International Journal of Accounting Information Systems xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

International Journal of Accounting Information Systems

journal homepage: www.elsevier.com/locate/accinf

SEC's cybersecurity disclosure guidance and disclosed cybersecurity risk factors

He Li[a], Won Gyun No[b,*], Tawei Wang[c]

[a] Southwestern University of Finance and Economics, 555, Liutai Avenue, Wenjiang District, Chengdu, Sichuan 611130, PR China
[b] Rutgers Business School, Rutgers, the State University of New Jersey, 1 Washington Park, Newark, NJ 07102, United States
[c] Driehaus College of Business, DePaul University, 1 E. Jackson Blvd. Chicago, IL 60604, United States

ARTICLE INFO

Keywords:
Cybersecurity
Cybersecurity risk disclosure
Risk factors
Disclosure guidance
Cybersecurity breach incident

ABSTRACT

Cybersecurity risk disclosure has received great attention in the past several years, especially after the passage of the Securities and Exchange Commission's (SEC's) cybersecurity disclosure guidance published on October 13, 2011. In this study, we examine the usefulness of cybersecurity-related risk factors disclosed in 10-K filings. We document that the presence of these risk factors in the pre-guidance period and length of these risk factors are related to future reported cybersecurity incidents. The association between the presence of cybersecurity risk disclosure and subsequently reported cybersecurity incidents becomes insignificant after the passage of the SEC's cybersecurity disclosure guidance. Our findings, in general, support the SEC's decision on emphasizing cybersecurity risk disclosure. However, SEC's disclosure guidance may unintentionally encourage firms to disclose cybersecurity risks regardless of the level of risks.

1. Introduction

Cybersecurity has attracted a lot of attention in the past ten years.[1] Both the general public and the business world are concerned about the growing cybercrimes that expose sensitive personal information, cause business disruptions, or steal trade secrets, especially after a series of high-profile data breaches such as the ones at Equifax, Sony, and Target.[2] According to a recent Annual Cybersecurity Report, > 20% of the breached firms experienced substantial loss of revenues, customer base, and business opportunities, and most of the breached firms spent millions of dollars improving security solutions and expanding security procedures following the attacks (CISCO, 2017). Due to the potential impact on firm value and operations, cybersecurity is becoming one of the top priorities for the board and executives. For instance, about 88% of U.S. Chief Executive Officers (CEOs) are concerned that cyber threats could hinder the growth of their firms (Loop, 2016). Likewise, investors are clamoring for more information about cybersecurity risks and data breaches, and how firms are addressing those risks (Shumsky, 2016).

To respond to the increasing cyber threats, the Securities and Exchange Commission (SEC) held a roundtable discussion to deliberate on cybersecurity landscape and cybersecurity disclosure issues (SEC, 2014). The Standing Advisory Group of the Public Company Accounting Oversight Board (PCAOB) also discussed the potential implications of cybersecurity on financial reporting and auditing (PCAOB, 2014). Particularly, the SEC's Division of Corporation Finance issued disclosure guidance regarding cybersecurity in 2011 to assist firms in

* Corresponding author at: 1 Washington Park, Room 993, Newark, NJ 07102-3122, United States.
E-mail addresses: lihe_stanley@swufe.edu.cn (H. Li), wgno@business.rutgers.edu (W.G. No), david.wang@depaul.edu (T. Wang).

[1] According to the U.S. Computer Emergency Readiness Team, cybersecurity is "[t]he activity or process, ability or capability, or state whereby information and communications systems and the information contained therein are protected from and/or defended against damage, unauthorized use or modification, or exploitation."

[2] For more detail, please see https://www.csoonline.com/article/2130877/data-breach/the-biggest-data-breaches-of-the-21st-century.html.

https://doi.org/10.1016/j.accinf.2018.06.003

1467-0895/ Published by Elsevier Inc.

assessing what, if any, disclosures they should provide related to cybersecurity risks (SEC, 2011). Although the guidance is not technically a ruling, the SEC has issued comment letters to several firms pointing out the inadequacies of their cybersecurity risk disclosures by referring to the guidance. Therefore, some have argued that the guidance is becoming a de facto ruling (Grant and Grant, 2014).

In this paper, we investigate the usefulness of cybersecurity risk disclosures in the risk factor section of 10-K filings (hereafter cybersecurity risk disclosure). We define the usefulness of cybersecurity risk disclosures as "*the ability to help stakeholders assess the possibility of the occurrence of future adverse events* (*i.e., cybersecurity breach incidents*)." Understanding the information conveyed by cybersecurity risk disclosures is important as it can help investors assess a firm's cybersecurity risk and provide regulators with information about whether additional legislative rules are necessary to encourage firms to disclose more on their cybersecurity risks. Two aspects of cybersecurity risk disclosure are considered in this study: presence and length. Specifically, we examine whether the presence of cybersecurity risk disclosure in a firm's 10-K filing implies higher cybersecurity risks as measured by subsequent cybersecurity incidents and whether the length of the disclosure is associated with increased likelihood of subsequent cybersecurity incidents. Consistent with Wang et al. (2013a), our results suggest that both the presence and the length of cybersecurity risk disclosure are associated with subsequent cybersecurity incidents. However, the association between the presence of disclosure and cybersecurity incidents becomes insignificant following the SEC's disclosure guidance. The release of the disclosure guidance leads to a substantial increase in the number of firms disclosing cybersecurity risks, suggesting that firms make cybersecurity disclosures regardless of their degree of cybersecurity risk in the post-guidance period.

The findings of this study make several contributions to the existing literature. First, the study contributes to the cybersecurity disclosure literature. Early research in the accounting and information systems domain primarily focuses on the market reaction following cybersecurity incidents and have examined a set of contingency factors such as type of breaches (Gordon et al., 2011; Yayla and Hu, 2011), firm characteristics (Ettredge and Richardson, 2003), and information disclosed through news articles (Wang et al., 2013b) and distribution channels (Benaroch et al., 2012) that could deepen or mitigate the market reaction, while only a few studies consider cybersecurity disclosure. Gordon et al. (2010) find that on average, voluntary disclosure relating to information security increases stock prices by > 6% and that voluntary disclosure concerning proactive security measures has the greatest impact on the firm's stock price, followed by the disclosure of vulnerabilities. This study complements Gordon et al. (2010) by exclusively focusing on the value of cybersecurity disclosures in terms of predicting future cybersecurity incidents. Wang et al. (2013a) examined the ex-post odds of cybersecurity incidents, revealing that firms that disclose information security risk factors in their 10-K filings with actionable information are less likely to be associated with future cybersecurity incidents. The paper complements Wang et al. (2013a) in at least two key ways. First, with the dramatic change in cybersecurity attacking behaviors in the past decade and the high profile breach incidents in recent years, it is worthwhile revisiting the usefulness of cybersecurity risk factor disclosures. More importantly, the sample in this study covers both the pre-guidance period and the post-guidance period, which enables us to examine the changes in disclosure usefulness. Second, our data collection approach enables analyses on a much larger scale to demonstrate that firms facing greater cybersecurity risks devote a greater portion of their disclosures towards describing cybersecurity risks.

Second, this research also contributes to the risk factor disclosure literature. While findings in the study are largely in line with recent accounting literature showing that risk factor disclosure is not boilerplate, we use the actual adverse event (i.e., cybersecurity incident) rather than market-based measures of firm risks (Campbell et al., 2014) or investors' risk perceptions (Kravet and Muslu, 2013) to capture risks that a firm faces. As the objective of providing risk factor disclosures is to discuss "the most significant factors that make the firm risky" (SEC, 2005), our risk measure that focuses on actual risk event is more consistent with the SEC's intention than measures based on the assumption of market efficiency. Our study, therefore, provides more direct evidence that risk disclosures are potentially informative of future operational failures. Furthermore, different from prior studies examining the variation of qualitative disclosures that are already included in risk factor disclosure section, our unique setting allows us to show that the presence or absence of risk disclosures could be informative of the risk.

Third, this paper makes contributions to the textual analysis literature. When examining disclosures related to cybersecurity, prior studies use manual collection (Wang et al., 2013a), take the number of several words around the keywords (Gordon et al., 2010), or simply count the number of predetermined keywords (Hilary et al., 2017). We develop methods that first locate individual risk factors from Item 1A and then identify security-related risk factors. These methods help us to examine the length of cybersecurity risk disclosure more accurately and are also consistent with recent research effort that calls for analysis at the individual risk factor level (Bao and Datta, 2014; Gaulin, 2017). In addition, the topic analysis using word-term patterns helps to obtain a thorough understanding with respect to the consequences of cybersecurity incidents that firms are most concerned about, which is not examined in prior studies.

