensity* as the log of the number of lawsuits that occur in a given fiscal year (plus one). I gather SEC comment letters from the Audit Analytics Comment Letter database, and classify the comment letter as relating to risk factors if the risk factor column in the database is non-empty (consistent with Brown, Tian, and Tucker 2018). I gather securities litigation events from the Stanford Law School's Securities Class Action Clearinghouse.[23] The data gathered include class start and end dates, as well as filing date and outcome (settled,

[21]The results are generally robust to using a 5% threshold, with some weakening in the significance of the definitiveness results.

[22]Because these outcomes may not be *conditionally* negative signals, in the robustness section I show that risk factors are also generally predictive of continuous future economic outcomes.

[23]Scraped from securities.stanford.edu/filings.html using code from github.com/gaulinmp/lit_scrape.

dismissed, or ongoing). A firm is said to have a securities litigation event if a filing date occurs within the fiscal year. To differentiate them in the tables, I refer to securities litigation as *Litigation* or *Securities Litigation*, and business lawsuits as *Lawsuit*.

[Table I about here.]

Table I presents the descriptive statistics for the final sample. The average firm in my sample has 29.5 risk factors, and adds 3.7 new risks and removes 2.5 obsolete risks per year. Firms increase their number of risk factors by 1.2 per year on average, although approximately one quarter of firm-years actually see a net decrease in the total number of risk factors, as shown in Figure I below.

[Figure I about here.]

When cutting the sample by firm age, younger firms (below the median firm age of 13 years) disclose 34.4 risk factors on average, and identify 4.4 new factors per year and remove 3.2 factors. Older firms (above median age) disclose significantly fewer risk factors, only 25.8, and identify new factors and remove old factors at significantly lower rates as well (3.4 and 2.2, respectively). This potentially suggests life cycle effects in risk factor disclosures, with older, more established, and less volatile firms disclosing fewer risks and experiencing less risk 'turnover.' However, when looking at the change in word count of the risk factors between these two groups, there is no significant difference.[24] This suggests that the time series evolution of individual risk factors may capture underlying economic differences that a word count approach might not.

IV. RESULTS

Supply of Risk Factors

To address the first two hypotheses, I employ a predictive framework in which I regress the adverse outcomes (in the following year) on the measures of risk factor disclosures and test

[24] The change in word counts are 135.5 and 136.4 for young and old firms respectively (t-stat = -0.16). In contrast, the t-stat for differences in new, dropped, and total individual risk factors are all significant at <0.001 level.

whether increases or decreases in risk factors and their definitiveness precede the events. To capture the time-series evolution of both the underlying economics and firm disclosure, I employ panel regressions and control for firm specific heterogeneity using firm fixed effects. My measure of adverse outcomes is an indicator variable for negative income, negative operating income, and significant sales decline of greater than ten percent of the previous year's sales.[25] To capture general business lawsuits, I use the natural log (assuming that there are diminishing marginal benefits to warning of subsequent lawsuits) of one plus the number of lawsuits brought against the firm in a given year.[26] To model the adverse outcomes that are coded as dichotomous variables, I use a probit specification with correlated random effects to control for firm specific heterogeneity (Wooldridge 2010), and otherwise use OLS with firm fixed effects.

[Table II about here.]

Because the hypotheses predict that managers warn of adverse outcomes, initially I do not include controls in the model beyond year and firm level fixed effects. Effectively, this tests whether managers warn of adverse outcomes independent of expectations or other potential disclosure/information channels. This is important to verify for any conclusion about average predictability, but also because in the absence of unconditional predictability, any further conditional results would be suspect. Consistent with my first hypothesis, Table II demonstrates that managers do increase the number of risk factors they identify in advance of adverse events. Managers are also less likely to remove old risk factors in advance of negative income and operating income, and marginally so for sales declines. Put differently, this is consistent with managers removing more risk factors when these adverse outcomes are less likely (due to the firm fixed effect nature of the regressions). The lagged total number of risk factors is also significant, suggesting that not only are the contemporary

[25] To avoid biasing towards small firms, I require the sales decline to be at least ten million dollars. Thus, firms with sales less than 100 million must experience a decrease in sales of more than ten million.

[26] In untabulated results, I also employ a Poisson model of the number of lawsuits (including firm fixed effects), and find my results are robust to this alternative specification.

changes in risk factors predictive of adverse outcomes, but lagged increases in risk factors are informative as well. Taken together, this evidence supports my first hypothesis that managers are informatively adding and removing risk factors in a timely fashion.[27]

Consistent with my second hypothesis, the results in Table II suggest that managers disclose more definitive risk factors in advance of adverse outcomes. For the three accounting measure outcomes, managers increase both the frequency of numbers used in the disclosures and the verbosity with which they describe them. Because this increase in length of risk factors is concomitant with increases in the other measures of definitiveness (specificity or numeric intensity), it is more likely that the verbosity of risk factors provided are in line with Bloomfield's (2008) ontology explanation rather than obfuscation. Interestingly managers do not use more numbers to describe the external risk of lawsuits, but they do use more specificity. This could suggest that managers disclose the more quantitative information stemming from their internal accounting information, and provide more qualitative detail about potentially less quantifiable external risks.

One limitation of testing the unconditional predictive ability of risk factors is that it does not speak to whether the information signal from observing risk factor decisions is subsumed by other financial accounting information. While it is important to test whether managers are reflecting their expectations of future events in the disclosure decisions, it is also of interest to study whether those expectations are fully captured elsewhere. To address this, I further include controls for ex-ante risk as identified in previous studies of risk and disclosure (e.g., Campbell, Chen, Dhaliwal, Lu, and Steele 2014; Nelson and Pritchard 2016). I also include a control for the lagged dependent variable using the contemporary continuous (not dichotomous) value measured at the time of the risk factor disclosure release, specifically: net income, operating income, and sales (scaled by total assets) for specifications (1)–(3), respectively, and the log number of suits for specification (4).

[Table III about here.]

[27]In untabulated results, I include the net change in risk factor disclosures, rather than the addition and removal separately, and find that the net change in disclosed risk factors also forecast these adverse events.

Consistent with the unconditional results, managers disclose the new risk factors in advance of adverse outcomes, incrementally to what is explained by observable ex-ante risk.[28] This suggests that the information compounded into the risk factor disclosure decision is not fully captured by other financial accounting disclosures, lending credence to the claim of risk factor disclosures being 'informative.'

Demand for Risk Factors

The third and fourth hypotheses seek to identify what dimensions of risk disclosures, if any, seem to be demanded by shareholders and regulators. To test these hypotheses, I measure shareholder and regulator demand using indicator variables representing the presence of a securities litigation and comment letter, respectively. I separately regress each of the six risk factor disclosure measures (the time series evolution and definitiveness variables) on the indicators, which are equal to one if the event occurred over the course of the fiscal year. I also include two lags of the indicators to capture the longevity of manager's responses to these events. In addition to the same controls used in the previous tables, including firm fixed effects, I also include lagged values of the dependent variable to control for serial correlation (i.e. 'sticky' disclosure). Table V presents the tests of information demand on the number of total, new, and removed risk factors.

[Table IV about here.]

The results in Table IV suggest that managers respond to private enforcement by increasing the number of new risk factors they identify, which results in an increase in the total number of risk factors disclosed. In subsequent years, firms continue identifying more new risk factors and also increase the number of risk factors they remove from their disclosures, resulting in an overall increase in total risk factors that persists for multiple years. Together, these results support my third hypothesis that private enforcement actions evoke the disclosure of more risk factors. Additionally, the subsequent increase in dropped risk factors is consistent

[28]Note the number of observations is reduced from the unconditional table due to both data availability of the control variables and the further omission of perfectly predicted outcomes.

with the suggestion in *Slayton vs. American Express* that risk factors should be continually updated. Together, the results in Table IV also support the conclusion that managers believe risk factors to be a potential litigation shield. This result is in contrast to the study by Rogers and Van Buskirk (2009), which finds that firms actually decrease their provision of voluntary disclosure in response to shareholder litigation. My novel evidence of disclosure increase is significant because it points to a potentially less risky substitute for the litigation shield disclosures studied in previous literature (e.g., Skinner 1994, 1997; Francis, Philbrick, and Schipper 1994; Field, Lowry, and Shu 2005; Rogers and Van Buskirk 2009; Cutler, Davis, and Peterson 2016).

Table IV also provides evidence consistent with the argument that the SEC comment letter process does not expand the set risk factors identified by firms. This may be due to the limited capacity or information set of the SEC (Cox, Thomas, and Kiku 2003; Jackson and Roe 2009), or may indicate that the SEC does not dictate risk factor disclosures on a regular basis.[29] While firms do increase the number of new risk factors they identify, they also increase the number they remove by roughly the same proportion, resulting in no net increase. This may be an indicator of the SEC requesting improved risk factor organization or more detailed headings from firms, which my proxies would pick up as equal number of new and removed disclosures.[30] The influence of the SEC seems to be short lived, as only the filing immediately following the comment letters are materially affected; subsequent filings return to the 'normal' firm level of disclosure updates.

