The Real Peer Effects of Data Breaches: Evidence from IT investment*

Wentao Yao†

November 2020

Abstract: This paper studies whether peer firms respond to data breach events occurred to other firms in their industry. Exploiting malicious hackings targeting U.S. public firms from 2008 to 2018, I find that peer firms defined as the same four-digit SIC industry level significantly increase IT investment following a hacking in their industry. The effect is stronger when more data records are lost and for firms that are more exposed to cyberattack, have higher risk management awareness (denoted by having a designated risk officer or a risk committee, or having previously discussed cyber risk in annual reports), and for the period after the 2011 SEC guidance on more cyber risk disclosure. I find that board interlocking with breach firms is a channel for the relevant information (e.g., on the effects of hacking and the resulting corporate responses) to be transmitted from hacked firms to peer firms. Overall, peer firms respond to hackings in the same industry by taking precautionary actions to manage the potential cyberattack risk and/or the potential litigation and reputation risk. In this way, a negative event (i.e., a data breach) results in positive externalities to the industry and its stakeholders.

Keywords: Cyber risk, Data breach, Peer effect, IT, Corporate disclosure, Risk management.
JEL codes: G30, G31, G32

* I'm deeply grateful to Hong Zou for his mentoring and suggestions. I thank Christopher Armstrong, Sumit Agarwal, Chen Lin, Thomas Schmid, Jennifer Tucker, Mao Ye, Wei Zhu, and doctoral students of the University of Hong Kong and of the University of Illinois at Urbana-Champaign for their helpful comments. Part of the research was conducted when I visited the University of Illinois at Urbana-Champaign. I'm also grateful to Clara Xiaoling Chen and University of Illinois at Urbana-Champaign for hosting me. All errors remain mine.
† Finance area, HKU Business school, The University of Hong Kong. Email:yaowt029@connect.hku.hk.

The Real Peer Effects of Data Breaches: Evidence from IT investment

"Any failure to maintain the security of the information relating to our company, customers, associates and vendors that we hold … could damage our reputation…, could cause us to incur substantial additional costs and to become subject to litigation, and could materially adversely affect our operating results … could result in the release to the public of confidential information about our operations and financial condition and performance … Moreover, a security breach could require us to devote significant management resources to address the problems created by the security breach and to expend significant additional resources to upgrade further the security measures."

(Extracted from the Data and Privacy Risks Section that was newly added in Walmart's 2014 annual report as a response to Target data breach in 2013)

1. Introduction

In today's internet-based world, voluminous personal and business information is stored online, and it is increasingly challenging to secure such information that can be lost due to insiders' errors or outside malicious attacks. Between 2007 and 2018, both the number of data breaches (Figure 1) and the total number of data records lost (Figure 2) have experienced a steady upsurge. On July 19, 2019, Capital One announced a data breach that affected about 100 million credit card customers – roughly 30% of the US population, but this data breach does not even break into the top-five worst data breaches (Youn, 2019). The largest data breach in history is Yahoo's loss of information for 3 billion user accounts in 2013, and for another 500 million user accounts in 2014. PricewaterhouseCoopers (2014) estimates that data breaches cost 1~3% of U.S. GDP annually. Data breach can result in significant remediation costs, potential litigation cost, regulatory fines, and/or reputation damages as pointed out by Walmart in the opening quote and shown by cases reported in Appendix A. For example, Equifax has spent more than a cost of $1.7 billion for the its data breach in 2017. The U.S. Securities and Exchange Commission (SEC) issued its first

guidance to public companies on disclosure of cybersecurity in October 2011, and released an updated interpretative guidance urging companies to be more transparent and timely in disclosing cybersecurity risks in February 2018. These regulatory efforts are motivated by the SEC's belief that the cybersecurity issue could be a material operational risk that can result in significant costs so that investors need to know.

Against this background, in this paper I examine whether peer firms respond to data breaches in their industry by increasing real economic activities proxied by IT investments.[1] Peer firms may do so to assure investors and other stakeholders who may have updated their Bayesian posterior assessment of the data breach risk of such firms after observing a data breach in another firm in the same industry and to lower the potential litigation risk and reputation risk should the same attack occur to them and cause a significant loss later. For example, as shown in the opening quote, in response to supermarket Target's serious data breach in 2013, Walmart added a new section "Data and Privacy Risks" in its 2014 annual report.[2] Securities and industry regulators also require a firm's board of directors to be ultimately responsible for the firm's risk management including the management of cybersecurity risk. The former SEC Commissioner Luis Aguilar spoke at a conference on "Cyber Risks and the Boardroom" hosted by the NYSE that "boards are responsible for overseeing how management implements cyber security programs….and directors [should be put] on notice to proactively address the risks associated with cyber-attacks" (Aguilar, 2014). A perceived failure may result in consumer lawsuits, regulatory sanction, and/or shareholder lawsuits. In the case of consumer lawsuits,

[1] Consistent with the Equifax data breach case reported in Appendix A, I find that firms that suffered a data breach on average increases their IT investment by about 19% in the three-year window around the data breach (see Appendix Table A2).
[2] See an additional example on Transunion's response to Equifax's data breach in 2017 in Appendix B.