Fourth, the results can also help the board of directors, executives, and policymakers to determine the benefits and consequences of cybersecurity risk disclosures and disclosure guidance.[3] Our findings support the decision to emphasize cybersecurity risk disclosures, as both the presence in the pre-guidance period and the length of cybersecurity risk disclosures are informative of subsequent cybersecurity incidents. However, the results also reveal that the SEC's disclosure guidance leads to an unintentional consequence that more firms make cybersecurity risk disclosures even though they do not face higher cybersecurity risks. As the SEC warned firms to "avoid generic risk factor disclosure that could apply to any company," the outcome is counter to the SEC' intention.[4] Such outcome arises from the ambiguity in the guidance and comment letters sent by the SEC to force firms to disclose cybersecurity

[3] Studies such as Kwon et al. (2013), Hsu and Wang (2014a), Hsu and Wang (2014b), Hsu and Wang (2015), and Feng and Wang (2017) have investigated various board/executive issues that may be related to information security risk management. Our study does not attempt to provide implications on the composition or characteristics of the board and executives but to highlight the importance of cybersecurity risk disclosures for the board and executives.

[4] This is consistent with the idea that disclosure is an inexpensive insurance policy. That is, firms will disclose their cybersecurity risks since if something happens, they can point to the disclosure and mitigate lawsuits, at least to some extent. We thank an anonymous reviewer for bringing this point out.

risks (Ferraro, 2013). Therefore, it may be necessary for the SEC to revise the guidance to encourage only firms which are facing a greater level of cybersecurity risks (e.g., experienced a Distributed Denial of Service attack) to make such disclosures. The SEC may not want to elevate the guidance to the commission level, a suggestion made by Senator Jay Rockefeller in 2013, as that may push more firms to issue cybersecurity risk disclosures without having high cybersecurity risks. Additionally, while Ferraro (2013) criticizes that the SEC did little to resolve the concern about publicly revealing too much information could provide potential hackers with a roadmap for successful attacks, we find no evidence supporting such claim.

The remainder of this paper is organized as follows. The next section provides research background and hypothesis development. This is followed by the details of sample selection procedures and research methodology. Next, empirical results and additional analyses are presented. The last section concludes this paper.

2. Background and hypotheses development

2.1. Risk factor disclosure

On June 29, 2005, the SEC mandated firms to describe “the most significant factors that make the offering speculative or risky” in Item 1A of 10-K filings after December 1, 2005 with the objective being “to provide investors with a clear and concise summary of the material risks to an investment in the issuer's securities” (SEC, 2005). Since firms are only required to provide qualitative descriptions and do not need to quantify the likelihood or impact of the disclosed risks, they have a great degree of discretion in what to disclose and how to disclose. Practitioners criticize that managers are likely to provide vague risk disclosure and simply list all uncertainties they face, providing little information for investors (Reuters, 2005). Similarly, Robbins and Rothenberg (2005) argue that risk factor disclosures are the cheapest form of insurance as “firms that cannot point to such a risk factor when faced with a lawsuit will wish they could turn back the clock and insert such language”, implying that firms have incentives to make uninformative risk factor disclosures for legal protection. Realizing the problem, the SEC has issued comment letters to require more risk information from firms (Johnson, 2010) and has warned firms to “avoid risk factor disclosure that could apply to any issuer or any offering” (SEC, 2010).

The concern that risk factor disclosures may be boilerplate is alleviated by recent studies. Campbell et al. (2014) show that firms disclose more risk factors when facing greater risks and devote a greater portion of the disclosures towards describing risks that are more significant. They also find that the unexpected portion of risk factor disclosures is associated with systematic risk, idiosyncratic risk, information asymmetry, and abnormal returns following the disclosure, indicating that market participants perceive the information conveyed by risk factor disclosures. Similarly, Kravet and Muslu (2013) reveal that increases in the number of risk-related sentences are positively associated with stock volatility, trading volume around and after the filings, and dispersed forecast revisions around the filings. However, the effect is largely driven by industry-level risk disclosures rather than firm-level disclosures. Hope et al. (2016) demonstrate that the level of specificity in risk factor disclosures is positively associated with the market reaction to 10-K filings and can help analysts assess firms' fundamental risk. Two contemporary papers examine the effect of comment letters. Brown et al. (2015) identify that firms significantly modify their risk factor disclosures after receiving comment letters. More importantly, spillover effect exists in that firms not receiving comment letters still revise their risk factor disclosures if industry leader, close rival, or industry peers receive comment letters regarding risk factor disclosures, suggesting a deterrence benefit of the SEC's review process. Beatty et al. (2015) find that financial constraints risk factor disclosures are associated with firms' expected level of financial constraints, ex-ante litigation risk, and realized financial constraints outcomes. However, the association is significantly reduced after firms increase disclosures to respond to comment letters, demonstrating the concerns that firms may make disclosures that they otherwise deem immaterial simply to fulfill regulatory requirements.

Several recent studies focus on risk factor updates. Filzen (2015) indicates that firms with risk factor updates in their quarterly reports have lower abnormal returns around the filing dates, lower future unexpected earnings, and larger likelihood of experiencing future negative earnings shock. A subsequent study by Filzen et al. (2016) documents that quarterly risk factor updates are negatively associated with future returns and that the association is stronger for firms using more direct words related to firm fundamentals. Finally, Gaulin (2017) emphasizes the importance of using individual risk factors by showing that managers add new risk factors and remove stale risk factors on a timely basis, and such activities predict future economic changes even after controlling for ex-ante risk and firm performance. In addition, firms respond to the SEC comment letters by improving the level of specificity while they respond to securities litigation by expanding the number of risks they identified without increasing the definitiveness of the disclosures, supporting the litigation shield hypothesis.

2.2. Cybersecurity disclosure guidance

In 2011, the SEC's Division of Corporation Finance issued disclosure guidance related to cybersecurity, pointing out sections that may be relevant for cybersecurity-related disclosure.[5] Regarding risk factor disclosure, the guidance states that “in determining whether risk factor disclosure is required, we expect registrants to evaluate their cybersecurity risks and take into account all

[5] Note that, in February 2018, the SEC issued an interpretation on the “Commission Statement and Guidance on Public Company Cybersecurity Disclosures” that goes beyond the 2011 guidance by addressing the application of disclosure controls and procedures to cybersecurity risks and incidents. However, given the time period this study covers, we are not able to consider this interpretation in our study.

available relevant information, including prior cybersecurity incidents and the severity and frequency of those incidents… Registrants should consider the probability of cybersecurity incidents occurring and the quantitative and qualitative magnitude of those risks, including the potential costs and other consequences resulting from misappropriation of assets or sensitive information, corruption of data or operational disruption" (SEC, 2011). Although the guidance explicitly specifies that it is not a ruling, the SEC has used comment letters to prompt cybersecurity risk disclosures. For example, in the comment letter addressing Freeport-McMoRan Copper & Gold Inc.'s annual report of 2011, the SEC states that: "We note that none of your risk factors, or other sections of your Form 10-K, specifically address any risks you may face from cyber attacks, such as attempts by third parties to gain access to your systems to compromise sensitive business information, to interrupt your systems or otherwise try to cause harm to your business and operations. In future filings, beginning with your next Form 10-Q, please provide risk factor disclosure describing the cybersecurity risks that you face or tell us why you believe such disclosure is unnecessary." Since comment letters are often considered as de facto rulings, one may argue that the disclosure guidance is becoming disclosure requirement (Grant and Grant, 2014).

However, cybersecurity disclosure has been criticized for being uninformative and boilerplate by both practitioners and academics. They argue that firms use boilerplate language every year (Bennett, 2015), a common criticism for risk factor disclosures in Item 1A. To examine the effectiveness of public firm disclosures, a panel, during the roundtable discussion organized by the SEC in 2014, was formed to discuss disclosures concerning cybersecurity risks and cybersecurity incidents by "focusing on what public firms are currently disclosing about their cybersecurity threats and breaches, both potential and those that have already occurred, and how they determine the appropriate disclosure, the timing of that disclosure, and what information about cybersecurity investors need to know to make informed voting and investment decisions."[6] Most panelists raised concerns that many cybersecurity disclosures are boilerplate and admitted the difficulty of striking a balance between providing meaningful disclosure and not adversely affecting the firm's reputation and performance. For instance, Keith Higgins, the director of the Division of Corporation Finance of the SEC, indicated in the panel discussion: "If you take boilerplate on the one hand and on the far side you take a look at the specific road map of the company's vulnerabilities and what the consequences of those vulnerabilities could be, where do you find the balance? How do you – is there somewhere in the middle that will be helpful to investors while at the same time not harmful to companies?" The lack of clarity in the SEC's cybersecurity disclosure guidance further complicates the issue. As the guidance acknowledges, there is no explicit requirement for disclosing cybersecurity risks or cybersecurity incidents so far. The guidance only pointed out several areas where cybersecurity disclosures may be necessary. Accordingly, firms have great discretion in deciding whether, what, and how much to disclose.

Research on the disclosure guidance is recently emerging. Ferraro (2013) argues that the disclosure guidance both procedurally overreaches and substantively underachieves. The author criticizes that the SEC is using the non-legislative guidance as a legislative rule. More importantly, the paper points out that the guidance is vague, similar across industries that will bring little information to the market. Consistent with this view, Hilary et al. (2017) fail to find a significant association between the market reaction following cybersecurity incidents and firms' prior cyber disclosures.

2.3. Hypotheses development

All the hypotheses in this study focus on the usefulness of cybersecurity risk disclosure associated with future cybersecurity incidents. The maintained assumption underlying the hypotheses is that managers are at least partially knowledgeable about the vulnerabilities, cyber threats firms face, and the security measures they have taken.