The third and fourth hypotheses also predict an increase in definitiveness in response to public and private enforcement actions. To test this, I again regress the textual features on indicator variables for public and private enforcement actions for the previous three years. Rather than regressing these features and the controls on the change in definitiveness, I allow

[29] A notable exception to this practice is the Cyber-security risk factor disclosure guidance issued in 2011. See: www.sec.gov/divisions/corpfin/guidance/cfguidance-topic2.htm.

[30] For example, the comment letter to Target Corporation, dated September 23, 2008, Exhibit (99)A. Accessed at: https://www.sec.gov/Archives/edgar/data/27419/000000000009001113/filename1.pdf.

for more flexible serial correlation by using the level as a dependent variable, with the lagged value as an additional control. These results are presented in Table V.

[Table V about here.]

The results in Table V provide only weak evidence in support of the third hypothesis with regards to increased definitiveness. The marginally significant coefficients on Securities Litigation$_{t-1}$ in specifications (1) and (2) suggest that private enforcement, at most, elicits a weak increase in level of detail with which managers disclose their risk factors. Together with the evidence from Table IV, this suggests that managers believe the legal value of risk factors stems from the disclosure of individual risk factors, rather than the level of detail. One potential reason is that firms facing litigation already have a high level of specificity, such that managers interpret the private enforcement as a signal of insufficient identification of risk factors, rather than insufficient definitiveness. Alternatively, firms could be withholding specificity because they believe that providing definitive expositional disclosure opens them up to litigation risk. Consistent with the former explanation, firms facing a securities litigation have on average 8.2% and 5.1% higher levels of specificity and numeric intensity, respectively, when compared to the non-litigation firm years, and both differences are significant at the five percent level. The insignificant results in Specification (3) suggests that managers do not believe that increasing the verbosity of risk factors is demanded by investors, which is consistent with the results in Campbell, Chen, Dhaliwal, Lu, and Steele (2014) that verbosity is associated with increased return volatility. Together, the results do not provide strong evidence that firms react to private enforcement by increasing the definitiveness of their risk factor disclosures, which is somewhat puzzling given the regulatory focus and court precedence, which both highlight the importance of detail in cautionary language.

The results in Table V provide strong evidence consistent with the fourth hypothesis that managers react to public enforcement by increasing the definitiveness of their risk factor disclosures. Firms significantly increase the specificity, numeric intensity, and length of each risk factor in response to SEC comment letters. These effects seem to be temporary, with the increase only occurring in the disclosure immediately succeeding the comment letter, with the exception of numeric intensity, which persists for two years. The results in Specification (3), which suggest that firms increase the verbosity of their risk factor disclosures, complements previous findings by Beatty, Cheng, and Zhang (2015) and Brown, Tian, and Tucker (2018), who both find that risk disclosure increases on average after an SEC comment letter. My results potentially suggest the channel for this increase, that while the length of disclosure increases, the number of identified factors does not. The regressions control for lagged values of the dependent variables, thus effectively demonstrate an increase in the *change* of definitiveness, not whether these increases in definitive detail persist. To test the persistence, in an untabulated regression I repeat this analysis without controlling for the lagged values of the dependent variables, and find that these increases in definitiveness do persist and are significant for multiple years, which suggests that firms do not undo the improvements in definitiveness that the SEC evokes.

Together, these results suggest that when managers receive an SEC comment letter about risk factor disclosure inadequacies, they respond by updating their disclosures in the immediately subsequent annual report, but then go back to their *normal* update procedure (i.e. the SEC does not cause firms to fundamentally to alter the underlying data generating process). However it is important to note that while a comment letter only temporarily changes how firms *update* their disclosures, these improvements to increased definitiveness persist: managers do not revert to the pre-comment letter disclosures. This is consistent with the expected costs of the SEC review process, which can evoke immediate changes, but pose few continuing costs once the *no further comment* letter is received. In contrast, private litigation does appear to have a continuing effect on the *updates* that persists for multiple years, consistent with the expectation of future litigation costs being significantly higher (Rose 2008). The results suggest that managers adjust the how they re-address their risk factors year over year (i.e. changing the underlying data generating process), for example by lowering the threshold of what qualifies as a *significant factor*, leading to an increase in total number of risks they identify. This contrast between the SEC evoking a 'one-off' change to disclosure updates and securities litigation evoking a more lasting change to disclosure updates is potentially of relevance in the ongoing legal literature on whether securities litigation is a useful governance mechanism (Rose 2008).

V. ROBUSTNESS

Alternative Specifications

Continuous Outcomes

One concern stemming from my approach to studying the predictive ability over extreme adverse outcomes is that these tail events may be considered too extreme to connote meaningfully informative disclosures. If, as I hypothesize, managers' disclosure decisions are driven by their subjective assessment of cash flow distributions (or other outcomes, such as general business litigation), it may equally reasonable to assume that the disclosure decision is related to the continuous underlying economic process, rather than just the discrete adverse events on which I focus. To test whether risk factor disclosures signal future cash flow changes in general, rather than just in the extreme, I regress continuous measures of accounting performance on the risk factor disclosure measures. To capture the potential for risk disclosures about both revenue declines and cost increases, I study the following accounting variables (all scaled by total assets): *Net Income*, *Operating Income*, *Sales*, and *ROA*. The regressions, presented in Table VI, also control for the contemporary value of the outcome variable, measured at the time of the risk factor disclosure, to capture the autoregressive nature of these variables.

[Table VI about here.]

The results in Table VI suggest that disclosing new risk factors is strongly correlated with declines in all four outcomes, consistent with expectations of risk disclosures containing adverse information. This relationship holds in both the cross section and the time series. Similarly, removing risk factor disclosures is associated with increases in the income measures (not so for $Sales_{t+1}$), and this relationship holds more strongly in the time series, suggesting that the information in the decision to remove disclosures is primarily idiosyncratic. There is a significant relationship between the *Specificity* of risk factor disclosures and future declines the accounting outcomes (and to a lesser extent the verbosity), but not so for the *Numeric Intensity*. One explanation may be that firms update their qualitative information across a broader spectrum of adverse outcomes, and provide more quantitative information in advance of significantly negative adverse outcomes, consistent with Heinle and Smith (2017). In general, these results support the idea underlying the hypotheses that managers are making risk factor disclosure decisions based on their expectations of future cash flows.

Repeated Adverse Outcomes

One concern of studying the predictive ability of risk factor disclosures is that managers' may actually be reacting to the observed adverse outcomes which happen to persist for multiple years. For example a firm suffering a net loss may contemporaneously disclose potential causes in their risk factor section, but if the loss persists into the next year, these risk factor disclosure changes may appear to 'predict' the loss. To mitigate this concern, I repeat the same predictive tests but reduce the sample of firm-year observations by omitting repeated observations. This effectively reduces the sample to firm years leading up to an adverse outcome, but does not continue to include firm years if the adverse outcomes persist. The results, presented in Table VII, show that even when only considering the first occurrence of an adverse event, managers still increase their provision of new risk factors and decrease their removal of prior risks. Interestingly, the significance of the lagged number of risk factors declines sharply, potentially because as adverse outcomes are repeated, the predictive risk factor disclosures immediately preceding the adverse outcome move into the 'lagged' risk factor disclosures, but are still predictive of the repeated outcomes. The decreased significance of the definitiveness results may suggest that firms increase in definitiveness is slower to respond to future expectations, and that firms continue to add more detail as adverse events persist. Together these results confirm the conclusion that firms on average disclose new risk factors *in advance* of adverse events, rather than merely reacting to the past occurrence thereof.

[Table VII about here.]

Specific Risk Factor Topics

The previous sections treat individual risk factors as fungible, and that they have equal weighting in their disclosure, as well as their predictive ability over adverse outcomes. This is a strong assumption because different risk factors address different adverse events, which potentially have differing impacts on firm outcome. To investigate to what extent the conclusions of predictive ability of risk factors generalizes to the various different risks being disclosed, I perform a topic analysis to attempt to relate specific topics to specific outcomes. To classify risk factor disclosures and determine what risks they disclose, I use the Latent Dirichlet Allocation (LDA) technique developed by Blei, Ng, and Jordan (2003) to discover and extract the 24 most prominent topics (see Internet Appendix **??** for the list of topics and top-weighted words). The idea behind LDA is that different topics have a linguistic 'fingerprint' corresponding to the words commonly used within the topic's domain, and that a document is composed of some mixture of these topics.[31] To provide generalizable evidence that firms disclose *specific* risk factors in advance of the *specific* adverse outcomes to which they relate, I chose two sets of topics and outcomes, one 'micro' factor which is related to the firm, and one 'macro' factor which is related to the economy as a whole.

The micro-level adverse outcome I focus on which is plausibly related to an extracted topic is the presence of a goodwill impairment. Since the introduction of SFAS 142 in 2001, firm's goodwill has to undergo a yearly impairment test.[32] Failing this test can be significantly detrimental to the firm's net income, thus early warning of potential impairments could provide legal benefit to managers. Hayn and Hughes (2006) suggests that the economic indicators of goodwill impairment lead the accounting write-off, sometimes by multiple years.

[31] For more information on the derivation details, see Internet Appendix Section **??**.
[32] See www.fasb.org/summary/stsum142.shtml.