Target has already paid over $35 million to settle class actions brought by consumers for its 2013 data breach. In the Yahoo's data breach, the SEC alleged that Yahoo's executives knowingly failed to act on the data breach and hided it from investors (see Appendix A).

Shareholder lawsuits targeting cybersecurity include derivative suits (brought by shareholders to challenge the failure of a firm's board to properly manage the cybersecurity risk and safeguard its valuable data) and securities class actions (that target misrepresentation or omission of cybersecurity risk in corporate filings) (e.g., lawsuit against Yahoo). Although shareholders need to overcome the business judgment rule that protects the board business decisions and satisfy the demand requirement that sets a higher bar in derivative suits (Lin et al., 2020), time and efforts needed in responding to such lawsuits and the potential reputation damage could constitute a concern to a firm's board of directors and managers (Masulis et al., 2020). Shareholder suits have targeted companies such as Heartland Payment Systems, TJX, Target, and Wyndham Worldwide Corp, Home Depot and named their directors and officers as defendants in the aftermath of data breaches (Messing and Zucke, 2015).

On the other hand, as discussed above, shareholder face a higher bar to win at least in derivative suits that target the board's negligence in managing cybersecurity risk (Messing and Zucke, 2015).[3] In addition, some company boards and managers may tend to believe that data breach events are largely isolated incidents and happen by chance to individual firms because the majority of the cyberattacks are non-targeted wherein hackers use malware to try to penetrate all computers they can find on the web and attacks against small firms (that have limited data) also account for about half of all attacks (Nicols, 2016;

[3] Shareholder derivative suits against Taget's board of directors and executives were rejected by the board and the defendants 'motion to dismiss was granted by the court (see Appendix C).

Kaiser, 2018).[4] If so, peer firms may not feel the credible threat of data breach or the pressure from investors and other stakeholders, and this could result in no significant peer responses to data breaches in their industry. Therefore, whether peer firms respond to data breaches is an empirical question.

Some authors (e.g., Ettredge and Richardson, 2003; Akey et al., 2020) argue that the majority of cyberattacks are sudden, unexpected, and idiosyncratic. Our statistics also show that data breach is common and occurs to each 1-digit SIC industry. The timing of such breaches is "plausibly random, and except in rare cases, the breaches themselves do not specifically affect the quality of the products or services offered by the affected company" (Akey et al., 2020, p.4). Kamiya et al. (2020) also argue that the risk of cyberattack is generally uncorrelated with other risks of a firm. Since I do not examine data breach firms per se but their peer firms, it is reasonable to assume that the occurrence of these unexpected data breach events is plausibly exogenous to peer firms especially when the major determinants of cyberattack are controlled for. The low or zero correlation between data breach risk and other firm risks (e.g., financial) also makes any observed peer effect less likely to be confounded by common factors that may independently drive the peer response (Manski, 1993).[5] This setting is thus cleaner than other settings like earnings management or financial restatements that are more likely to be affected by the same unobserved economic shocks such as an industry downturn.

We gather data breach events from Privacy Rights Clearinghouse (PRC), which is a common data source of studying cyber risk (Akey et al., 2020; Kamiya et al., 2020). I

[4] Nichols, Angela, May 12, 2016. Targeted attack vs. Untargeted attack: Knowing the difference. https://www.anomali.com/blog/targeted-attack-vs-untargeted-attack-knowing-the-difference. Needless to say, which system eventually gets penetrated also depends on whether the system is strong enough or not.
[5] Nevertheless, I will also take care to further address the potential reflection problem in our research design.

only include 129 hacking events that are categorized as malicious cyberattack from outsiders and also result in a loss of customer and/or employee records over the period 2007-2018 into our analysis, as such hacking events are more damaging and salient to peer firms.[6] Our IT investment data are from the Computer Intelligence Technology Database (CiTDB) (see Section 2 for detail). In our DID analysis, I define treatment firms as peer firms that share the same four-digit SIC industry as a hacked firm, and all other firms that have never encountered a hacking event over the sample period as control firms. I also use alternative control firms such as firms that share the same three-digit SIC code as the hacked firm, but have a different four-digit SIC code and have never encountered a hacking event over our sample period or matched control firms with a different four-digit SIC code and have never encountered a hacking event over our sample period.

We start our analysis by conducting a short-term market reaction of peer stocks to the 129 hacking events, and find that peers firms on average experience a 0.5% loss to shareholder wealth within the 11-day window centered at the public disclosure date of a hacking event. This is consistent with Kamiya et al.'s (2020) finding that short-term negative effects could be contagious to industry competitors. This negative market reaction may be a reason leading to peer firms' reaction to hacking events in the same industry.