The first hypothesis centers on the presence of cybersecurity risk disclosures. The disclosure literature suggests that managers have incentives to withhold negative information (Beyer et al., 2010; Verrecchia, 2001) because disclosing bad news could reduce market value, increase cost of capital, damage future career opportunities, and reveal proprietary information to competitors (Ke et al., 2003; Kothari et al., 2009a; Kothari et al., 2009b). However, managers also have incentives to disclose negative information to reduce potential litigation cost and reputational damage (Skinner, 1994). In the context of cybersecurity risk disclosure, if a firm facing high cybersecurity risk fails to alert the investors about the risk in advance and the risk materialized to an actual cybersecurity incident, the firm may be exposed to lawsuits. For example, Heartland Payment Systems was sued for "misrepresenting or failing to disclose that the company's safety and security measures designed to protect consumers' financial records and data from security breaches were inadequate and ineffective."[7] Therefore, managers are likely to disclose cybersecurity risk if they believe that the likelihood of future cybersecurity incident is high, and the potential implication of the incident is significant. By contrast, managers are less likely to disclose cybersecurity risk if they foresee that the probability of future cybersecurity incident is low, and the impact of the incident is insignificant. When the potential litigation and reputational damage is remote, managers are not willing to bear the cost of disclosing negative information.

Taken together, we argue that firms that provide cybersecurity risk disclosures face higher cybersecurity risk, and thus are more likely to experience cybersecurity incidents. Accordingly, we introduce the following hypothesis.

H1. The presence of cybersecurity risk disclosure is positively associated with the likelihood of subsequently reported cybersecurity incident.

The second hypothesis examines the length of cybersecurity risk disclosure. While the presence of cybersecurity risk disclosure

[6] For more details, visit https://www.sec.gov/spotlight/cybersecurity-roundtable/cybersecurity-roundtable-transcript.txt.
[7] For more details, visit http://securities.stanford.edu/filings-case.html?id=104260.

Graco Inc.

Security Breaches – Intrusion into our information systems may impact our business.

Security breaches or intrusion into our information systems, and the breakdown, interruption in or inadequate upgrading or maintenance of our information processing software, hardware or networks may impact our business. Security breaches or intrusion into the systems or data of the third parties with whom we conduct business may also harm our business.

(Excerpted from https://www.sec.gov/Archives/edgar/data/42888/000119312514056452/d675621d10k.htm.)

Diodes Inc.

System security risks, data protection breaches, cyber-attacks and other related cybersecurity issues could disrupt our internal operations, and any such disruption could reduce our expected revenue, increase our expenses, damage our reputation and adversely affect our stock price.

Experienced computer programmers and hackers may be able to penetrate our security controls and misappropriate or compromise our confidential information or that of third parties, create system disruptions or cause shutdowns. Computer programmers and hackers also may be able to develop and deploy viruses, worms and other malicious software programs that attack our websites, products or otherwise exploit any security vulnerabilities of our websites and products. The costs to us to eliminate or alleviate cyber or other security

problems, bugs, viruses, worms, malicious software programs and security vulnerabilities could be significant, and our efforts to address these problems may not be successful and could result in interruptions, delays, cessation of service and loss of existing or potential customers that may impede our sales, manufacturing, distribution or other critical functions.

We manage and store various proprietary information and sensitive or confidential data relating to our business and third party business. Breaches of our security measures or the accidental loss, inadvertent disclosure or unapproved dissemination of proprietary information or sensitive or confidential data about us or our partners or customers, including the potential loss or disclosure of such information or data as a result of fraud, trickery or other forms of deception, could expose us, our partners and customers or the individuals affected to a risk of loss or misuse of this information, result in litigation and potential liability for us, damage our brand and reputation or otherwise harm our business. In addition, the cost and operational consequences of implementing further data protection measures could be significant.

Delayed sales, significant costs or lost customers resulting from these system security risks, data protection breaches, cyber-attacks and other related cybersecurity issues could adversely affect our financial results, stock price and reputation.

(Excerpted from https://www.sec.gov/Archives/edgar/data/29002/000119312514073365/d633786d10k.htm.)

Fig. 1. Examples of Cybersecurity Risk Disclosure.

signals elevated cybersecurity risk that prompts the firm to disclose, the variation of the disclosure length could also be informative. Fig. 1 shows two cybersecurity risk disclosures.

It may be misleading to treat these two cybersecurity risk disclosures the same as they differ significantly in the amount of information provided. Practitioners, regulators, and academics have expressed concerns that cybersecurity risk disclosures may be boilerplate (Bennett, 2015; Hilary et al., 2017). If the concern is true, the length of cybersecurity risk disclosure is not likely to be associated with the likelihood of reported future cybersecurity incidents. On the other hand, Campbell et al. (2014) show that the level of risk determines the volume of disclosure firms devote to address that risk. Similarly, Filzen (2015) argues that the more discussions of potential negative outcomes, the greater the likelihood of the negative event. If cybersecurity risk disclosure is informative, firms facing higher cybersecurity risks are more likely to devote a greater portion of the disclosures to describe their cybersecurity risks. Therefore, it is an interesting empirical question whether the length of cybersecurity risk disclosure, as measured by a relative length to capture the relative importance of the risk in firm's risk portfolio, is useful. This leads to the following hypothesis.

H2. The length of cybersecurity risk disclosure is positively associated with the likelihood of subsequently reported cybersecurity incident.

The last hypothesis explores the effect of the SEC's cybersecurity disclosure guidance. Firms are increasingly disclosing their cybersecurity risks following the guidance. The percentage of firms providing cybersecurity risk disclosure jumps from 27.29% in

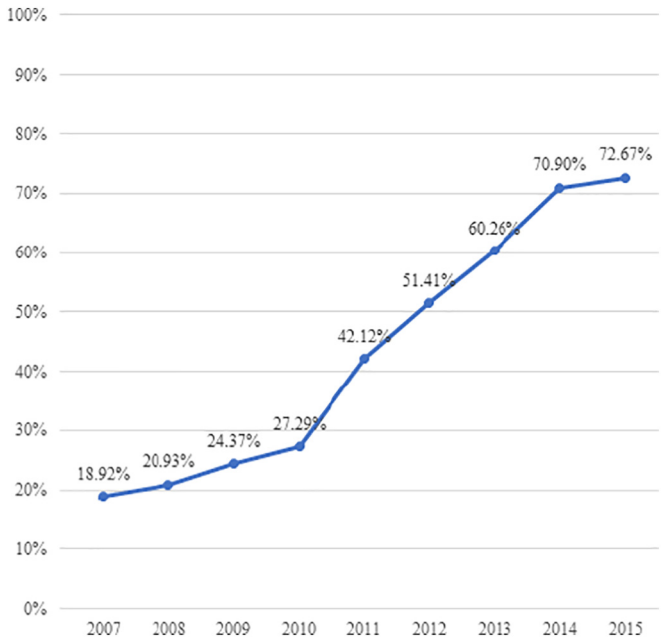

Fig. 2. Percentage of Cybersecurity Risk Disclosures Across Years.

2010 to 42.12% in 2011 (see Fig. 2). However, it is unclear whether the increase in regulatory pressure will result in uninformative disclosures. Since risk factor disclosure in Item 1A is qualitative and does not require assessment of probability, firms may disclose all possible risk factors to fulfill regulatory requirements (Campbell et al., 2014). Consistent with this view, Beatty et al. (2015) document that disclosures become less reflective of future financial constraints following the SEC comment letters. To the extent that the SEC's cybersecurity disclosure guidance could be viewed as a regulatory shock, this study examines the following hypothesis.

H3. The association between the presence of cybersecurity risk disclosure and subsequent cybersecurity incident is different before and after the introduction of the SEC's cybersecurity disclosure guidance.

3. Empirical design and sample selection

3.1. Empirical design

The first hypothesis predicts that the presence of cybersecurity risk disclosure is associated with subsequent cybersecurity incidents. We construct a variable, *Disclosure*, that takes one if there is any cybersecurity risk disclosure in that fiscal year, zero otherwise. To examine the second hypothesis, we create another variable, *Length*, which measures the total word count of cybersecurity risk disclosure, normalized by the average word count of individual risk factors disclosed in Item 1A for that firm-year. The normalization is important as it controls for a firm's tendency to provide longer disclosure. A logit model is estimated with *Breach* as the dependent variable that takes one if the firm experiences cybersecurity incident in year $t + 1$, zero otherwise.

$$\begin{aligned} P(Breach_{it} + 1 = 1) = {} & Cyber\,Dislosure_{it} + Past\,Breach_{it} \\ & + Size_{it} + LN\,Segments_{it} \\ & + Age_{it} + Loss_{it} + LN\,Analyst_{it} + Foreign_{it} + Merger_{it} \\ & + Growth_{it} + ICW_{it} \end{aligned} \quad (1)$$

Appendix A provides a detailed definition of each variable. *Cyber_Disclosure* is the variable of interest, be it either *Disclosure* or *Length*. A positive coefficient on this variable would support the hypotheses. A set of control variables based on prior literature are also included (Hilary et al., 2017; Sheneman, 2017; Wang et al., 2013a). Specifically, we control for consumer industry as it is the sector that witnesses most cybersecurity incidents. Positive coefficients are expected for firm size, age, growth, and the number of analysts following, as these variables control for the visibility of the firm. Further, firm's financial conditions are controlled using *Loss*. As financially constrained firms are less likely to invest sufficiently into their financial reporting control systems (Doyle et al., 2007), it is expected that firms with losses are also less likely to make sufficient investment in their internal controls over operations. In addition, we include *Foreign*, *Merger*, and *LN_Segments* to control the complexity of a firm's business. Positive coefficients are expected on these variables as more complex and dispersed operations are likely to result in ineffective and inconsistent controls (Sheneman, 2017). We also include *ICW* to control for a firm's internal control environment. Since internal controls over financial reporting and internal controls over operations are correlated (Lawrence et al., 2016), firms with material weaknesses in internal controls over

financial reporting are more likely to experience cybersecurity incidents. Finally, we add an indicator variable, *Past_Breach*, to capture whether the firm had cybersecurity incidents in any previous year (i.e., going back to the first year of the sample, 2007).