This suggests that managers may reasonably predict the presence or probability of future adverse impairment test outcomes. This potential for predictability of goodwill impairments presents an interesting setting in which I can test whether managers systematically warn of these adverse events through the risk factor disclosures, and specifically those topics pertaining to goodwill and impairments: *Strategic Alliance* and *Accounting*. The results of these tests, mirroring those performed in Table III, are presented in Table VIII.

[Table VIII about here.]

The results in Table VIII suggest that firms increase their disclosure of risks relating to *Strategic Alliance* (e.g. M&A activity) in advance of a goodwill write-off However there is not significant evidence of an increase in the discussion of *Accounting* terms, which includes goodwill and impairments. Given the findings in Hayn and Hughes (2006) that economic indicators of goodwill impairments precede the observed write-off by multiple years, it may be the case that firms disclose of goodwill impairments more than one year in advance. To test this, in untabulated results, I add a second lag of the accounting topic disclosures, and find a significant positive relationship (T-stat=2.19), suggesting managers potentially warn of impending impairments further in advance.

The micro-level adverse outcome I focus on is future changes in the price of oil. Because many industries are affected by the price of oil, the expected cost of significant shifts in the market price may be sufficient to warrant a risk factor disclosure. Testing the disclosure of risks relating to the *Energy* topic against future changes in the price of oil lends evidence as to whether managers incorporate their expectations about future outcomes in their disclosure decisions. The tests of the change of oil price on risk factors are presented in Table IX.[33]

[Table IX about here.]

[33]The oil price is measured using the end of month spot price for West Texas Intermediate (WTI), a common benchmark for oil pricing. Data are downloaded from the U.S. Energy Information Administration. Url: www.eia.gov/dnav/pet/pet_pri_spt_s1_d.htm.

The results in Table IX suggest that on average, firms disclose more risk factors specifically about energy topics in advance of price declines. Similarly, firms remove more risk factors in advance of subsequent oil price increases, consistent with managers 'informatively' updating their set of disclosed risk factors by actively removing obsolete risk factor disclosures. The net increase in the disclosure of risk factors relating to the *Energy* topic is also negatively associated with oil price changes, suggesting firms increase their overall level of energy related risk factors ahead of price declines.

In general, the results of robustness tests conducted on the individual topic level are consistent with managers disclosing *specific* risk factors in advance of adverse outcomes they address. However in untabulated results, replacing the aggregated risk factor measures in the main regressions with the topic-level measures results in only a subset of topics being significantly predictive. This may suggest that only some topics have systematic relationships with the adverse outcomes I test in this study. An alternative interpretation is that some topics are less informative of short-term cash-flow outcomes, and the predictive framework employed herein is inflexible in accounting for differing horizons of forecasting. While future research looking at predictability by topic might find that specific subsets of risk factor disclosures are more predictive of specific outcomes, it could also be the case that aggregate change in risk factors is a more 'informative' measure overall because it provides a more complete picture of managers assessed (and disclosed) beliefs. However, the evidence presented herein generally suggests that both the aggregate and individual topic level risk factors are disclosed in a manner consistent with the hypotheses that managers disclose individual risk factors to warn of specific adverse outcomes.

VI. CONCLUSION

This study addresses the question of whether risk factor disclosures are informative by reconsidering both the measure used to capture the risk factors, and the test employed to determine 'informativeness.' My results suggest managers disclose timely and definitive risk factors to warn of adverse outcomes, in accordance with the regulatory requirement. I find that the addition and removal of risk factors predicts future adverse economic outcomes even after controlling for ex-ante risk and firm performance, and that firms do in fact *remove* disclosures in an timely manner as well. My approach differs from previous literature in that researchers have typically aggregated the risk disclosure using a bag of words approach, and primarily focused on market outcomes. My results suggest that treating risk factors as distinct units may be a more faithful representation of the information being conveyed by managers.

Given the validation of this time-series approach to measuring the information in risk factors, I use my approach study the demand for these disclosures from two significant sources: investors and regulators. I focus on demand 'shocks' through public enforcement via SEC comment letters, and private enforcement from investors via securities litigation. Consistent with my predictions, I find that firms respond to private enforcement by expanding the set of risk factors they identify, but surprisingly do not increase the definitiveness of those disclosures. In contrast, I find firms respond to public enforcement not by increasing their identified risk factors as suggested in previous literature, but by improving their level of specificity and detail. In line with the cost benefit tradeoff to managers, I find that the response to private enforcement actions persists for multiple years, while public enforcement actions generally only effect changes in the subsequent filing. Together, these results shed light on the beliefs of managers as to the demand for risk factors disclosures from both regulators and investors. These results suggest that managers believe risk factor disclosures to be effective litigation shields for investors, and I provide evidence consistent with this belief.

My study offers three main takeaways. First, my paper suggest that the SEC effectively elicits more detail in disclosures. While this may lead to the conclusions found in other studies, that the length of disclosure is less correlated with underlying economic activity, my findings suggest that this focus on length may be falsely attributing regulatory impact to less informative disclosures. Additionally, my results point to an interesting role that risk factor disclosures may play as litigation shields, potentially substituting for voluntary information when the cost of the latter increases. Second, this study demonstrates that managers update their disclosures in a timely fashion, suggesting that they are not merely 'copy and pasting.' I find that firms are informatively choosing when to add and, importantly, remove information, and actively supply disclosure to meet information demand. These results suggest caution against interpreting the increasing length of disclosure as a sufficient signal of information overload; in some cases, a more targeted measure of information may be warranted. Last, my study demonstrate the potential value of considering the time series evolution of disclosures, and carefully constructing textual measures based on the data generating process underlying the disclosure.

References

Amihud, Y. 2002. Illiquidity and Stock Returns: Cross-Section and Time-Series Effects. *Journal of Financial Markets* 5 (1): 31–56.

Bao, Y. and A. Datta. 2014. Simultaneously Discovering and Quantifying Risk Types from Textual Risk Disclosures. *Management Science* 60 (6): 1371–1391.

Beatty, A., L. Cheng, and H. Zhang. 2015. Sometimes Less Is More: Evidence from Financial Constraints Risk Factor Disclosures. *Working Paper* The Ohio State University.

Beatty, A., L. Cheng, and H. Zhang. 2018. Are Risk Factor Disclosures Still Relevant? Evidence from Market Reactions to Risk Factor Disclosures before and after the Financial Crisis. *Contemporary Accounting Research* (forthcoming).

Bird, A. and S. A. Karolyi. 2016. Do Institutional Investors Demand Public Disclosure? *Review of Financial Studies* 29 (12): 3245–3277.

Blackwell, D. 1951. Comparison of Experiments. In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, volume 1, 93–102.

Blei, D. M., A. Y. Ng, and M. I. Jordan. 2003. Latent Dirichlet Allocation. *Journal of Machine Learning Research* 3: 993–1022.

Bloomfield, R. 2008. Discussion of "Annual Report Readability, Current Earnings, and Earnings Persistence". *Journal of Accounting and Economics* 45 (2–3): 248–252.

Bloomfield, R., M. W. Nelson, and E. Soltes. 2016. Gathering Data for Archival, Field, Survey, and Experimental Accounting Research. *Journal of Accounting Research* 54 (2): 341–395.

Boone, A. and J. White. 2015. The Effect of Institutional Ownership on Firm Transparency and Information Production. *Journal of Financial Economics* 117 (3): 508–533.

Bozanic, Z., J. R. Dietrich, and B. A. Johnson. 2017. SEC Comment Letters and Firm Disclosure. *Working Paper* The Ohio State University.

Brown, S. V., X. Tian, and J. W. Tucker. 2018. The Spillover Effect of SEC Comment Letters on Qualitative Corporate Disclosure: Evidence from the Risk Factor Disclosure. *Contemporary Accounting Research* 35 (2): 622–656.

Brown, S. V. and J. W. Tucker. 2011. Large-Sample Evidence on Firms' Year-over-Year MD&A Modifications. *Journal of Accounting Research* 49 (2): 309–346.

Campbell, J. L., M. Cecchini, A. M. Cianci, A. C. Ehinger, and E. M. Werner. 2018. Tax-Related Mandatory Risk Factor Disclosures, Future Profitability, and Stock Returns. *Review of Accounting Studies* (forthcoming).

Campbell, J. L., H. Chen, D. S. Dhaliwal, H.-m. Lu, and L. B. Steele. 2014. The Information Content of Mandatory Risk Factor Disclosures in Corporate Filings. *Review of Accounting Studies* 19 (1): 396–455.

Cheng, C. A., H. H. Huang, and Y. Li. 2016. Does Shareholder Litigation Deter Insider Trading? *Journal of Law, Finance, and Accounting* 1 (2): 275–318.

Chiu, T.-T., Y. Guan, and J.-B. Kim. 2015. The Effect of Mandatory Risk Factor Disclosures on the Pricing of Credit Default Swaps. *Working Paper* .

Cox, J. D., R. S. Thomas, and D. Kiku. 2003. SEC Enforcement Heuristics: An Empirical Inquiry. *Duke Law Journal* 53 (2): 737–779.

Cutler, J., A. K. Davis, and K. Peterson. 2016. Disclosure and the Outcome of Securities Litigation. *Working Paper* University of Houston.