In our baseline DID analysis, I find that compared to firms in industries that have never encountered a hacking event, peer firms of hacked firms on average increase their IT investment by about 20% following a hacking event in their industry. Controlling for high dimensional headquarter state-by-year fixed effects to capture unobserved time-varying headquarter state-specific shocks and two-digit SIC industry-by-year fixed effects to

[6] I hereafter use 'hacking events' and 'data breaches' interchangeably unless stated otherwise.

capture time-varying industry-specific unobserved shocks, the magnitude of IT investment drops to about an 10% increase, but the change remains statistically significant. These results suggest that our observed peer responses are unlikely to be due to unobserved common factors that may independently drive the peer response (a sort of the 'reflection problem') (Manski, 1993). Importantly, the divergence in IT investment between peer (treatment) firms and control firms do not exist before the occurrence of a hacking event, suggesting that the parallel assumption underlying a causal interpretation is met.

The results also hold for several robustness tests, including using an alternative propensity score matched sample, only the first hacking event in each industry, an alternative definition of control groups, or a subsample excluding financial crisis years 2008-2009. Importantly, I also conduct a placebo test showing that peer firms do *not* respond to other seven types of non-salient data breaches other than hacking, supporting our argument that peer firms only respond to severe cybersecurity risk caused by malicious hackings. The IT investment is not limited to a certain component but extends to components including communications, hardware, software, and IT services.

If peer firms indeed pay attention to hacking events in their industries and are concerned about the cybersecurity risk, I expect to observe heterogeneities in firms' responses. I first show that increase in IT investment is predictably larger for hacking events that involve more losses of data records and for firms that are more exposed to cyberattack as proxied by the significant positive determinants of being hacked reported in Kamiya et al. (2020) (e.g., a larger firm size, higher valuation, higher ROA, or less financial constraints) or for firms with more publicity (as proxied by advertising intensity). Second, I find that increase in IT investment is more pronounced in firms that have higher risk

management awareness denoted by having a designated risk officer, a risk committee, or by discussing cyber risk in corporate filings before a hacking.

As a further way to mitigate the concern that our observed baseline result is an artifact of the reflection problem (Manski, 1993), I explicitly show an information transmission channel from the hacked firm to some peer firms. I hypothesize that the relevant information (e.g., on the effects of hacking and the resulting corporate responses) to be transmitted from hacked firms to peer firms via interlocked directors, leading peer firms to be responsive. Consistent with this hypothesis, I show that peer firms experience a larger increase in IT investment when there are interlocked directors between peer firms and a hacked firm.

Finally, I investigate whether the SEC guidance on cyber security disclosure in 2011 causes a structural change in peer effects in IT investment. This guidance has led to a rapid increase in cybersecurity disclosures in firms' 10-K reports (Berkman et al., 2018), though several studies argue that the guidance is largely a boilerplate (Hilary et al., 2016). Our results show that the response of peer firms' IT investment is only significant after the 2011 SEC guidance, therefore the guidance does play an important role. Relatedly, I also show that peer firms increase their frequency of cyber risk discussion in 10-K filings by about 6% following the occurrence of a hacking event in their respective industry. This also helps explain and corroborate our finding of an increase in IT investment.

Overall, peer firms respond to hackings in the same industry by taking precautionary actions to manage potential cyberattack risk, litigation and/or reputation risk. In this way, a negative event (i.e., a data breach) may result in positive externalities to the industry and its stakeholders.

The remainder of this paper proceeds as follows. Section 2 discusses the relation with the literature. Section 3 discusses data and sample used as well as the identification strategies. Section 4 presents the empirical results. I conclude in Section 5.

2. Relation with the literature

Our paper first extends the studies of cyber risk, an emerging and increasingly important operational risk in the risk management literature by documenting the real effects of malicious hackings on peer firms' IT investment. In contrast, prior studies have focused on the determinants and consequences of data breach from the breached firm's point of view. Kamiya et al. (2020) find that firms with higher visibility, higher valuations, less financially constraints, or higher asset intangibility are more likely to be the target of a hacking. They also document a significant loss to shareholders of hacked firms (short-run stock returns and sales growth) arising from hackings involving the loss of personal information. In addition, they find that firms after being hacked decrease risk-taking incentives as afforded by option grants to executives. Huang and Wang (2020) document that banks will price in the data breaches and require higher loan spreads for the firms that just get data breached. The loans also demand more collateral and covenants. Akey et al. (2020) show that breached firms invest more in CSR to restore reputation loss after being hacked. Makridis and Dean (2018) report that breach is negatively related to firms' productivity. Nordlund (2019) finds that firms experience directors' turnover after a data breach. These papers show that data breach has significant consequences for breached firms.

On the other hand, Hilary et al. (2016) report that breached firms do not experience significant changes in operational performance, executive departure likelihood, shareholder clientele, or disclosures except for a modest increase in the proportion of firms disclosing cybersecurity issues after the issuance of the 2011 SEC disclosure guidance. In summary, there is mixed evidence on the effects of data breach, but more importantly, except Kamiya et al.'s (2020) examination of the short-term stock reactions to breached firms' peers, no study has examined whether and how peer firms respond to a data breach event in their respective industry by changing their IT investment.