3.2. Sample selection

The cybersecurity incident data comes from Privacy Rights Clearinghouse (privacyrights.org) and Audit Analytics cybersecurity database. Privacy Rights Clearinghouse publishes data breaches that involve individual's identity while the Audit Analytics cybersecurity database collects hacking incidents. To identify cybersecurity risk disclosures, we first extract Item 1A from each 10-K using an approach similar to Campbell et al. (2014).[8] After obtaining the whole Item 1A section, we obtain individual risk factors using information provided in HTML tags. The SEC requires that each risk factor should be preceded by a subcaption that summarizes that risk.[9] Similar to Gaulin (2017), we identify each subcaption that is highlighted (i.e., bold, italic, or underlined) and is located at the beginning of a paragraph or isolated on a separate line. The content between two highlighted subcaptions is considered to be a unique risk factor. A detailed description of the procedure can be found in Appendix B.

We then use a keyword search to identify risk factors related to cybersecurity. Keywords are identified from prior research (Gordon et al., 2010; Wang et al., 2013a) and have been refined to prevent misidentification.[10] Risk factors that contain any of these keywords are considered cybersecurity risk disclosure. To ensure the quality of the identification, we randomly selected 200 risk factor disclosures and manually validated the identified risk factors related to cybersecurity. All of them are accurately identified. Appendix C provides a list of these keywords.

Table 1 summarizes the sample selection procedure. We start with 755 cybersecurity incidents for the period between 2007 and 2015 that can be mapped to Compustat.[11] We exclude observations in the software and computer industries (SIC between 3570 and 3579 and 7370–7379) because their cybersecurity risk disclosures may not be accurately captured by the keyword search.[12] For firms that experienced more than one cybersecurity incidents in the same year, we keep only one incident in our sample. Also, observations of which Item 1A cannot be extracted are removed from our sample.[13] Lastly, we eliminate observations that have missing values on any of the independent variables. In total, the final sample contains 26,335 non-breached observations and 291 breached observations.

4. Results

4.1. Descriptive statistics

Fig. 2 shows the percentage of firms providing cybersecurity risk disclosures in the sample. While the overall trend is upward, there is an unusual jump following the SEC's cybersecurity disclosure guidance in 2011. In addition, the annual increase in the percentage of firms providing cybersecurity risk disclosures is much larger following the disclosure guidance.

Table 2 presents the descriptive statistics of the variables used in this study. The mean for *Breach* is 0.011, suggesting that only about 1% of the firms in the sample experience cybersecurity incidents. The percentages of firms making cybersecurity risk disclosures are 73.5% (with cybersecurity incidents) and 36.7% (without cybersecurity incidents), providing initial support for our argument that firms with high cybersecurity risks are more likely to provide cybersecurity risk disclosures. Similarly, cybersecurity risk disclosures of firms with cybersecurity incidents are much longer than that of firms without cybersecurity incidents (2.134 vs. 1.548).

Table 3 reports the univariate correlations among variables examined in this study. The variables of interest, *Disclosure* and *Length*, are both positively correlated with the dependent variable *Breach*.[14]

4.2. Main findings

Panel A of Table 4 shows the results of testing H1. Consistent with our expectation, the coefficient of *Disclosure* is positive and significant (0.695, $p < 0.01$). The result suggests that firms with prior cybersecurity risk disclosures are more likely to experience subsequent cybersecurity incidents. As for control variables, larger firms, firms undergoing merger, firms operating in consumer section, and firms with a history of cybersecurity incidents are more likely to have future cybersecurity incidents. Panel B of Table 4 presents the test results for H2. The coefficient of *Length* is positive and significant (0.181, $p < 0.01$), revealing that firms providing lengthy cybersecurity risk disclosure are more likely to experience subsequent cybersecurity incidents. In untabulated test, we also

[8] All 10-K filings filed between January 2005 and December 2015 are downloaded.

[9] Item 503(c) of Regulation S-K.

[10] We exclude several keywords that could generate false positives. For example, while Trojan typically refers to malicious program that is used to hack into a computer, it can also refer to a condom brand. In addition, we add some new keywords such as ransomware and key logger.

[11] Many incidents reported in the database occur in non-profit or private firms, and thus cannot be linked to Compustat.

[12] Business of firms in these industries could include providing security solutions to customers. The keyword search method is likely to misidentify these security solutions as risk factors related to their business. For instance, disclosure regarding how the sales of intrusion detection products would influence stock price is incorrectly identified as risk factor, which has nothing to do with cybersecurity risk.

[13] We are not able to extract all Item 1As due to HTML tagging issues in some of the 10-K filings.

[14] Multicollinearity is not a concern as none of the correlation is larger than 0.5 and that Variance Inflation Factors (VIFs) for our models are < 2.5.

Table 1
Sample Selection.

Number of firm-years with cybersecurity incidents	291
Original number of cybersecurity incidents	755
Minus: observations that are in the computer and software industry (SIC 3570–3579, 7370–7379)	(94)
Minus: observations that have more than one cybersecurity incidents in a year (keep each firm-year only once)	(107)
Minus: observations for which Item 1A cannot be extracted	(184)
Minus: observations that have missing values on any one of the variables used in the study	(79)
Number of firm-years without cybersecurity incidents	26,335
Total number of firm-years	26,626

Table 2
Descriptive Statistics for Variables in Eq. (1).

Variable	Total sample (N = 26,626)			Firms without cybersecurity incidents (N = 26,335)			Firms with cybersecurity incidents (N = 291)		
	Mean	Standard deviation	Median	Mean	Standard deviation	Median	Mean	Standard deviation	Median
Breach	0.011	0.104	0.000	0.000	0.000	0.000	1.000	0.000	1.000
Past_breach	0.037	0.189	0.000	0.033	0.178	0.000	0.430	0.496	0.000
Disclosure	0.371	0.483	0.000	0.367	0.482	0.000	0.735	0.442	1.000
Length	1.560	1.211	1.153	1.548	1.198	1.147	2.134	1.583	1.552
Size	6.458	2.360	6.619	6.426	2.346	6.591	9.326	1.769	9.314
LN_segments	1.449	0.490	1.386	1.447	0.490	1.386	1.590	0.472	1.609
Age	22.824	16.331	18.000	22.733	16.293	18.000	31.031	17.583	28.000
Loss	0.424	0.494	0.000	0.426	0.495	0.000	0.196	0.398	0.000
LN_analyst	0.293	0.764	0.000	0.293	0.763	0.000	0.275	0.841	0.000
Foreign	0.260	0.439	0.000	0.260	0.439	0.000	0.241	0.428	0.000
Merger	0.195	0.396	0.000	0.193	0.395	0.000	0.316	0.466	0.000
Growth	0.153	0.642	0.049	0.153	0.645	0.050	0.078	0.238	0.043
ICW	0.085	0.278	0.000	0.085	0.279	0.000	0.031	0.173	0.000

Note: All variables are winsorized at 1 and 99%. All variables are defined in Appendix A.

Table 3
Correlations among Variables in Eq. (1).