D'Souza, J., K. Ramesh, and M. Shen. 2009. Disclosure of GAAP Line Items in Earnings Announcements. *Review of Accounting Studies* 15 (1): 179–219.

Dyer, T., M. Lang, and L. Stice-Lawrence. 2017. The Evolution of 10-K Textual Disclosure: Evidence from Latent Dirichlet Allocation. *Journal of Accounting and Economics* 64 (2): 221–245.

Field, L., M. Lowry, and S. Shu. 2005. Does Disclosure Deter or Trigger Litigation? *Journal of Accounting and Economics* 39 (3): 487–507.

Filzen, J. 2015. The Information Content of Risk Factor Disclosures in Quarterly Reports. *Accounting Horizons* 24 (4): 887–916.

Filzen, J., G. McBrayer, and K. Shannon. 2016. Risk Factor Disclosures: Do Managers and Markets Speak the Same Language? *Working Paper* Boise State University.

Francis, J., D. Philbrick, and K. Schipper. 1994. Shareholder Litigation and Corporate Disclosures. *Journal of Accounting Research* 32 (2): 137–164.

Hayn, C. and P. J. Hughes. 2006. Leading Indicators of Goodwill Impairment. *Journal of Accounting, Auditing & Finance* 21 (3): 223–265.

Heinle, M. S. and K. C. Smith. 2017. A Theory of Risk Disclosure. *Review of Accounting Studies* 22 (4): 1459–1491.

Hope, O.-K., D. Hu, and H. Lu. 2016. The Benefits of Specific Risk-Factor Disclosures. *Review of Accounting Studies* 21 (4): 1005–1045.

Huang, K.-W. and Z. Li. 2008. A Multilabel Text Classification Algorithm for Labeling Risk Factors in SEC Form 10-K. *ACM Transactions on Management Information Systems* 2 (3): 18:1–18:19.

Jackson, H. E. and M. J. Roe. 2009. Public and Private Enforcement of Securities Laws: Resource-Based Evidence. *Journal of Financial Economics* 93 (2): 207–238.

Johnston, R. and R. Petacchi. 2017. Regulatory Oversight of Financial Reporting: Securities and Exchange Commission Comment Letters. *Contemporary Accounting Research* forthcoming.

Kravet, T. and V. Muslu. 2013. Textual Risk Disclosures and Investors' Risk Perceptions. *Review of Accounting Studies* 18 (4): 1088–1122.

Lukomnik, J. 2016. Risk Factor Disclosure Landscape. Technical report, Investor Responsibility Research Center Institute.

Naughton, J. P., T. O. Rusticus, C. Wang, and I. Yeung. 2015. Private Litigation Costs and Voluntary Disclosure: Evidence from the Morrison Ruling. *Working Paper* Northwestern University.

Nelson, K. and A. C. Pritchard. 2016. Carrot or Stick? The Shift from Voluntary to Mandatory Disclosure of Risk Factors. *Journal of Empirical Legal Studies* 13 (2): 266–297.

Robbins, R. B. and P. L. Rothenberg. 2005. Writing Effective Risk Factor Disclosure in Offering Documents and Exchange Act Reports. *Insights: The Corporate and Securities Law Advisor* 19 (5): 9–19.

Rogers, J. L. and A. Van Buskirk. 2009. Shareholder Litigation and Changes in Disclosure Behavior. *Journal of Accounting and Economics* 47 (1–2): 136–156.

Rose, A. M. 2008. Reforming Securities Litigation Reform: Restructuring the Relationship Between Public and Private Enforcement of Rule 10b-5. *Columbia Law Review* 108 (6).

Skinner, D. J. 1994. Why Firms Voluntarily Disclose Bad News. *Journal of Accounting Research* 32 (1): 38–60.

Skinner, D. J. 1997. Earnings Disclosures and Stockholder Lawsuits. *Journal of Accounting and Economics* 23 (3): 249–282.

White, M. J. 2013. The Path Forward on Disclosure.

Wooldridge, J. 2010. Correlated Random Effects Models with Unbalanced Panels. *Working Paper* Michigan State University.

Appendix A Variable Descriptions

The following table defines the variables used in this paper. Variable names and calculations provided in brackets correspond to source database. For the regressions presented in the tables, the continuous variables are winsorized at the 1st and 99th percentiles.

Variable	**Description**
Assets	Total assets $\{AT\}$
Market Equity	Market value of equity at fiscal year end. $\{CSHPRI * PRCC\_F\}$
Net Income/AT	Net income to lagged assets. $\{NI_t/AT_{t-1}\}$
ROA	Return on assets. $\{IB_t/AT_{t-1}\}$
Operating Inc./AT	Operating income (after depreciation) to lagged assets. $\{OIADP_t/AT_{t-1}\}$
Sales Growth	Ratio of change in sales to lagged assets. $\{(SALE_t - SALE_{t-1})/AT_{t-1}\}$
Sales / AT	Sales to lagged assets. $\{SALE_t/AT_{t-1}\}$
Book Equity	Book value of equity. $\{(TEQ_t\|AT_t - LT_t) + TXDITC_t - PSTK_t\}$
Book-to-Market	Book to market. $\{\text{Book Equity}_t/\text{Market Equity}_t\}$
Leverage	Leverage $\{(DLTT_t + DLC_t)/AT_t\}$
Tangibility	Tangibility $\{PPENT_t/AT_t\}$
Turnover	Ratio of average daily volume (CRSP) to outstanding shares at fiscal year end (Compustat).
Beta	Market loading from CAPM model of daily returns on value weighted index, for all available days in the fiscal year. $\{RET = \alpha + \beta \cdot VWRETD + \epsilon\}$
Excess Returns	Cumulative excess daily returns during fiscal year. $\{(\prod RET - VWRETD + 1) - 1\}$
Excess Ret. Std.	Standard deviation of daily excess returns. $\{Std.\ Dev(RET - VWRETD)\}$
Min. Excess Ret.	Minimum daily excess return during fiscal year. $\{Min(RET - VWRETD)\}$
Excess Ret. Skew	Skewness of daily excess returns. $\{Skew(RET - VWRETD)\}$
$\text{CAR}_{+3\ day}$	Cumulative abnormal return spanning three business days starting on the filing date of the annual report, using a Fama-French Carhart factor model. Data from Kenneth French's website. $\{RET = MKTRF + SMB + HML + UMD + MOM + \epsilon\}$
$\text{CAR}_{+3\ months}$	Cumulative abnormal return spanning 60 business days starting on the filing date of the annual report, using a Fama-French Carhart factor model. Data from Kenneth French's website. $\{RET = MKTRF + SMB + HML + UMD + MOM + \epsilon\}$
$\text{Bid-Ask Spread}_{+1\ year}$	Average daily bid ask spread at closing for 240 business days starting on the filing date of the annual report, as percentage of average bid and ask. $\{200 * (ASK - BID)/(ASK + BID)\}$
Illiquidity	Amihud Illiquidity (2002), calculated as $\frac{1}{D_{iy}} \sum_{t=1}^{D_{iy}} \|R_{iyd}\| / Volume_{iyd}$, where D_{iy} is the number of days over which the illiquidity is calculated, and R_{iyd} and $Volume_{iyd}$ are the return and trading volume, respectively, on a given day.

Continued on next page...

...continued from previous page.

Variable	Description
Negative NI	Indicator variable equal to 1 if next year's net income is negative. $\{NI_{t+1} < 0\}$
Negative Op. Inc.	Indicator variable equal to 1 if next year's operating income is negative $\{OIADP_{t+1} < 0\}$
Sales Decline	Indicator variable equal to 1 if next year's sales are lower than the current year's sales by 10% or 10 million dollars, whichever is larger. $\{SALE_{t+1} - SALE_t < -10\% * max(100, SALE_t)\}$
Security Litigation	Indicator variable equal to 1 if a securities litigation is filed in the subsequent year.
Lawsuit Intensity	Natural log of number of litigation events found in CapitalIQ Key Developments database in a given fiscal year.
RF Comment	Indicator variable equal to 1 if an SEC comment letter is received in a given fiscal year that references risk factors.
Comment (Any)	Indicator variable equal to 1 if any SEC comment letter is received in a given fiscal year, including those with a reference to risk factors.
FPS Industry	Indicator variable equal to 1 if the firm has an SIC code in one of the high litigation risk industries defined in Francis, Philbrick, and Schipper (1994).
# Risk Factors	Total number of risk factors disclosed under Item 1A of an annual report.
# New RF	Number of new risk factors which were not present in the previous year's annual report.
# Dropped RF	Number of risk factors which were in the previous year's annual report, but are no longer included in the present year.
# Kept RF	Number of risk factors which were in the previous year's annual report and persist in the current annual report.
Δ # RF	Net change in the number of risk factors, equal to *# New RF - # Dropped RF*.
# of Words	Total number of words in Item 1A of an annual report (excluding stop words).
# of Sentences	Total number of sentences in Item 1A of an annual report.
# of Specific Words	Total number of words identified by the Stanford Named Entity Recognition algorithm as being in categories: Location, Person, Organization
# of Numerics	Total number of numbers in Item 1A of an annual report.
# of Words/RF	Ratio of the total number of words to the total number of risk factors in Item 1A of an annual report (excluding stop words).
Specificity	Ratio of the total number of specific words to the total number of risk factors in Item 1A of an annual report (excluding stop words).
Numeric Intensity	Ratio of the total number of numbers to the total number of risk factors in Item 1A of an annual report (excluding stop words).
FOG Index	Gunning Fog score for the text in Item 1A of an annual report (excluding stop words). Calculated as $\left\{0.4\left(\frac{\#\ of\ words}{\#\ of\ sentences} + 100\frac{\#\ of\ complex\ words}{\#\ of\ words}\right)\right\}$

Appendix B Risk Factor Extraction

To derive an initial list of filings, I extract the gvkey and historical CIKs from the Compustat annual file, starting from 89,687 firm years. I then merge the non-missing CIKs from Compustat with the EDGAR filings index file provided by the SEC.[34] I filter the index files to include only form 10-Ks, excluding amended 10-K/As, leaving 67,648 filings from 9,632 firms. Form 10-KSB is also included, but are no longer filed after 2009.