Second, our paper adds to the growing literature on contagion effects within industries. Gleason et al. (2008) show that non-restating peer firms also suffer a decline in stock price upon restatement announcements in the same industry, suggesting that investors are concerned about the quality of reporting in peer firms and need to reassess the financial statement information previously released by non-restating firms. Gande and Lewis (2009) report a significantly negative stock price reaction to firms targeted by shareholder litigation and their industry peers. Kedia et al. (2015) find that peer firms are more likely to begin managing earnings after the restatement announcement of another firm in their industry or neighborhood if the restating firm is not disciplined by the SEC or class action lawsuits. Files and Gurun (2018) document that restatements in an industry increase peer firms' loan spread regardless of the restatement severity. Beatty et al. (2013) examine how high-profile accounting frauds affect peer firms' investment and show that industry peers react to the overstated earnings of fraudulent firms by increasing investment during fraud periods particularly in industries with higher private benefits of control. Dimmock et al.

(2018) show that fraud is contagious among coworkers in the financial advisor industry. All these studies point to a negative externality of a negative incident.

There is also evidence on the positive externality of a positive event. Lin et al. (2018) find that an increase in institutional ownership in new constituent firms of the Russell 2000 index creates pressures on these firms' industry peers to increase voluntary disclosures. Our study is different from these studies by showing a positive externality of a negative event: peer firms take the occurrence of a data breach event as a warning and making precautionary IT investment to manage the potential cyberattack risk, litigation and/or reputation risk. In this way, a negative data breach event engenders positive externalities by improving peer firms' risk management, thereby lowering the expected costs of data breaches to the economy. Recall that annual data breach costs 1~3% of U.S. GDP (PricewaterhouseCoopers, 2014). The existence of peer effects in IT investment may help lower this loss figure. Our study thus helps enrich the contagion literature. I know of only one study in the literature that also shows a positive externality of a negative event: Cheng et al. (2019) find firms are less likely to report a material internal control weakness if one of their audit committee members is concurrently on the board of a firm that previously disclosed a material weakness.[7]

[7] Another literature examines whether a firm's financial policy is shaped by industry average (see Armstrong et al. (2019) and Grennan (2019) for a review). While loosely related, it is different from our focus in this study – i.e., whether peer firms respond to a negative event occurred to another firm.

3. Data, sample and research design

3.1 IT data

We obtain IT investment data from the Computer Intelligence Technology Database (CiTDB), a proprietary database that provides IT annual spending data for companies around the globe. The database contains detailed plant-level IT information and has been used in several studies on IT technology spending decisions and is believed to be a reliable source for IT spending (Brynjolfsson and Hitt, 2003; Bloom et al., 2014; Tuzel and Zhang, 2017).

The variable I are mostly interested in is "Total IT spending", which is from the IT spending section of the CiTDB database and the value equals the sum of four IT components: Communication spending, hardware spending, software spending, and services spending. Since the variables are recorded at the plant/site level, I first match company names of the plant/site to Compustat companies as well as their subsidiary names using the LexisNexis Affiliation Database. I then aggregate the plant-level variables into Compustat firm-year level data. In the end, I are able to match the IT data to 7,893 Compustat firms, approximately 60% of the firms in the Compustat population over our sample period from 2007 to 2018, which is the longest sample period for the IT data.

3.2 Data breach events

We gather data breach events from the Privacy Rights Clearinghouse (PRC).[8] PRC is an independent nonprofit organization aiming to protect individuals' privacy and

[8] https://www.privacyrights.org/

advocate for positive changes within the ever-changing privacy landscape in the U.S. PRC not only assists consumers with privacy issues, but also collects a comprehensive list of data breach events from public information source since 2005. I download 6,962 data breach events as of December 2018. Figure 1 presents the annual number of data breach events from 2005 to 2018. Figure 2 presents the volume of records lost in data breach over the years. These figures show that both the number of events and the number of customer/employee record loss have been increasing steadily in the last decade. Figure 3 shows the geographic distribution of data breaches. As can be seen, California, New York State, and Texas take the most hit.

The data include detailed information such as data breach organization name and type, public disclosure date, breach type, information source, etc. For the data breach organization types, I exclude "Educational Institutions" (EDU), "Government & Military (GOV)", and "Nonprofits (NGO)".[9] PRC categorizes data breach events into eight types, including "Payment card fraud (CARD)", "Malicious electronic entry by an outside party (HACK)", "Insider taking advantage of company system (INSD)", "Physical loss, stolen, or discarded (PHYS)", "Loss, stolen, or discarded portable device (PORT)", "Lost, stolen, or discarded electronic devices not designed to be moved (STAT)", "Unintended/accidental disclosure (DISC)", and "Unknown all other types (UNKN)".

As our goal is to investigate how peer firms respond to data breaches, I only focus on data breach events under "HACK" and involving a loss of customer and/or employee records into our analysis. This because hackings are malicious attempts of outsiders and

[9] The rest of breach organization types, which I use, are "Businesses-Financial and Insurance Services (BSF)", "Businesses – Other (BSO)"," Businesses-Retail/Merchant - Including Online Retail (BSR)", "Healthcare, Medical Providers & Medical Insurance Services (MED)", and "Unknown (UNKN)".

are more damaging and salient to peer firms. Anecdotal evidence shows that serious hacking events with a gigantic number of customer/employee records loss often result in costly lawsuit settlement with the victims.[10] In their examination of the effect on firms that are breached, Kamiya et al. (2020) also focus on hackings because they are hard to predict and more damaging once occurred.[11] I finally identify 129 hackings in 72 unique four-digit SIC industries over the period of 2008 to 2018.[12] I keep events starting from 2008 to allow at least one year with available IT data that start from 2007.