	Breach	Past_breach	Disclosure	Length	Size	LN_segments	Age	Loss	LN_analyst	Foreign	Merger	Growth	ICW
Breach	1.000												
Past_breach	**0.218**	1.000											
Disclosure	**0.079**	**0.172**	1.000	.									
Length	**0.070**	**0.143**	.	1.000									
Size	**0.128**	**0.239**	**0.258**	0.018	1.000								
LN_segments	**0.030**	**0.051**	**0.059**	−0.018	**0.309**	1.000							
Age	**0.053**	**0.112**	**0.104**	**−0.048**	**0.346**	**0.266**	1.000						
Loss	**−0.049**	**−0.085**	**−0.132**	**−0.042**	**−0.408**	**−0.151**	**−0.239**	1.000					
LN_analyst	−0.003	−0.007	**−0.034**	0.018	**0.039**	**0.012**	**−0.104**	**0.049**	1.000				
Foreign	−0.005	−0.008	0.008	**−0.029**	**0.107**	**0.352**	**0.060**	**−0.026**	**0.036**	1.000			
Merger	**0.032**	**0.053**	**0.131**	**0.056**	**0.175**	**0.177**	**0.044**	**−0.078**	**0.019**	**0.117**	1.000		
Growth	**−0.012**	**−0.031**	**−0.050**	−0.014	**−0.108**	**−0.118**	**−0.145**	**0.077**	**0.012**	**−0.034**	**0.024**	1.000	
ICW	**−0.020**	**−0.040**	**−0.069**	−0.006	**−0.254**	**−0.087**	**−0.131**	**0.180**	**−0.018**	**−0.021**	**−0.033**	**0.073**	1.000

Note: All variables are defined in Appendix A. Bold indicates significance at 0.05 level.

use alternative measures of *Length*. Specifically, *Length* is replaced with the log number of words in cybersecurity risk disclosure as well as the number of words in cybersecurity risk disclosure normalized by the total number of words in Item 1A. Similar results are obtained using both measures. Overall, results reported in Table 4 suggest that both the presence and length of cybersecurity risk disclosures as measured by adjusted length are informative of future cybersecurity incidents, providing support for the SEC's intention to encourage cybersecurity risk disclosures.

To examine H3, we partition the sample into a pre-guidance period group and a post-guidance period group and reexamine Eq. (1). Results are presented in Table 5. Panel A and Panel B of Table 5 reveal that *Disclosure* is only significant in the pre-guidance period group, but not significant in the post-guidance period group. We also create a dummy variable, *Guidance*, that takes the value of 1 if the observation is in the post-guidance period, 0 otherwise. The interaction between *Guidance* and *Disclosure* is negative and

Table 4
Logit Regression of Cybersecurity Risk Disclosure on Cybersecurity Incidents.

Independent variables	Panel A		Panel B	
	Estimates	z-statistics	Estimates	z-statistics
Disclosure	0.695	3.17***		
Length			0.181	3.69***
Past_breach	1.454	8.11***	1.317	7.21***
Size	0.621	10.68***	0.568	9.01***
LN_segments	0.112	0.57	0.173	0.68
Age	−0.004	−0.80	−0.005	−1.00
Loss	−0.024	−0.15	0.043	0.21
LN_analyst	0.050	0.50	0.062	0.52
Foreign	−0.069	−0.42	−0.086	−0.45
Merger	0.246	1.72**	0.097	0.61
Growth	−0.044	−0.27	−0.028	−0.13
ICW	0.051	0.14	0.061	0.13
Consumer	0.966	2.56***	0.958	1.91**
Intercept	−10.230	−19.54***	−9.522	−14.90***
Year effects	Included		Included	
Industry effects	Included		Included	
Pseudo R square	0.276		0.238	
# Observations	26,626		9884	

Note: *, **, *** represent significance at the 0.10, 0.05, and 0.01 levels based on two-tailed *p*-values (one-tailed when predicted), respectively. Test statistics are based on robust standard errors clustered by firm. All variables are defined in Appendix A.

significant (Panel C), demonstrating that there is a differential effect of cybersecurity risk disclosures in the pre- and post-guidance period. Overall, these findings support the argument that the introduction of the SEC's cybersecurity disclosure guidance leads to disclosures by firms that do not have material cybersecurity risks.

4.3. Sensitivity tests

As robustness tests, we explore alternative selections for the control firms. Following Wang et al. (2013a), we consider two methods to select 1:1 control groups: a randomly sampled control group and a size-matched group. Results using the randomly sampled control group and using the size-matched group are presented in Table 6 and Table 7, respectively. Findings in these two tables are consistent with our main results.

Table 5
Logit Regression of Cybersecurity Risk Disclosure on Cybersecurity Incidents by Period.

Independent variables	Panel A Pre-guidance		Panel B Post-guidance		Panel C Interaction	
	Estimates	z-statistics	Estimates	z-statistics	Estimates	z-statistics
Disclosure	0.892	3.51***	0.356	1.19	0.872	3.63***
Guidance					0.163	0.70
Disclosure*guidance					−0.508	−1.74**
Past_breach	1.566	4.99***	1.389	6.42***	1.435	7.99***
Size	0.580	6.04***	0.647	8.75***	0.625	10.79***
LN_segments	0.232	0.83	−0.053	−0.23	0.121	0.62
Age	0.001	0.08	−0.005	−0.85	−0.003	−0.77
Loss	−0.113	−0.42	0.044	0.19	−0.083	−0.51
LN_analyst	0.188	1.52*	−0.169	−0.96	0.069	0.67
Foreign	−0.072	−0.27	−0.020	−0.10	−0.075	−0.46
Merger	0.273	0.96	0.312	1.80**	0.232	1.61*
Growth	−0.119	−0.51	0.031	0.15	−0.006	−0.04
ICW	0.068	0.10	−0.031	−0.06	0.069	0.19
Consumer	0.670	1.09	1.145	1.85**	0.969	2.59***
Intercept	−11.453	−14.77***	−10.037	−14.07***	−10.240	−19.64***
Year effects	Included		Included		Included	
Industry effects	Included		Included		Included	
Pseudo R square	0.267		0.276		0.271	
# Observations	11,849		14,777		26,626	

Note: *, **, *** represent significance at the 0.10, 0.05, and 0.01 levels based on two-tailed *p*-values (one-tailed when predicted), respectively. Test statistics are based on robust standard errors clustered by firm. All variables are defined in Appendix A.

Table 6
Sensitivity Tests Using Random Non-Breach Sample.

Independent variables	Panel A		Panel B		Panel C	
	Estimates	z-statistics	Estimates	z-statistics	Estimates	z-statistics
Disclosure	1.202	3.99***			1.867	3.97***
Length			0.429	2.54***		
Guidance					0.552	1.52*
Disclosure*guidance					−1.448	−2.47***
Past_breach	2.486	3.68***	2.345	2.87***	2.518	3.72***
Size	0.669	6.34***	0.672	5.14***	0.656	6.56***
LN_segments	−0.246	−0.85	0.170	0.38	−0.258	−0.90
Age	−0.001	−0.17	−0.013	−1.11	0.001	0.15
Loss	−0.237	−0.77	−0.063	−0.13	−0.300	−1.01
LN_analyst	0.325	2.33**	0.465	1.98**	0.342	2.45***
Foreign	0.028	0.09	0.212	0.46	0.044	0.15
Merger	0.041	0.13	−0.425	−0.86	−0.041	−0.14
Growth	−0.540	−1.69**	−0.827	−1.68**	−0.507	−1.51*
ICW	0.434	0.75	1.350	1.27	0.427	0.79
Consumer	1.112	1.70**	1.634	1.45*	0.880	1.33*
Intercept	−6.001	−6.92***	−2.500	−1.95**	−5.964	−7.24***
Year effects	Included		Included		Included	
Industry effects	Included		Included		Included	
Pseudo R square	0.636		0.603		0.633	
# Observations	582		315		582	

Note: *, **, *** represent significance at the 0.10, 0.05, and 0.01 levels based on two-tailed *p*-values (one-tailed when predicted), respectively. Test statistics are based on robust standard errors clustered by firm. All variables are defined in Appendix A.

4.4. Additional tests

4.4.1. Firm-specific disclosure

In this section, we address the concern that more firm-specific cybersecurity risk disclosures could lead to more attacks. The SEC stated that "we are mindful of potential concerns that detailed disclosures could compromise cybersecurity efforts – for example, by providing a 'roadmap' for those who seek to infiltrate a registrant's network security – and we emphasize that disclosures of that nature are not required under the federal securities laws" (SEC, 2011). Ferraro (2013) criticizes the SEC's failure to address this issue and argues that any disclosure that is meaningful for investors is likely to contain information for hackers seeking future attacks. To test if the claim is valid, we use eq. (1) but substitute *Cyber_Disclosure* with two measures: *Score* and *Uniqueness*. Both measures are

Table 7
Sensitivity Tests Using Size-matched Sample.

Independent variables	Panel A		Panel B		Panel C	
	Estimates	z-statistics	Estimates	z-statistics	Estimates	z-statistics
Disclosure	1.519	3.15***			2.662	3.70***
Length			0.859	2.25**		
Guidance					0.341	0.62
Disclosure*guidance					−2.033	−2.23**
Past_breach	4.425	5.27***	4.776	2.70***	4.333	5.44***
Size	0.352	2.13**	0.916	2.97***	0.345	2.11***
LN_segments	0.468	0.81	0.859	1.00	0.774	1.33*
Age	−0.040	−2.41***	−0.089	−3.08***	−0.042	−2.57***
Loss	−0.508	−0.89	−0.627	−0.91	−0.450	−0.76
LN_analyst	11.373	8.04***	5.162	12.10***	11.645	8.86***
Foreign	−0.483	−1.03	−0.376	−0.52	−0.502	−1.10
Merger	1.394	2.36***	1.242	1.54*	1.150	1.86**
Growth	2.063	1.63*	−1.916	−2.33**	1.807	1.54*
ICW	0.059	0.06	−0.388	−0.36	−0.292	−0.24
Consumer	2.217	2.35***	2.979	2.41***	1.691	1.69**
Intercept	−4.092	−2.83***	−4.487	−1.55*	−4.338	−2.96***
Year effects	Included		Included		Included	
Industry effects	Included		Included		Included	
Pseudo R square	0.625		0.758		0.626	
# Observations	582		344		582	

Note: *, **, *** represent significance at the 0.10, 0.05, and 0.01 levels based on two-tailed *p*-values (one-tailed when predicted), respectively. Test statistics are based on robust standard errors clustered by firm. All variables are defined in Appendix A.