I then search through these filings to extract *Item 1A: Risk Factor* section. Filings, as they are submitted to the EDGAR system, comprise an SGML header (with information about the filing such as company identifier, name, SIC code, and period with which the filing is associated), and multiple 'documents' that are typically the main form and exhibits. I extract the first document in a filing, which is the 10-K report, but ignore the accompanying exhibits. This potentially biases against extracting the risk factor section if it is included in Exhibit 13, but is in line with the other studies of risk factors that omit sections included by reference. Similar to the approach in Campbell et al. (2014), I extract sections by assuming visual prominence of Item headings through font and whitespace delineation. Thus, I also exclude filings that are not submitted in HTML format because they omit the visual features required to reliably identify the sections in annual reports.

The general algorithm I use for extracting a specific section in an annual or quarterly report is as follows. First, I identify the Item headings, which should occur in order throughout the document (omitting them when they occur in the table of contents, whether it is at the beginning or end of the document). I identify headings as being the only text on a 'line' of text, in the format Item 1A: Risk Factors (with flexibility in punctuation), and emphasized with either bold or underlined font. Once I identify Item 1A and either Item 1B or Item 2 (whichever is found and comes first), I extract the text between the first instance of Item 1A

[34]Located, for example, at ftp.sec.gov/edgar/full-index/2005/QTR1/master.idx. Downloaded using script from github.com/gaulinmp/pyedgar

to first instance of the next non 1A Item header. Some firms repeat the Item 1A header at the top of each page of the section, thus I include all the text until the next item number.

This method is based on visual identification of features that stand out to a human reader of the rendered document. The HTML format of annual reports allows for programmatic identification of visual elements through the Document Object Model, or DOM. HTML DOM is a platform independent graph based method for describing content, primarily textual, in the case of HTML (but also arbitrary data in XML, one example of which being XBRL). Each element in a page is a node in the DOM graph, with parents that fully contain it, and potentially children it fully contains. Nodes in the graph, called HTML tags, have multiple types that each have different features. For example, the paragraph tag <p> defines a paragraph which is separated/isolated visually from text above and below (children nodes or parent nodes), while the font tag <font> defines a subset of words (potentially children of a parent paragraph tag) that have a specific font, but are not isolated with whitespace from the surrounding text. These features of the different tag types allow for a researcher to programmatically extract data based on what a human reader sees as whitespace or emphasis when viewing the rendered document in a web browser. Specifically, I use the beautifulsoup library to handle the HTML DOM which allows for traversing tag nodes in a graph-like format.

For each HTML annual report, I iterate through all instance of an HTML tag containing the text *Item 1A*, *1B*, or *2*, using the case insensitive regular expression (`item(?:[^ a-z0-9]|&(?:nbsp|#160);)*(1[AB]|2)`).[35] The ` ` and ` ` are HTML encoded representations of a non-breaking space, which are necessary to include when searching through raw HTML encoded documents. By only searching for Items 1A, 1B, and 2, I impose the assumption that items are presented in order. The result of this search is then checked for emphasis, based on either HTML emphasis tags or CSS styles.[36] Those items that are

[35] Acknowledging and accepting the limitations to parsing HTML with regular expressions.

[36] HTML emphasis tags: b, em, strong, h1, h2, h3, h4, h5, h6, u. CSS Styles *bold* or *underline* within HTML tags: p, font, div, span, li.

emphasized are then checked for whitespace separation. Whitespace separation is defined as the tag being at the beginning of a visual line, or as described in Campbell, Chen, Dhaliwal, Lu, and Steele (2014) as *segmentation*.

To identify whitespace separation, I iterate up the HTML DOM to find the first parent node that is visually separated. This includes the tags: `h1`, `h2`, `h3`, `h4`, `h5`, `h6`, `p`, `div`, `ul`, `ol`, `tr`, or `table`. While it is possible that `div` tags are not visually isolated (using `float` CSS styling), in a non-exaustive search of HTML documents I did not find any instances of floats. Additionally, while some documents use paragraphs and CSS as mentioned above to lay out their annual report, other filings employ tables to do so. The difference is akin to using tab-stops in Microsoft word to put text on both the left and right side of a document, or alternatively creating a table with two columns. For firms that use the latter table method, merely checking within the immediate parent table cell may result in false conclusions as to the visual layout. Therefore, when an item is included in a table, I include the entire row as the parent element, rather than just the cell. This separating parent element thus contains the 'block' of text in which the word]textitItem occurs and is emphasized, be it just the header *Item 1A: Risk Factors*, or a paragraph containing a reference to the Item 1A section. To filter out the latter false positives (including table of contents matches), I require that the full (plain) text in the separating element comprise solely the Item number and description using regular expressions.[37]

The result of this step is the elements containing the Item headers for 1A, and 1B or 2. I then keep all of the HTML code between the beginning of the first Item 1A header and the beginning of the first Item 1B or 2 header, whichever occurs first. These extracted sections represent the Item 1A for each given filing. To match the filings to their associated Compustat fiscal years, I extract the period assigned with the filing from the heading of the EDGAR document. To do so, I search in the first 5,000 bytes of the SGML header

[37] Item 1A: `/^\s*item[^a-z0-9]*1A[^a-z]*risk[^a-z]*factors?\s*$/`,
Item 1B: `/^\s*item[^a-z0-9]*1B[^a-z]*((?:unresolved|sec|staff|comments)[^a-z]*)+\s*$/`,
Item 2: `/^\s*item[^a-z0-9]*2.{,20}propert(y|ies)\D{,35}$/`

provided in the raw daily feed file for the `PERIOD` header. I keep the filing if the `PERIOD` header contains a string which conforms to an eight character date format (YYYYMMDD). I match the EDGAR filings to the Compustat data based on CIK (exact match) and period date (within a date window). I require the period listed in the filing header be within five business days of the `datadate` variable in Compustat. While greater than 90% of the filings have an exact match between the reported period and `datadate`, there are some discrepancies because Compustat always sets the `datadate` variable equal to the last day of a month. To verify this assumption, I manually selected a random sample of 50 reports with a date match 5 days apart, and found that all were matched to the appropriate fiscal year in Compustat.

Figure I: Net Change in Risk Factors

Figure I shows the histogram of net change in total risk factors. The sample comprises firm-years with a non-missing risk factor section in the previous year.

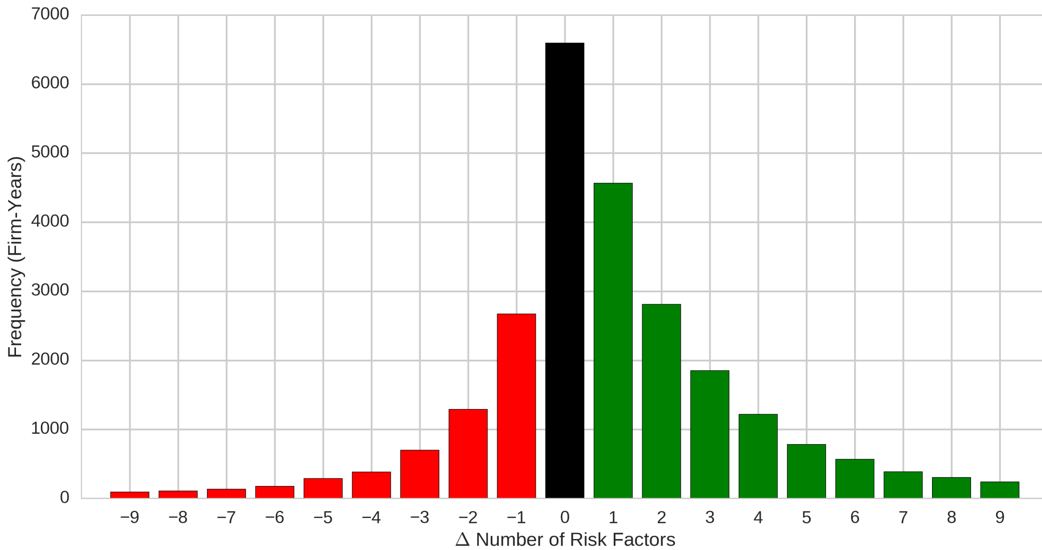

Table I: Summary Stats

Table I reports the summary statistics for the variables used in the regressions as defined in Appendix A.