We then hand match the organization names that are hacked to the historical company names ever showed up in the Center for Research in Security Prices (CRSP) name tracking file. I also read the item "information source" in the PRC data to assist our matching. Table 1 describes the industry distribution of hacking events in our sample, and it shows that data hacking happened to each of the one-digit SIC industry, suggesting that it is a common issue. It also shows that hacking is more pronounced in banking, insurance, health services, retail, communication, industrial and electronic equipment, and wholesale industries.

---Insert Figure 1 to Figure 3 here---

---Insert Table 1 Here---

[10] Appendix A provides examples of litigation costs for data breaches.
[11] As a placebo test, I use all other types, excluding "hacking", of data breaches to construct a new sample of peer and control firms, and replicate our baseline results and find no significant effect, which is in line with our expectation. See Section 3.4.4.
[12] 25 out of the 72 unique 4-digit SIC codes experience hacking events multiple times.

Financial data are obtained from Standard & Poor's (S&P) Compustat database, stock price data come from CRSP. I winsorize all the continuous variables at 1% and 99% to mitigate the outlier concern. Data on disclosure of cyber risk are crawled from firms' 10-K filings.

3.3 Identifying peer firms

Following prior studies on peer effects (Leary and Roberts, 2014; Grennan, 2019) and on contagion effects (Beatty et al., 2013; Kedia et al., 2015), I define peer or treatment firms as firms in the same four-digit SIC industries as breach firms. I use the strictest industry classification (i.e., four-digit SIC code) to define peers to mitigate concerns that firms within a broader industry classification may not be close product market competitors. I define control groups as firms operating in four-digit SIC industries that have never encountered a hacking event over our sample period. Firms in industries that have never encountered a hacking event should be least affected when firms in other industries encounter a hacking event and thus serve a valid control group. Our final sample consists of 1,628 unique peer firms and 2,160 unique control firms from 2007 to 2018.

Summary statistics of the variables used in the analysis are presented in Table 2 Panel A. The mean annual spending on IT is 24.1 million dollars, but the median spending is only 2.69 million dollars. Since the IT spending variable is highly skewed, I use the natural logarithm of raw spending plus one dollar as the dependent variable in our analysis.

In Panel B of Table 2, I compare the major determinants of being hacked that are found to be significant in Kamiya et al. (2020) between treatment/peer firms and control firms in the year prior to a hacking event. I find that control firms have a lower leverage,

higher ROA, and lower asset intangibility. Importantly, treatment firms and control firms have statistically similar IT spending and the annual growth rate in IT spending in the year before a hacking event, suggesting the assumption on the existence of a parallel trend in IT investment between treatment firms and control firms is prima facie met. Since treatment firms and control firms exhibit some differences, I also report results from controlling for lagged firm characteristics as well as results from using a matched sample of control firms in a robustness check to ensure that our inference is not driven by the differences between treatment firms and control firms.

---Insert Table 2 Here---

4. Empirical design and results

4.1 Does a hacking event engender a negative short-term contagion effect on peer firms?

We start the analysis by conducting a short-term event study of peer firms' stock market reaction to the public disclosures of hacking events in the same industry. After observing hacking event in the same industry, investors of peer firms may have updated their Bayesian posterior assessment of the data breach risk of such firms and worry that peer firms are also likely to be targeted by hackers or are subject to the same loophole in IT systems, which predict a negative contagion effect on peer firms.

In order to have a clean test, I screen peer firms for confounding events by excluding peer firms that have any of the following announcements within the 11-day (-5, 5) window centered at the public annoucement date of a hacking event: (1) release of financial reports, such as 10-Qs and 10-Ks; and (2) M&A annoucements in 8-K filings. The procedure yields 580 events with available stock price data for the event study. I

compute cumulative abnormal returns (CARs) for peer firms during the event windows (-1, 1) , (-3, 3), and (-5, 5), where event day 0 is the disclosure date of the hacking event. The market model parameters are estimated over the period (-280, -61) with the CRSP value-weighted return as the market index. The results are reported in Table 3.

---Insert Table 3 Here---

Table 3 shows that peer firms' market reactions to the 129 hacking events occurred to other firms in the same four-digit SIC industry are significantly negative. In column (1), the means of CARs over the period of (-1, 1) is -0.285%, statistically siginificant at the 10% level, suggesting on average, firms lose around 22 million dollars (0.285%*7,714 million, 7,714 is the mean market capitalization for peer firms used in this test at day = -2) upon the public disclosure of a hacking event in the same industry. The mean of CARs over (-3, 3), and (-5, 5) is -0.673% and -0.779%, respectively, and significant at the 1% level. The median tests in column (2) show a similar pattern. The results suggest that although data hacking does not directly affect peer firms, peer firms on average bear a 0.3~0.8% loss of shareholder wealth. Our results are statistically and economically similar to the finding in Kamiya et al (2020). Note that I do not expect a super-large effect as a countervailing factor is that peer firms may actually benefit from such hacking events in the same industry if investors expect that some customers of hacked firms may switch to rivalry peer firms. In addition, given the long-tail nature of litigation claims for breach

firms (that often lasts for multiple years), the peer firms' investors could underestimate the total cost of a hacking event to the firm attacked by the hacking in the short term.[13]

Therefore, hacking events in the same industry appear to engender an immediate negative spillover effect on peer firms, and this negative effect may alter peer firms to take precautionary actions to increase IT investment following the occurrence of a hacking event in another firm of the same industry.