Table 8
Logit Regression of Firm-specific Disclosure on Cybersecurity Incidents.

Independent variables	Panel A		Panel B	
	Estimates	z-statistics	Estimates	z-statistics
Score	−1.037	−1.01		
Uniqueness			0.495	0.45
Past_breach	1.367	7.47***	1.389	7.44***
Size	0.568	8.57***	0.558	8.11***
LN_segments	0.160	0.63	0.100	0.39
Age	−0.005	−0.95	−0.003	−0.61
Loss	0.080	0.41	0.104	0.52
LN_analyst	0.060	0.49	0.047	0.36
Foreign	−0.018	−0.09	−0.011	−0.05
Merger	0.131	0.81	0.082	0.49
Growth	−0.041	−0.20	−0.012	−0.06
ICW	0.019	0.04	−0.160	−0.32
Consumer	1.007	2.08**	0.933	1.95**
Intercept	−9.385	−14.55***	−9.233	−12.54***
Year effects	Included		Included	
Industry effects	Included		Included	
Pseudo R square	0.228		0.229	
# Observations	9686		8900	

Note: *, **, *** represent significance at the 0.10, 0.05, and 0.01 levels based on two-tailed *p*-values (one-tailed when predicted), respectively. Test statistics are based on robust standard errors clustered by firm. All variables are defined in Appendix A.

based on the bag-of-words approach that represents documents as vectors with each dimension representing a unique word. The first measure, *Score*, is adapted from Brown and Tucker (2011), which is calculated as one minus the cosine similarity between a firm's disclosure and industry-year's average disclosure, adjusted by document length using Tayler expansion.[15] The variable captures how a firm's disclosure deviates from the average industry practice. The second measure, *Uniqueness*, is calculated as the percentage of unique words that are not used by any other firms in the same industry for the same fiscal year. The variable represents how a firm's disclosure includes firm-specific information regarding word usage. Table 8 shows the regression results. Neither of these two measures is statistically significant. While the results do not invalidate the concern, they seem to suggest that the level of firm-specific information in Item 1A of 10-K is not informative enough to jeopardize a firm's cybersecurity endeavor.

4.4.2. Topic analysis

To further understand cybersecurity risk disclosure, a topic analysis is conducted to investigate firm's concerns about cybersecurity. Specifically, all two-word phrases that occur in at least 2% but no > 98% of all cybersecurity disclosures are extracted, which gives us 1042 phrases in total.[16] We manually read these phrases, choose 211 phrases that are meaningful, and classify these phrases into five topics of consequences: business operations, financial performance, reputation, lawsuit and litigation, and intellectual property. Appendix A lists the phrases used for classification.

Fig. 3 shows the percentage of firms that mention each type of risk across years. The figure offers two important findings. First, the disruption of business operations is the biggest concern regarding cybersecurity. Among the firms that disclose cybersecurity risks in Item 1A, > 85% of them report the potential impact of cybersecurity incidents on business operations, and the rate remains relatively stable over time. Impact on financial performance is the second biggest concern, with > 70% of the firms mentioning this topic. Second, while intellectual property is the least mentioned topic, we observe a significant jump in the recent years. Similarly, concerns over reputation are steadily increasing over years, which is consistent with the public perception that cybersecurity is attracting greater attention in recent years.

4.4.3. Market reactions to incidents following disclosures

It is ex-ante not clear whether the market incorporates information conveyed by the disclosure that describes cybersecurity risk. We explore whether the market reacts differently with and without the presence of cybersecurity risk disclosures as in Wang et al. (2013a). Our un-tabulated findings demonstrate that the market reaction following cybersecurity incidents is less severe for firms with prior cybersecurity risk disclosures as in Wang et al. (2013a). However, this is not the case for lengthy disclosures. Future research can explore more on this direction.

[15] Brown and Tucker (2011) analytically prove that the similarity score between two documents is a function of document length. Accordingly, they propose to use Tayler expansion to adjust the similarity score.

[16] All the words are stemmed, and stop words are removed. Phrases that consist of two words are used to increase the interpretability of the outcome. In addition, we use 2% as the threshold to get rid of specific phrases such as firm names as well as 98% threshold to filter out uninformative phrases that are used by all disclosures. The results do not change when these parameters are varied.

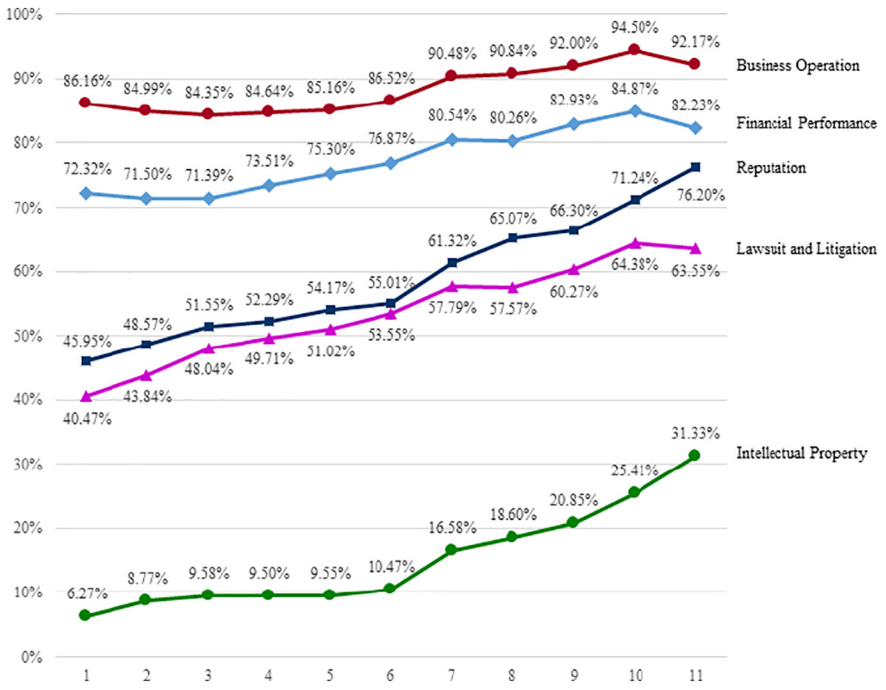

Fig. 3. Percentage of Firms Disclosing Different Topics Across Years.

5. Concluding remarks

In this paper, we examine whether cybersecurity risk disclosure is informative for future cybersecurity incidents. Specifically, we focus on two measures: the presence of cybersecurity risk disclosure and the length of cybersecurity risk disclosure. Consistent with our expectation, the presence in the pre-guidance period and length of cybersecurity risk disclosure are positively associated with subsequent cybersecurity incidents, suggesting that cybersecurity risk disclosure is not boilerplate. Furthermore, the presence of cybersecurity risk disclosure is no longer associated with subsequent cybersecurity incidents after the SEC's cybersecurity disclosure guidance, revealing that the SEC's emphasis on cybersecurity risk disclosures results in more disclosures by firms not having material cybersecurity risks. We fail to find a significant association between firm-specific disclosure and cybersecurity incidents, providing some relief for the regulator's concern that more firm-specific disclosure may provide information for hackers. Finally, the topic analysis indicates that firms are more concerned about the disruption of business operations and impact on financial performance when encountering cybersecurity issues. Moreover, there is a growing concern regarding reputation damage and loss of intellectual property due to cybersecurity incidents. Collectively, results in this paper should be valuable to practitioners, regulators, and academics who are interested in the usefulness of cybersecurity risk disclosures. We stand with the SEC to emphasize the importance of cybersecurity risk disclosure but raise the question of the unintended consequence resulted from cybersecurity disclosure guidance.

There are several limitations. We maintain the assumption that managers have knowledge of the cybersecurity risks firms face, which may not necessarily hold. If firms are not aware of the level of cyber threats, they are less likely to provide meaningful disclosures. In addition, cybersecurity incidents are used as the proxy for cybersecurity risks, which may not be the most accurate measure as theoretically any system can be breached. Firms that did not experience cybersecurity incidents may have good controls in place to prevent such incident or may just be lucky.

There are several possible future research directions. First, future studies may benefit by using information at more disaggregated level, such as data from intrusion detection system (IDS), to determine whether firm's disclosure practice is consistent with the actual cyber threats it faces. Alternatively, one can interview managers to better understand firms' risk assessments and disclosure decisions and whether the disclosed information reflects the firms' cybersecurity risk management concerns. In addition, disclosure itself can be explored further to understand how firms determine whether a specific piece of information should be disclosed and how it should be disclosed. Second, another avenue for future research is to examine whether the current disclosed content is informative enough to help investor's decision making as well as whether there is a change in the content of cybersecurity disclosures before and after the SEC guidance. The AICPA is proposing new disclosure guidelines regarding cybersecurity. Different disclosure activities may provide a more holistic understanding of a firm's cybersecurity risk management program. Third, it would be interesting to explore different ways to capture the characteristics of incidents and investigate how it may affect disclosure decisions.