	Mean	Std. Dev	Min	25%	50%	75%	Max	N
$Log(Assets)_t$	6.70	2.01	2.12	5.34	6.74	8.05	11.56	26,547
$Log(Market\ Equity)_t$	6.33	1.96	1.84	4.99	6.35	7.70	10.56	26,537
$Book-to-Market_t$	0.73	0.90	-0.71	0.31	0.58	0.96	4.60	26,518
$Net\ Income/AT_t$	-0.02	0.32	-1.09	-0.02	0.03	0.07	0.40	26,536
$Opr.\ Inc./AT_t$	0.02	0.28	-1.01	0.01	0.06	0.12	0.48	25,276
$Sales/AT_t$	0.90	0.85	0.00	0.26	0.70	1.27	4.04	26,536
$Leverage_t$	0.23	0.25	0.00	0.02	0.17	0.35	0.96	26,547
$Tangibility_t$	0.21	0.24	0.00	0.03	0.12	0.32	0.89	25,479
$Turnover_t$	0.87	0.78	0.04	0.33	0.66	1.13	4.15	26,538
$Beta_{t-1}$	1.04	0.54	-0.10	0.69	1.05	1.40	2.31	26,547
$Excess\ Ret._{t-1}$	0.02	0.48	-0.84	-0.24	-0.04	0.19	1.92	26,547
$Ex.\ Ret.\ Std_{t-1}$	0.03	0.02	0.01	0.02	0.02	0.04	0.10	26,547
$Min.\ Excess\ Ret._{t-1}$	-0.12	0.09	-0.47	-0.15	-0.09	-0.06	-0.02	26,547
$Ex.\ Ret.\ Skew_{t-1}$	0.44	1.45	-3.91	-0.16	0.34	0.95	6.18	26,545
$CAR_{+3\ days}$	-0.00	0.06	-0.21	-0.02	-0.00	0.02	0.21	26,375
$CAR_{+3\ months}$	0.00	0.25	-0.72	-0.12	-0.00	0.11	0.86	26,382
$Bid-Ask\ Spread_{t+1}$	0.75	1.57	0.02	0.07	0.16	0.58	8.40	26,386
Indicator and Negative Outome Variables								
$Negative\ NI_{t+1}$	0.31	0.46	0	0	0	1	1	22,026
$Negative\ OperatingInc._{t+1}$	0.23	0.42	0	0	0	0	1	20,970
$Sales\ Decline_{t+1}$	0.21	0.41	0	0	0	0	1	22,025
$Sec.\ Litigation_{t-1}$	0.02	0.14	0	0	0	0	1	26,546
$Lawsuit\ Intensity_{t-1}$	0.19	0.44	0	0	0	0	2.08	26,547
$RF\ Comment_{t-1}$	0.07	0.25	0	0	0	0	1	26,224
$Comment\ (Any)_{t-1}$	0.53	0.50	0	0	1	1	1	26,223
Textual Variables								
$\#\ Risk\ Factors_t$	29.52	14.05	6	19	27	37	69	26,547
$\#\ New\ RF_t$	3.67	5.03	0	1	2	5	25	26,547
$\#\ Dropped\ RF_t$	2.45	3.73	0	0	1	3	21	26,547
$\Delta\#\ RF_t$	1.22	4.56	-11	0	1	2	18	26,547
$\#\ of\ Words_t$	2731.17	1707.36	396	1478	2364	3570	8847	26,547
$\#\ of\ Sentences_t$	257.39	150.96	43	147	228	335	729	26,547
$\#\ of\ Specific\ Words_t$	210.10	182.34	12	89	162	273	887	26,547
$\#\ of\ Numerics_t$	58.24	47.43	3	25	46	77	225	26,547
$\#\ of\ Words/RF_t$	88.95	25.77	38.71	71.30	86.32	102.93	176.54	26,547
$Specificity_t$	5.45	3.25	0.60	3.21	4.76	6.90	17.88	26,547
$Numeric\ Intensity_t$	2.11	1.05	0.49	1.37	1.92	2.63	5.86	26,547
$FOG\ Index_t$	20.68	1.23	17.64	19.91	20.69	21.45	23.89	26,547

Table II: Adverse Outcome Predictability

Table II reports results from predictive regressions of future adverse outcomes on current risk factor disclosures. Specifications (1)–(3) present probit regressions with correlated random effects to control for average firm effects. Specification (4) presents an OLS regression with firm fixed effects, and the # of Events listed represents the number of firm years with at least one lawsuit. The coefficients in Specifications (1)–(3) represent the average marginal effects, evaluated at the mean of the dependent variables. The *Sign* column denotes the expected sign of the coefficients based on the hypotheses. Variables are defined in Appendix A.***, **, and * indicate significance at the 0.01, 0.05, and 0.10 level, respectively.

	Sign	Negative NI	Negative Op. Inc.	Sales Decline	Lawsuit Intensity
		(1)	(2)	(3)	(4)
$Log(\#\ New\ RF)_t$	+	0.022***	0.014***	0.014***	0.013***
		(5.54)	(4.09)	(3.63)	(3.58)
$Log(\#\ Dropped\ RF)_t$	–	-0.024***	−0.015***	−0.008*	−0.005
		(5.62)	(4.16)	(1.90)	(1.31)
$Log(\#\ RF)_{t-1}$	+	0.092***	0.061***	0.047***	0.044***
		(5.41)	(3.75)	(3.39)	(3.24)
$\Delta Specificity_t$	+	0.001	0.000	−0.002	0.004**
		(0.79)	(0.22)	(1.45)	(2.12)
$\Delta Numeric\ Intensity_t$	+	0.018***	0.011***	0.014***	−0.001
		(3.69)	(2.76)	(3.29)	(0.26)
$\Delta Log(\#\ of\ Words/RF)_t$	+	0.052**	0.039**	0.059***	0.064***
		(2.18)	(2.06)	(2.83)	(2.88)
Year F.E.		Y	Y	Y	Y
Correlated R.E.		Y	Y	Y	
Firm F.E.					Y
# Observations		21,683	20,618	21,682	26,984
# of Events		6,774	4,793	2,944	2,306

Table III: Adverse Outcome Incremental Predictability

Table III reports results from predictive regressions of future adverse outcomes on current risk factor disclosures, and controls for ex-ante risk. Specifications (1)–(3) present average marginal effects from probit regressions with correlated random effects to control for average firm effects. Specifications (4) presents an OLS regression with firm fixed effects. The $Dependent\ Var.\ (Level)_t$ variable is different for each specification. In specification order, it is: net income / total assets, operating income / total assets, sales / total assets, and lawsuit intensity. Variables are defined in Appendix A.***, **, and * indicate significance at the 0.01, 0.05, and 0.10 level, respectively.

	Negative NI	Negative Op. Inc.	Sales Decline	Log(# Suits)
	(1)	(2)	(3)	(4)
$Log(\#\ New\ RF)_t$	0.017***	0.009***	0.009**	0.012***
	(4.14)	(2.81)	(2.33)	(3.13)
$Log(\#\ Dropped\ RF)_t$	−0.018***	−0.011***	−0.007*	−0.005
	(4.14)	(2.93)	(1.71)	(1.24)
$Log(\#\ RF)_{t-1}$	0.036**	0.018	0.042***	0.045***
	(2.40)	(1.37)	(3.01)	(3.09)
$\Delta Specificity_t$	0.001	0.001	−0.002	0.004**
	(0.86)	(0.41)	(1.15)	(2.29)
$\Delta Numeric\ Intensity_t$	0.010*	0.006	0.010**	−0.001
	(1.92)	(1.39)	(2.31)	(0.15)
$\Delta Log(\#\ of\ Words/RF)_t$	0.061**	0.042**	0.042**	0.050**
	(2.48)	(2.19)	(1.98)	(2.25)
$Dependent\ Var.\ (Level)_{t-1}$	−0.209***	−0.113***	0.089***	−0.027**
	(5.60)	(7.29)	(9.12)	(2.47)
$Log(Market\ Equity)_t$	0.007	−0.000	−0.004	0.046***
	(0.76)	(0.02)	(0.64)	(6.45)
$Big\ N_t$	−0.000	−0.007	0.022	−0.031*
	(0.02)	(0.43)	(1.30)	(1.75)
$Book-to-Market_t$	0.052***	0.027***	0.011**	0.001
	(7.92)	(5.65)	(2.17)	(0.25)
$Tangibility_t$	0.168**	0.237***	0.071	0.132**
	(2.22)	(4.03)	(1.24)	(2.36)
$Leverage_t$	−0.116***	−0.079**	−0.011	−0.001
	(2.98)	(2.50)	(0.51)	(0.04)
$Turnover_t$	0.014*	0.019***	0.032***	0.040***
	(1.89)	(3.09)	(6.17)	(6.10)
$Beta_{t-1}$	0.002	−0.004	−0.007	−0.005
	(0.25)	(0.71)	(1.00)	(0.61)
$Excess\ Ret._{t-1}$	−0.059***	−0.028***	−0.063***	−0.022***
	(8.60)	(5.03)	(8.77)	(3.30)
$Ex.\ Ret.\ Std_{t-1}$	−0.194	−0.015	0.107	0.352
	(0.57)	(0.06)	(0.39)	(1.21)
$Ex.\ Ret.\ Skew_{t-1}$	−0.000	0.002	0.001	−0.005**
	(0.11)	(1.03)	(0.75)	(2.17)
Year F.E.	Y	Y	Y	Y
Correlated R.E.	Y	Y	Y	
Firm F.E.				Y
# Observations	20,303	19,945	20,302	25,252
# of Events	6,372	4,604	2,780	2,235

Table IV: Enforcement Response: Risk Factors

Table IV reports results from firm fixed effect regressions of risk factor evolution –total, new, and dropped– on public and private enforcement events. Variables are defined in Appendix A.***, **, and * indicate significance at the 0.01, 0.05, and 0.10 level, respectively.