4.2 The identification strategy

We conduct a DID analysis to estimate peer firms' response to a hacking attack occurred to another firm in the same industry using the following baseline model specification:

$$IT\ investment_{i,j,t} = \theta_i + \delta_t + \beta Breach\ peer_{i,j,t} + X_{i,j,t-1} + \epsilon_{i,j,t} \qquad (1)$$

Where *i* indexes firm, *j* indexes 4-digit SIC industry, *t* indexes year. $IT\ Investment_{i,j,t}$ is the total amount of firm's IT investment in natural logarithm plus one dollar. θ_i is firm fixed effect to control for any time-invariant omitted variables, and δ_t is year fixed effect to control for year-specific factors. One concern in peer-effect studies is that any observed peer responses may be due to unobserved common factors that may independently drive the peer response and hence the 'reflection problem' (Manski, 1993). I therefore also examine the sensitivity of our results and mitigate the concern over the 'reflection problem' by including high-dimensional headquarter state-by-year fixed effects to capture

[13] For example, it takes at least two years for Equifax to settle the liability from the data breach in 2017. https://www.washingtonpost.com/technology/2019/07/22/equifax-pay-up-million-settle-state-federal-investigations-into-security-breach/

unobserved time-varying headquarter state-specific shocks and two-digit SIC industry-by-year fixed effects to capture time-varying industry-specific unobserved shocks in some model specification.

$Breach\ peer_{i,j,t}$ equals one if firm *i* belongs to the four-digit SIC industry *j* where another firm in the same four-digit SIC industry experiences a hacking attack in the previous 12 months and zero otherwise. β captures the average differences in the pre-to-post change of IT investment between the treatment/peer firms (i.e., firms in the same four-digit SIC industry as the hacked firm) and control firms (i.e., firms in different four-digit SIC industries that have never encountered a hacking event over our sample period).

$X_{i,t-1}$ are a set of control variables measured in one-period lag relative to the dependent variables including firm size (*Size*), leverage (*Leverage*), market-to-book ratio (*M/B*), return on assets (*ROA*), institutional block ownership (*Institutional block ownership*), asset intangibility (*Intangibility*), and cash holdings (*Cash holdings*). The choice of these variables are informed by the finding of Kamiya et al. (2020) who find that firms are more likely to be hacked when they have a larger firm size, higher valuation, higher ROA, higher institutional block ownership, higher asset intangibility, less financial constraints. These variables are included to control for the effect of differences in some firm characteristics between the treatment and control firms as shown in Panel B of Table 2. But since these control variables themselves may be affected by the hacking event and hence can be endogenous, I also report results without control variables. Finally, $\epsilon_{i,j,t}$ is the error term, and I cluster standard errors at the four-digit SIC industry level to account for potential cross-sectional correlations among firms in the same industry.

4.3 Do peer firms respond to hacking events by increasing IT investment?

Table 4 reports the baseline results from estimating Equation (1). The coefficient of the DID estimate, *Breach peer*, equals 0.164 and is significant at the 1% level in column (1), suggesting that peer firms respond to hacking attacks occurred other firms in the same industry by increasing IT investment by about 16% to lower the chance of being hacked, potential litigation and/or reputation risk. I then add high-dimensional fixed effects in column (2) and column (3). The magnitude of *Breach peer* is similar and statistically significant at the 1% level in column (2) when I include headquarter state-by-year fixed effects, and drops to about 10% when I further include two-digit SIC industry-by-year fixed effects in column (3). Therefore, our results are robust to controlling for unobserved time-varying local economic shocks and time-varying industry-specific shocks, which mitigates the concern that our observed peer responses in IT investment may be due to unobserved common factors that may independently drive the peer response as argued by the 'reflection problem' (Manski, 1993).

We include control variables in columns (4) to (6), and the statistical and economic significance of the DID estimate remain similar to the respective specification in column (1) to (3). This mitigates the concern that our observed increase in IT investment in peer firms is due to the differences in firm characteristics between peer/treatment firms and control firms.

For control variables, firms that are larger, have a higher institutional block ownership, or lower leverage are more likely to invest more IT. The sign of ROA, however, is unexpectedly negative. No other control variables are found to have significant effect on IT investment.

Overall, our baseline results show that peer firms vigilantly respond to the hacking event in their industry by increasing IT investments to strengthen their cybersecurity and reduce future litigation and/or reputation risk.

---Insert Table 4 Here---

Underpinning a valid DID analysis is the existence of a parallel trend between the treatment and control groups in the pre-event period. To verify this, I conduct a dynamic DID analysis in Table 5.