Acknowledgements

The authors are grateful for the valuable comments provided by the discussants and participants at University of Waterloo Centre on Information Integrity and Information Assurance 10th Biennial Symposium. The authors are also thankful for the financial supports from China Scholarship Council, Southwestern University of Finance and Economics, Rutgers Business School, and DePaul University.

Appendix A. Variable definitions

Variable	Definition
Breach	Indicator variable, equal to 1 if the firm experiences cybersecurity incident(s) during fiscal year t, 0 otherwise;
Past_breach	Indicator variable, equal to 1 if the firm experiences cybersecurity incident(s) in any year preceding fiscal year t, 0 otherwise;
Disclosure	Indicator variable, equal to 1 if the firm has cybersecurity risk disclosure in fiscal year t, 0 otherwise;
Length	Total number of words in cybersecurity risk disclosure in fiscal year t, normalized by the average number of words in individual risk factor disclosed in Item 1A;
Size	Natural log of total assets in millions in fiscal year t;
LN_segments	Natural log of number of business and geographic segments in fiscal year t;
Age	Number of year firms are included in CompuSmart in fiscal year t;
Loss	Indicator variable, equal to 1 if the firm reports negative net income in fiscal year t, 0 otherwise;
LN_analyst	Natural log of number of analysts following in fiscal year t;
Foreign	Indicator variable, equal to 1 if the firm has foreign operations (based on FCA) in fiscal year t, 0 otherwise;
Merger	Indicator variable, equal to 1 if the firm is involved in merger activity in fiscal year t (based on AQP), 0 otherwise;
Growth	One-year growth rate in sales in fiscal year t;
ICW	Indicator variable, equal to 1 if the auditor reports an internal control weakness in fiscal year t, 0 otherwise;
Consumer	Indicator variable, equal to 1 if the firm operates in consumer goods industry (i.e., SIC between 5200 and 5999);
Guidance	Indicator variable, equal to 1 after the issuance of SEC cybersecurity disclosure guidance, 0 otherwise;
Score	One minus the cosine similarity score between firm's cybersecurity risk disclosure and industry's average disclosure for fiscal year t, adjusted by length using Taylor expansion proposed by Brown and Tucker (2011)
Uniqueness	Percentage of unique words that are not used by any other firms in the same industry for the same fiscal year

Appendix B. Risk factor extraction

We first download all available 10-K filings filed between December 2005 to December 2015 from the SEC's Electronic Data Gathering, Analysis, and Retrieval (EDGAR) system because the SEC mandated risk factor disclosure on December 2005. Similar to Campbell et al. (2014) and Gaulin (2017), our procedure for extracting risk factor disclosure (i.e., Item 1A) is based on the assumption that 10-K filings in HyperText Markup Language (HTML) format contain visual clues (e.g., emphasis or whitespace separation) for readers to recognize item boundaries easily. The HTML filings are parsed into a tree structure using Beautifulsoup package in Python (https://pypi.python.org/pypi/beautifulsoup4). The leaf nodes of the tree are textual information while the internal nodes of the tree are HTML tags that can be used for identifying headings. For example, a tag, ⟨p⟩, defines a paragraph that is visually separated and isolated from text below and above. By assuming items are presented in order, all the HTML tags that contain the text “ITEM 1A”, “ITEM 1B”, “ITEM 2” (case insensitive) are identified. From all the candidates, the ones that are emphasized are first selected (i.e., the ones include tag ‘b’, ‘em’, ‘strong’, ‘h1’, ‘h2’, ‘h3’, ‘h4’, ‘h5’, ‘h6’, ‘u’, ‘p’, ‘font’, ‘div’, ‘span’, or ‘li’ if using HTML emphasis tags, or ‘bold’, ‘italic’, or ‘underline’ if using Cascading Style Sheets within HTML tags). For all the candidates that satisfy the emphasis criteria, we identify their first parent node that is one of the following: ‘h1’, ‘h2’, ‘h3’, ‘h4’, ‘h5’, ‘h6’, ‘p’, ‘div’, ‘ul’, ‘ol’, and ‘table’. For the ones that are not separated by ‘table,’ we obtain the plain text in the separated paragraph which contains the phrase “ITEM 1A RISK FACTOR” without any other words. For the ones that are separated by ‘table,’ we gather the entire row and obtain the plain text in the entire row which contains the phrase “ITEM 1A RISK FACTOR” without any other words.

Following the procedure, a list of elements that contain the headers for Item 1A, Item 1B or Item 2 is located. Risk factor disclosures are identified by extracting all the contents between the first Item 1A header and the first Item 1B or Item 2 header (in case there is no Item 1B). Individual risk factors are also extracted using HTML tags, similar to the approach used in Gaulin (2017). The SEC requires that each risk factor should be preceded by a subcaption that summarizes the risk. We identify these subcaptions based on such requirement: i.e., they are emphasized (bold, underline, or italic), and are at the beginning of each paragraph or isolated on its own line. The identified subcaptions are further filtered by applying a threshold (i.e., there are at least 10 words below that subcaptions). Contents between subcaptions represent individual risk factors.

Appendix C. Keywords and phrases

Keywords to identify cybersecurity risk disclosures

encryption
computer (virus|breach|break-in|attack|security)
security (breach|incident)
(information|network|computer) security
intrusion
hacking|hacker
denial of service
cyber(−|)(attack|fraud|threat|risk|terrorist|incident|security)
cyber-based attack
cybersecurity
infosec
system security
information technology (security|attack)
data theft
phishing
malware
data confidentiality
confidentiality of data
confidential data
unauthorized access
data corruption
corruption of data
network break-in
espionage
cyber(−|)insurance
data breach
crimeware
ransomware
keylogger
keystroke logging
social engineering

Phrases to Identify Topics (Stemmed)

Lawsuit and litigation	'addit-regulatori', 'applic-law', 'civil-crimin', 'civil-litig', 'compli-applic', 'compli-law', 'complianc-cost', 'contractu-oblig', 'crimin-penalti', 'enforc-action', 'expo-civil', 'expo-litig', 'fail-compli', 'failur-compli', 'feder-state', 'fine-penalti', 'govern-regul', 'law-govern', 'law-protect', 'law-regul', 'legal-claim', 'legal-liabil', 'legisl-regulatori', 'liabil-claim', 'liabil-law', 'litig-liabil', 'litig-regulatori', 'loss-litig', 'possibl-liabil', 'potenti-liabil', 'privaci-law', 'regulatori-action', 'regulatori-approv', 'regulatori-environ', 'regulatori-interv', 'regulatori-penalti', 'regulatori-requir', 'regulatori-scrutini', 'result-legal', 'result-litig', 'secur-law', 'signific-legal', 'state-feder', 'state-law', 'state-local', 'subject-litig', 'violat-applic'
Business operations	'abil-conduct', 'abil-oper', 'abil-perform', 'act-vandal', 'affect-oper', 'busi-continu', 'busi-damag', 'busi-disrupt', 'busi-failur', 'busi-harm', 'busi-interrupt', 'caus-disrupt', 'caus-interrupt', 'compromis-network', 'compromis-secur', 'comput-equip', 'comput-hardwar', 'comput-network', 'comput-telecommun', 'conduct-busi', 'continu-oper', 'continu-plan', 'creat-disrupt', 'critic-busi', 'damag-disrupt', 'damag-failur', 'damag-interrupt', 'deliv-product', 'denial-servic', 'disast-power', 'disast-recoveri', 'disast-terror', 'disast-terrorist', 'disrupt-busi', 'disrupt-compani', 'disrupt-inform', 'disrupt-oper', 'disrupt-servic', 'disrupt-shutdown', 'effect-oper', 'electr-telecommun', 'enterpri-resourc', 'experi-interrupt', 'failur-disrupt', 'failur-interrupt', 'failur-network', 'hardwar-failur', 'harm-oper', 'impact-oper', 'infrastructur-vulner', 'intern-control', 'intern-oper', 'internet-telecommun', 'interrupt-busi', 'interrupt-failur', 'interrupt-malfunct', 'interrupt-oper', 'interrupt-power', 'interrupt-servic', 'jeopard-secur', 'loss-telecommun', 'malfunct-oper', 'materi-disrupt', 'network-disrupt', 'network-failur', 'network-infrastructur', 'oper-disrupt', 'oper-failur', 'oper-infrastructur', 'oper-interrupt', 'penetr-network', 'power-loss', 'power-outag', 'properti-damag', 'resourc-plan', 'result-disrupt', 'result-interrupt', 'servic-attack', 'servic-disrupt', 'servic-interrupt', 'signific-disrupt', 'signific-interrupt', 'similar-disrupt', 'softwar-hardwar', 'softwar-network', 'subject-disrupt', 'suppli-chain', 'technolog-disrupt',