	Log(# Total RFs)	Log(# New RFs)	Log(# Dropped RFs)
	(1)	(2)	(3)
Sec. Litigation$_{t-1}$	0.031***	0.178***	0.027
	(2.67)	(4.38)	(0.76)
—$_{t-2}$	0.004	0.099**	0.096**
	(0.44)	(2.48)	(2.48)
—$_{t-3}$	0.023***	0.145***	0.065*
	(3.05)	(4.04)	(1.87)
RF Comment$_{t-1}$	0.004	0.072***	0.063***
	(0.88)	(3.11)	(2.88)
—$_{t-2}$	0.005	0.006	0.004
	(0.98)	(0.26)	(0.21)
—$_{t-3}$	−0.005	−0.050**	−0.026
	(1.17)	(2.27)	(1.31)
Log(# RF)$_{t-1}$	0.532***	−0.438***	1.069***
	(33.00)	(10.38)	(27.15)
Log(# New RF)$_{t-1}$		−0.027***	
		(3.81)	
Log(# Dropped RF)$_{t-1}$			−0.091***
			(11.87)
Log(# of Words/RF)$_{t-1}$			
Log(Market Equity)$_t$	0.010***	0.010	−0.050***
	(2.66)	(0.68)	(3.65)
$Book-to-Market_t$	0.013***	0.037***	−0.001
	(4.80)	(3.32)	(0.09)
Comment (Any)$_{t-1}$	0.005**	0.031***	0.022**
	(2.03)	(2.79)	(2.22)
Sales Growth$_t$	0.018***	0.086***	0.067***
	(3.42)	(3.15)	(2.68)
Leverage$_t$	0.070***	0.285***	−0.017
	(4.74)	(4.59)	(0.34)
Turnover$_t$	0.006**	0.079***	0.059***
	(2.13)	(6.18)	(4.94)
Excess Ret.$_{t-1}$	−0.003	−0.018	0.039***
	(1.05)	(1.32)	(3.22)
Ex. Ret. Std$_{t-1}$	0.288*	4.327***	2.706***
	(1.94)	(6.54)	(4.33)
Ex. Ret. Skew$_{t-1}$	−0.002**	−0.000	0.004
	(2.11)	(0.12)	(1.02)
Year & Firm F.E.	Y	Y	Y
R^2	0.497	0.056	0.113
# Observations	23,209	23,209	23,209

Table V: Enforcement Response: Textual Features

Table V reports results from firm fixed effect regressions of risk factor textual features on public and private enforcement events. Variables are defined in Appendix A.***, **, and * indicate significance at the 0.01, 0.05, and 0.10 level, respectively.

	Specificity	Number %	Words / RF
	(1)	(2)	(3)
Sec. Litigation$_{t-1}$	0.154*	0.049*	0.666
	(1.87)	(1.82)	(1.03)
—$_{t-2}$	0.006	0.034	0.671
	(0.07)	(1.21)	(1.12)
—$_{t-3}$	−0.060	−0.036	−0.199
	(0.83)	(1.32)	(0.39)
RF Comment$_{t-1}$	0.125**	0.046***	0.895***
	(2.49)	(2.86)	(2.95)
—$_{t-2}$	−0.039	0.040***	0.373
	(0.87)	(2.75)	(1.29)
—$_{t-3}$	0.015	0.002	0.352
	(0.36)	(0.16)	(1.29)
Specificity$_{t-1}$	0.456***		
	(36.69)		
Numeric Intensity$_{t-1}$		0.448***	
		(38.94)	
# of Words/RF$_{t-1}$			0.501***
			(26.98)
Log(Market Equity)$_t$	−0.033	−0.073***	0.176
	(1.13)	(6.63)	(0.84)
$Book-to-Market_t$	0.008	0.012	0.329*
	(0.33)	(1.56)	(1.71)
Comment (Any)$_{t-1}$	−0.005	−0.005	0.032
	(0.24)	(0.68)	(0.22)
Sales Growth$_t$	−0.062	−0.061***	−0.204
	(1.35)	(3.58)	(0.65)
Leverage$_t$	0.364***	0.178***	3.611***
	(2.84)	(4.09)	(4.16)
Turnover$_t$	0.011	0.020**	0.309*
	(0.47)	(2.14)	(1.77)
Excess Ret.$_{t-1}$	−0.104***	−0.020**	−0.864***
	(4.01)	(2.18)	(4.63)
Ex. Ret. Std$_{t-1}$	4.471***	1.437***	30.408***
	(3.47)	(2.84)	(3.38)
Ex. Ret. Skew$_{t-1}$	0.005	−0.004	−0.029
	(0.59)	(1.54)	(0.56)
Year & Firm F.E.	Y	Y	Y
R^2	0.251	0.277	0.358
# Observations	23,201	23,201	23,201

Table VI: Risk Factor Disclosures and Continuous Accounting Outcomes

Table VI reports results from predictive regressions of future continuous accounting outcomes on the risk factor disclosure measures, and controls for ex-ante risk and accounting performance. Odd and even specifications correspond to cross sectional and firm fixed effects models respectively, and all dependent variables are scaled by total assets. Each specification includes $Dependent\ Var.\ (Level)_{t-1}$, which is the contemporary value of the continuous dependent variable. The results are robust to the omission of this variable. Variables are defined in Appendix A.***, **, and * indicate significance at the 0.01, 0.05, and 0.10 level, respectively.

	Net Income$_{t+1}$		Op. Inc.$_{t+1}$		Sales$_{t+1}$		ROA$_{t+1}$	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$Log(\#\ New\ RF)_t$	-0.015***	-0.015***	-0.013***	-0.012***	-0.040***	-0.028***	-0.015***	-0.015***
	(5.89)	(5.60)	(4.94)	(4.46)	(6.57)	(4.21)	(5.94)	(5.63)
$Log(\#\ Dropped\ RF)_t$	0.006*	0.012***	0.003	0.008***	0.003	0.007	0.006**	0.012***
	(1.94)	(4.01)	(0.99)	(2.59)	(0.49)	(0.99)	(1.99)	(3.93)
$Log(\#\ RF)_{t-1}$	-0.017***	-0.004	-0.011***	0.012	-0.009	0.041	-0.016***	-0.003
	(4.14)	(0.37)	(2.72)	(1.09)	(0.95)	(1.51)	(4.04)	(0.29)
$\Delta Specificity_t$	-0.003**	-0.002*	-0.003**	-0.002*	-0.007**	-0.004	-0.003**	-0.002**
	(2.52)	(1.85)	(2.49)	(1.88)	(2.35)	(1.51)	(2.56)	(1.99)
$\Delta Numeric\ Intensity_t$	-0.002	-0.001	-0.003	-0.001	-0.009	-0.006	-0.002	-0.001
	(0.63)	(0.32)	(0.89)	(0.23)	(0.99)	(0.78)	(0.62)	(0.22)
$\Delta Log(\#\ of\ Words/RF)_t$	-0.011	-0.029*	-0.007	-0.018	-0.097**	-0.068*	-0.012	-0.028*
	(0.62)	(1.83)	(0.38)	(1.14)	(2.43)	(1.89)	(0.71)	(1.81)
$Dependent\ Var.\ (Level)_{t-1}$	0.514***	0.037*	0.672***	0.174***	0.725***	0.266***	0.523***	0.041**
	(36.07)	(1.83)	(60.00)	(7.17)	(75.10)	(10.40)	(36.71)	(1.99)
$Log(Market\ Equity)_t$	0.017***	0.034***	0.014***	0.033***	0.011***	-0.033***	0.017***	0.034***
	(11.71)	(6.73)	(9.83)	(6.51)	(3.59)	(2.76)	(11.73)	(6.87)
$Big\ N_t$	-0.007	-0.026*	-0.008*	-0.026*	0.023**	-0.003	-0.007	-0.025
	(1.30)	(1.70)	(1.66)	(1.70)	(2.11)	(0.09)	(1.39)	(1.62)
$Book-to-Market_t$	0.000	-0.009**	0.002	-0.002	-0.013**	-0.036***	0.001	-0.008**
	(0.15)	(2.31)	(0.74)	(0.51)	(2.06)	(3.61)	(0.51)	(1.98)
$Tangibility_t$	0.031***	-0.035	0.029***	-0.035	0.068***	0.385***	0.032***	-0.035
	(3.15)	(0.85)	(2.95)	(0.85)	(3.18)	(3.78)	(3.22)	(0.84)
$Leverage_t$	-0.006	0.034	-0.024**	-0.006	-0.091***	-0.185***	-0.008	0.025
	(0.63)	(1.53)	(2.53)	(0.26)	(3.78)	(3.17)	(0.89)	(1.17)
$Turnover_t$	-0.001	0.006	0.000	0.010**	-0.016**	-0.008	-0.001	0.006
	(0.50)	(1.23)	(0.02)	(2.29)	(2.57)	(0.85)	(0.33)	(1.32)
$Beta_{t-1}$	0.007*	0.003	0.007*	0.002	0.023***	-0.004	0.007*	0.004
	(1.76)	(0.62)	(1.88)	(0.32)	(2.77)	(0.30)	(1.87)	(0.71)
$Excess\ Ret._{t-1}$	0.036***	0.017***	0.035***	0.015***	0.048***	0.066***	0.037***	0.017***
	(8.12)	(3.72)	(7.90)	(3.29)	(4.79)	(6.46)	(8.31)	(3.74)
$Ex.\ Ret.\ Std_{t-1}$	-1.120***	-0.317	-0.872***	-0.721***	-0.232	-0.923*	-1.120***	-0.346
	(5.74)	(1.35)	(4.63)	(3.06)	(0.56)	(1.71)	(5.79)	(1.49)
$Ex.\ Ret.\ Skew_{t-1}$	-0.010***	-0.009***	-0.009***	-0.009***	-0.017***	-0.012***	-0.010***	-0.009***
	(6.27)	(5.52)	(5.84)	(5.41)	(5.10)	(3.64)	(6.21)	(5.56)
$Year\ F.E.$	Y	Y	Y	Y	Y	Y	Y	Y
$Industry\ F.E.$	Y		Y		Y		Y	
$Firm\ F.E.$		Y		Y		Y		Y
# Observations	25,252	25,252	25,252	25,252	25,252	25,252	25,252	25,252