Specifically, I break *Breach peer* down into several variables based on the timing relative to the event year: *Breach peer (-2)* is a dummy variable that equals one if there is a hacking two years later. *Breach peer (-1)* is a dummy variable that equals one if there is a hacking one year later. *Breach peer (0)* is a dummy variable that equals one if the hacking event happened during the past 12 months. *Breach peer (1)* is a dummy variable that equals one if the hacking event happened one year ago. *Breach peer (2)* is a dummy variable that equals one if the hacking event happened two years ago. *Breach peer (3)* is a dummy variable that equals one if the hacking event happened three years ago. *Breach peer (4+)* is a dummy variable that equals one if the hacking event happened at least four years ago. The omitted group is years that are at least two years before the hacking event.

Results show that there is no significant pre-trend for the DID analysis before the hacking event as indicated by the insignificant coefficient on *Breach peer (-2)* and *Breach peer (-1)*, and hence there is no violation of the parallel trend assumption. The significant positive effects only appear in years from *Breach peer (0)* (the year when the hacking event

is publicized) onward and are persistent in the following three years. I also find that three years after, the magnitude of peer firms' IT investments tends to drop back to the pre-period level. The results suggest that the effects could be overreactions to salient risk but not permanent (Bordalo et al., 2012).

---Insert Table 5 Here---

4.4 Robustness of the results

4.4.1 Results from using a matched sample

We conduct several robustness checks of our baseline results. First, I use a matched sample as an alternative way of addressing the differences in firm characteristics between treatment and control firms (see Panel B of Table 2). I match a control firm (with replacement) to a peer firm one year prior to a hacking event based on the nearest propensity score estimated from logit matching regressions. As suggested by Kamiya et al. (2020), I include variables that could affect the probability of firms being hacked: *Size*, *M/B*, *ROA*, *Institutional block ownership*, *Intangibility*, and *Cash holdings*. Following Fang et al. (2014), I also include the level and the annual growth rate of IT investment in the matching process so that treatment and matched control firms not only have similar IT spending but also a parallel trend before hacking events. The procedure yields 898 matched pairs of peer firms and control firms for this analysis. The results of DID estimation using the matched sample are reported in Table 6. From Panel A, I can see that after matching, treatment firms and control firms are balanced in the matching covariates (i.e., the major determinants of being hacked). In Panel B, in all six columns of Table 6, the coefficients

of *Breach peer* consistently suggest an average of 13% increase in IT investment. Taken together, our baseline results are not attributable to the differences in firm characteristics between treatment and control firms.

---Insert Table 6 Here---

4.4.2 Results from using the first hacking event in an industry

Our previous baseline results are obtained using all available hacking events and one industry may be subject to multiple treatments. In this section, I provide a robustness check by focusing on the first hacking event (if any) of each industry. Specifically, for peer firms, I keep three years before and three years after the first hacking event in their respective industry. Observations falling in the same window for control firms operating in a different four-digit SIC industry that has never encountered a hacking in the sample period are also included in the analysis. *Breach peer* equals one if by year *t* a firm's four-digit SIC industry has experienced a hacking event and zero for otherwise. This setting is similar to studies investigating the effects of staggered law adoption, such as Bank Combination Law (Bertrand and Mullainathan, 2003) and Universal Demand Law (Lin et al. 2020). The results are reported in Table 7. As the table shows, the results are robust and the magnitude of the coefficient of *Breach Peer* in column (6) is about twice of that in Table 4 column (6), suggesting that the first hacking event in an industry engenders a larger peer response.

---Insert Table 7 Here---

4.4.3 Results from looking at different components of IT investment

We also take advantage of the richness of the IT data and examine the changes in different components of IT investment (i.e., communications, hardware, software, and IT services). I do not have predictions and so the analysis is explorative. The results reported in Table 8 suggest that increase in IT investment is not limited to a certain component of IT, but extends to all components.

---Insert Table 8 Here---

4.4.4 Results from a placebo test using non-hacking data breaches

We also conduct a placebo test by using non-hacking data breaches. As I discuss in Section 3, non-hacking data beaches such as lost laptops or internal errors are more ad-hoc, and not as serious as external hacking attack and are not expected to engender a significant peer response. I include data breaches with a loss of records other than type "HACK" and reconstruct our sample using the same specification to estimate. The results reported in Table 9 show that there is no significant peer effect when an industry encounters non-hacking data breaches, suggesting that only malicious hacking events can engender a peer effect as I predict.

---Insert Table 9 Here---

4.4.5 Results from using alternative control group

In Appendix Table A3, I define the control group as firms in the same three-digit SIC industry, but in a different four-digit SIC industry that has never encountered a hacking event over our sample period. Using this version of control variable sets a higher bar for generating a significant DID effect since one can argue that firms sharing the same three-digit SIC industry code as the firm that was hacked may also respond to the hacking event. Nevertheless, the results reported in Appendix Table A3 are consistent with our baseline results.

4.4.6 Results from excluding the financial crisis period

The negative demand shock imposed by the 2008 financial crisis may change peer firms' investment decisions and thus bias the estimates on IT investments. I replicate our baseline results in Table 4 using a subsample excluding years 2008-2009, the financial crisis period. The results reported in Appendix Table A4 are consistent with our baseline results.