	'technolog-fail', 'technolog-failur', 'technolog-infrastructur', 'technolog-network', 'telecommun-failur', 'telecommun-outag', 'transmiss-distribut', 'uninterrupt-oper'
Reputation	'abil-attract', 'affect-reput', 'attract-new', 'attract-retain', 'busi-reput', 'compani-reput', 'custom-relationship', 'damag-brand', 'damag-reput', 'effect-reput', 'harm-reput', 'impact-reput', 'negat-public', 'relationship-custom', 'relationship-manag', 'reput-brand', 'reput-damag', 'reput-expo', 'reput-financi', 'reput-harm', 'reput-loss', 'reput-suffer'
Intellectual property	'competit-posit', 'intellectu-properti', 'proprietari-busi', 'research-develop', 'trade-secret'
Financial performance	'addit-cost', 'addit-resourc', 'affect-financi', 'capac-constraint', 'capit-expenditur', 'capit-resourc', 'cash-flow', 'common-stock', 'compen-loss', 'decreas-revenu', 'effect-financi', 'financi-condit', 'financi-liabil', 'financi-loss', 'financi-oper', 'financi-perform', 'financi-posit', 'financi-result', 'impact-financi', 'increas-cost', 'increas-expen', 'incur-liabil', 'loss-liabil', 'loss-revenu', 'lost-revenu', 'oper-cash', 'oper-cost', 'oper-expen', 'oper-financi', 'proceed-liabil', 'reduc-revenu', 'remedi-cost', 'revenu-profit', 'signific-capit', 'signific-cost', 'signific-expen', 'signific-invest', 'signific-liabil', 'signific-loss', 'substanti-cost', 'suffer-loss'

References

Bao, Y., Datta, A., 2014. Simultaneously discovering and quantifying risk types from textual risk disclosures. Manag. Sci. 60 (6), 1371–1391.
Beatty, A., Cheng, L., Zhang, H., 2015. Sometimes Less is More: Evidence from Financial Constraints Risk Factor Disclosures. (Working Paper).
Benaroch, M., Chernobai, A., Goldstein, J., 2012. An internal control perspective on the market value consequences of IT operational risk events. Int. J. Account. Inf. Syst. 13 (4), 357–381.
Bennett, C., 2015. SEC Weighs Cybersecurity Disclosure Rules. The Hill.
Beyer, A., Cohen, D.A., Lys, T.Z., Walther, B.R., 2010. The financial reporting environment: review of the recent literature. J. Account. Econ. 50 (2–3), 296–343.
Brown, S.V., Tucker, J.W., 2011. Large-sample evidence on firms' year-over-year MD&A modifications. J. Account. Res. 49 (2), 309–346.
Brown, S.V., Tian, X.S., Tucker, J.W., 2015. The Spillover Effect of SEC Comment Letters on Qualitative Corporate Disclosure: Evidence from the Risk Factor Disclosure. (Working Paper).
Campbell, J.L., Chen, H., Dhaliwal, D.S., Lu, H.-m., Steele, L.B., 2014. The information content of mandatory risk factor disclosures in corporate filings. Rev. Acc. Stud. 19 (1), 396–455.
CISCO, 2017. Annual Cybersecurity Report.
Doyle, J., Ge, W., Mcvay, S., 2007. Determinants of weaknesses in internal control over financial reporting. J. Account. Econ. 44 (1), 193–223.
Ettredge, M.L., Richardson, V.J., 2003. Information transfer among internet firms: the case of hacker attacks. J. Inf. Syst. 17 (2), 71–82.
Feng, C., Wang, T., 2017. CIO risk appetite and information security management. In: Pre-ICIS Workshop on Accounting Information Systems. Seoul, Korea.
Ferraro, M.F., 2013. Groundbreaking'or broken? In: An Analysis of SEC Cyber-security Disclosure Guidance, its Effectiveness, and Implications.
Filzen, J.J., 2015. The information content of risk factor disclosures in quarterly reports. Account. Horiz. 29 (4), 887–916.
Filzen, J.J., McBrayer, G., Shannon, K., 2016. Risk Factor Disclosures: do Managers and Markets Speak the Same Language? (Working Paper).
Gaulin, M., 2017. Risk Fact or Fiction: the Information Content of Risk Factor Disclosures. (Working Paper).
Gordon, L.A., Loeb, M.P., Sohail, T., 2010. Market value of voluntary disclosures concerning information security. MIS Q. 34 (3), 567–594.
Gordon, L.A., Loeb, M.P., Zhou, L., 2011. The impact of information security breaches: has there been a downward shift in costs? J. Comput. Secur. 19 (1), 33–56.
Grant, G.H., Grant, C.T., 2014. SEC cybersecurity disclosure guidance is quickly becoming a requirement. CPA J. 84 (5), 69.
Hilary, G., Segal, B., Zhang, M.H., 2017. Cyber-risk Disclosure: who Cares? (Working Paper).
Hope, O.-K., Hu, D., Lu, H., 2016. The benefits of specific risk-factor disclosures. Rev. Acc. Stud. 21 (4), 1005–1045.
Hsu, C., Wang, T., 2014a. Composition of the top management team and information security breaches. In: Cruz-Cunha, M.M. (Ed.), Handbook of Research on Digital Crime, Cyberspace Security, and Information Assurance. IGI Global, Pennsylvania.
Hsu, C., Wang, T., 2014b. Exploring the association between board structure and information security breaches. Asia-Pac. J. Inf. Sys. 24 (4), 531–557.
Hsu, C., Wang, T., 2015. Board busyness and information security risk management. In: Pre-ICIS Workshop on Accounting Information Systems. Fort Worth, TX.
Johnson, S., 2010. SEC Pushes Companies for More Risk Information. (CFO Magazine).
Ke, B., Huddart, S., Petroni, K., 2003. What insiders know about future earnings and how they use it: evidence from insider trades. J. Account. Econ. 35 (3), 315–346.
Kothari, S.P., Li, X., Short, J.E., 2009a. The effect of disclosures by management, analysts, and business press on cost of capital, return volatility, and analyst forecasts: a study using content analysis. J. Account. Econ. 84 (5), 1639–1670.
Kothari, S.P., Shu, S., Wysocki, P.D., 2009b. Do managers withhold bad news? J. Account. Res. 47 (1), 241–276.
Kravet, T., Muslu, V., 2013. Textual risk disclosures and investors' risk perceptions. Rev. Acc. Stud. 18 (4), 1088–1122.
Kwon, J., Rees, J., Wang, T., 2013. The association between top management involvement and compensation and information security breaches. J. Inf. Syst. 27 (1), 219–236.
Lawrence, A., Minutti-Meza, M., Vyas, D., 2016. Is Operational Control Risk Informative of Undetected Financial Reporting Deficiencies? (Working Paper).
Loop, P., 2016. Cybersecurity and the board: 8 issues keeping directors up at night. Wall Street J. https://sponsoredcontent.wsj.com/pwc/broader-perspectives/cybersecurity-and-the-board-8-issues-keeping-directors-up-at-night/.
Public Company Accounting Oversight Board, 2014. Standing Advisory Group Meeting: Cybersecurity. Retrieved from. http://pcaobus.org/News/Events/Documents/0624252014_SAG_Meeting/06252014_Cybersecurity.pdf.
Reuters, 2005. Refco Risks Boiler-plate Disclosure. (By Scott Malone).
Robbins, R.B., Rothenberg, P.L., 2005. Writing effective risk factor disclosure in offering documents and exchange act reports. In: Insights: the Corporate & Securities Law Advisor. 19(5).
Securities and Exchange Commission (SEC), 2005. Release #33-8591: Securities Offering Reform (Section VII: Additional Exchange act Disclosure Provisions).
Securities and Exchange Commission (SEC), 2010. 17 CFR PARTS 211, 231 and 241. Release Nos. 33-9106; 34-61469; FR-82.
Securities and Exchange Commission (SEC), 2011. CF Disclosure Guidance: Topic no. 2: Cybersecurity.
Securities and Exchange Commission (SEC), 2014. Cybersecurity Roundtable. Retrieved from. https://www.sec.gov/spotlight/cybersecurity-roundtable.shtml.
Sheneman, A.G., 2017. The Effect of Operating Control Failures on the Cost of Capital-evidence from Data Breaches. (Working Paper).

Shumsky, T., 2016. Corporate judgment call: when to disclose you've been hacked. Wall Street J. https://www.wsj.com/articles/corporate-judgment-call-when-to-disclose-youve-been-hacked-1474320689.
Skinner, D.J., 1994. Why firms voluntarily disclose bad news. J. Account. Res. 32 (1), 38–60.
Verrecchia, R.E., 2001. Essays on disclosure. J. Account. Econ. 32 (1), 97–180.
Wang, T., Kannan, K., Rees, J., 2013a. The association between the disclosure and the realization of information security risk factors. Inf. Syst. Res. 24 (2), 201–218.
Wang, T., Rees, J., Kannan, K., 2013b. Investors' reactions to information security incidents and profitable short-term investment opportunities. J. Organ. Comput. Electron. Commer. 23, 1–24.
Yayla, A.A., Hu, Q., 2011. The impact of information security events on the stock value of firms: the effect of contingency factors. J. Inf. Technol. 26 (1), 60–77.