Table VII: Adverse Outcome Predictability Omitting Repeated Outcomes

Table VII reports results from predictive regressions of future adverse outcomes on current risk factor disclosures, and controls for ex-ante risk. The sample consists of firm-years where the contemporaneous value of the adverse outcome is zero. Specifications (1)-(3) present average marginal effects from probit regressions with correlated random effects to control for average firm effects. Specification (4) presents coefficients from an OLS regression with firm fixed effects. The $Dependent\ Var.\ (Level)_t$ variable is different for each specification, and is equal to the continuous dependent variable scaled by total assets. Specification (4) does not have a lagged dependent variable by definition of the removal of repeated adverse outcomes. Variables are defined in Appendix A.***, **, and * indicate significance at the 0.01, 0.05, and 0.10 level, respectively.

	Negative NI (1)	Negative Op. Inc. (2)	Sales Decline (3)	Log(# Suits) (4)
$Log(\#\ New\ RF)_t$	0.019***	0.010***	0.012***	0.008**
	(4.27)	(3.41)	(3.18)	(2.29)
$Log(\#\ Dropped\ RF)_t$	-0.016***	-0.010***	-0.010**	0.000
	(3.28)	(2.87)	(2.47)	(0.07)
$Log(\#\ RF)_{t-1}$	0.018	0.022*	0.060***	0.021
	(1.15)	(1.92)	(4.18)	(1.61)
$\Delta Specificity_t$	0.003*	0.001	-0.001	0.004**
	(1.82)	(0.94)	(0.42)	(2.53)
$\Delta Numeric\ Intensity_t$	0.003	0.004	0.009**	-0.001
	(0.48)	(0.89)	(2.05)	(0.29)
$\Delta Log(\#\ of\ Words/RF)_t$	0.058**	0.028*	0.032	0.057**
	(2.29)	(1.65)	(1.45)	(2.56)
$Dependent\ Var.\ (Level)_{t-1}$	-0.361***	-0.220***	0.056***	
	(5.50)	(5.36)	(5.77)	
$Log(Market\ Equity)_t$	0.047***	0.004	-0.007	0.022***
	(4.44)	(0.64)	(0.97)	(3.31)
$Big\ N_t$	0.015	0.006	-0.000	-0.019
	(0.66)	(0.38)	(0.01)	(1.24)
$Book-to-Market_t$	0.115***	0.027***	0.014**	0.001
	(8.84)	(5.44)	(2.47)	(0.15)
$Tangibility_t$	0.107	0.185***	0.059	0.016
	(1.51)	(3.75)	(1.03)	(0.33)
$Leverage_t$	0.011	-0.032	-0.015	0.032
	(0.31)	(1.09)	(0.60)	(1.05)
$Turnover_t$	0.005	0.013***	0.033***	0.043***
	(0.62)	(2.61)	(6.10)	(6.26)
$Beta_{t-1}$	-0.014	-0.010*	-0.008	-0.009
	(1.57)	(1.72)	(1.15)	(1.30)
$Excess\ Ret._{t-1}$	-0.063***	-0.029***	-0.059***	-0.006
	(7.29)	(5.17)	(8.06)	(0.91)
$Ex.\ Ret.\ Std_{t-1}$	0.939**	0.326	0.743**	0.102
	(2.18)	(1.39)	(2.40)	(0.36)
$Ex.\ Ret.\ Skew_{t-1}$	-0.001	0.001	0.000	-0.004**
	(0.37)	(0.82)	(0.10)	(2.06)
Year F.E.	Y	Y	Y	Y
Correlated R.E.	Y	Y	Y	
Firm F.E.				Y
# Observations	13,837	15,334	17,628	20,130
# of Events	1,649	939	2,038	817

Table VIII: Risk Factors and Future Goodwill Impairments

Table VIII reports the average marginal effects from a regression of the presence of a goodwill impairment on risk factor disclosures under the *Strategic Alliance* (*SA*) and *Accounting* topics. The topics are described in Appendix **??**. The sample omits repeated goodwill impairment observations, similar to the filter in Table VII. The $Log(\Delta\#SA)$ and $Log(\Delta\#Acct.)$ coefficients are calculated as the log of one plus the difference in new and removed risk factor disclosures of the respective topics, less the minimum difference across all firms (to maintain a positive value in the logarithm function). The controls included are those used in Table III: *Log(Market Equity)*, *Big N*, *Book-to-Market*, *Tangibility*, *Leverage*, *Share Turnover*, *Beta*, and *Excess Returns*, *Standard Deviation*, and *Skewness*. Variables are defined in Appendix A.***, **, and * indicate significance at the 0.01, 0.05, and 0.10 level, respectively.

	Goodwill Impairment (1)	Goodwill Impairment (2)	Goodwill Impairment (3)
$Log(\#\ New\ RF)_t$	0.025***		
	(6.92)		
$Log(\#\ New\ SA)_t$		0.074**	
		(2.16)	
$Log(\#\ New\ Acct.)_t$		0.031	
		(1.07)	
$Log(\Delta\#SA)$			0.072**
			(2.08)
$Log(\Delta\#Acct.)$			0.036
			(1.17)
$Log(\#\ Dropped\ RF)_t$	−0.024***		
	(6.14)		
$Log(\#\ Dropped\ SA)_t$		−0.050	
		(0.95)	
$Log(\#\ Dropped\ Acct.)_t$		−0.029	
		(0.82)	
$Log(\#\ RF)_{t-1}$	0.108***		
	(7.59)		
$Log(\#\ SA)_{t-1}$		0.014	0.003
		(0.52)	(0.12)
$Log(\#\ Acct.)_{t-1}$		0.026	0.021
		(1.37)	(1.18)
Controls	Y	Y	Y
Year F.E.	Y	Y	Y
Correlated R.E.	Y	Y	Y
# Observations	21,969	21,617	21,617

Table IX: Risk Factors and Future Oil Price

Table IX reports the results of a regression of the change in oil price on risk factor disclosures under the *Energy* topic. The topic is described in Appendix **??**. Δ*Oil Price* is measured as the difference between the monthly price of oil 3 months after the fiscal year end, less the price of oil at the fiscal year end. The controls included are those used in Table III: *Log(Market Equity)*, *Big N*, *Book-to-Market*, *Tangibility*, *Leverage*, *Share Turnover*, *Beta*, and *Excess Returns*, *Standard Deviation*, and *Skewness*. Variables are defined in Appendix A.***, **, and * indicate significance at the 0.01, 0.05, and 0.10 level, respectively.

	Δ*Oil Price*	Δ*Oil Price*	Δ*Oil Price*
	(1)	(2)	(3)
Log(# New RF)$_t$	-0.634^{***}		
	(5.68)		
Log(# New Energy)$_t$		-0.349^{**}	
		(2.18)	
Log(Δ#Energy)			-0.914^{***}
			(2.88)
Log(# Dropped RF)$_t$	0.769^{***}		
	(6.00)		
Log(# Dropped Energy)$_t$		0.425^{**}	
		(2.21)	
Log(# RF)$_{t-1}$	-5.448^{***}		
	(16.52)		
Log(# Energy)$_{t-1}$		-1.701^{***}	-1.690^{***}
		(7.63)	(7.63)
Controls	Y	Y	Y
Firm F.E.	Y	Y	Y
# Observations	25,445	25,044	25,044