4.5 An information transmission channel for the peer effect: board interlocking

In this section, as a further attempt to mitigate the concern that Manski's (1993) reflection problem (i.e., peer firms' changes simply reflect something in common with the hacked firms in the same industry) drives our baseline results, I conduct tests to directly show that board interlocking is a plausible networking channel through which relevant information is transmitted from the hacked firms to their peers, thereby resulting in peer firms' responses. Such information may include after a firm is hacked, the losses and costs

caused by the hacking, the possible pressure (e.g., from consumers, investors, suppliers, regulators and other stakeholders) the firm and its board of directors have received and gone through, the remedial measures the hacked firm has taken, issues and difficulties encountered, etc. If peer firms have a director or an executive sitting on the board of the hacked firm, they can learn the above information and has a duty to advise the corresponding peer firms on cybersecurity. I therefore predict that when a peer firm has a director interlocked with the hacked firm, the peer firm has a stronger response to a hacking event via IT investment.

In Table 10, I interact *Breach peer* with *Interlock board*, a dummy variable that equals one if there is board interlocking between a peer firm and a hacked firm in the year when a hacking event occurs and zero otherwise. Consistent with our expectation, the interaction term is loaded positively. Therefore, board interlocking is a plausible channel through which peer firms become responsive to hacking events occurred to another firm in the same industry. The result thus mitigates the concern that our observed baseline result is an artifact of the reflection problem (Manski, 1993).

---Insert Table 10 Here---

4.6 Heterogeneity in peer firms' IT investment response – the moderating effect of the exposure to a hacking attack

4.6.1 Hacking events characteristics as proxies for exposures

Hacking events that involve a larger loss of personal information should be more salient and have a larger impact on peer firms. I define a dummy variable *High record loss*

and interact with *Breach peer* to perform triple-difference estimations. *High record loss* equals one if the loss of record from a hacking event is in the top tercile of record loss among 129 hacking events included in our analysis and zero otherwise. The results are reported in column (1) of Table 11 and are consistent with our predictions.

---Insert Table 11 Here---

4.6.2 Firm financial characteristics as proxies for exposures

If peer firms are indeed vigilant to cybersecurity risk and respond to hacking events in the same industry, the response in IT investment should be stronger in peer firms that are more exposed to hacking risk. I measure such exposure in several ways.

First, Kamiya et al. (2020) suggest that firms with higher visibility, higher valuation, higher profitability, less financially constrained, and more intangible assets are more likely to become the target of hacking. In addition, firms that have higher advertising spending are likely to have wide publicity and a strong brand name. These firms are likely to suffer a larger reputation loss once a hacking occurs to them, and therefore are more incentivized to reduce cyber risk by increasing IT investment. Using these characteristics, I define several high exposure dummies and interact with *Breach peer* so that I can perform triple-difference estimations. Specially, *Big firm* equals one if a firm's total asset is above the top tercile of the sample distribution in the previous year and zero otherwise. *High valuation* equals one if a firm's Tobin's Q is above the top tercile of the sample distribution in the previous year and zero otherwise. *High ROA* equals one if a firm's return on assets is above the top tercile of the sample distribution in the previous year and zero otherwise. Following

Kamiya et al. (2020) I use the Whited and Wu's (2006) index to measure financial constraints. *Less constrained* equals one if a firm's Whited and Wu's (2006) index is below the bottom tercile of the sample distribution in the previous year and zero otherwise. *High intangibility* equals one if a firm's tangible assets ratio is above the top tercile of the sample distribution in the previous year and zero otherwise. *High advertising* equals one if a firm's advertising expenses scaled by total assets is above the top tercile in a given year and zero otherwise. The results are reported in Table 11. I find that the interaction effects are significantly positive, except for high intangibility (untabulated), and the standalone item of *Breach peer* is positively loaded.

4.6.3 Industry characteristics as exposure proxies

Firms operating in certain industries (e.g., financial firms, hospitality industries such as hotels and restaurants; telecommunication firms, retail and wholesale firms) have large amounts of user data and hence are more likely to be targeted by hackers. I introduce a dummy variable that equals one for these industries and interact it with *Breach peer* and the triple-difference results are reported in column (7) of Table 11. The results on positive and significant coefficients of the interaction term are consistent with our expectation.

---Insert Table 12 Here---

4.7 Heterogeneity in peer firms' IT investment response – the moderating effect of risk management awareness

A peer firm's response to a hacking event in the same industry also depends on its risk management awareness. I proxy a firm's risk management awareness in two ways: the awareness is higher if a firm has a designated risk officer or a risk committee overseeing a firm's risk management, or has explicitly discussed cybersecurity risk in 10-K filings before a hacking.

We construct a dummy variable *Risk position* that equals one if a firm has an executive with a risk title or has a risk committee. I collect the data from Execucomp and identify the names of risk title by hand.

To measure whether a firm discusses cybersecurity in corporate filings, I build a data library of keywords related to data risk from firms' 10-K filings and use total word count as proxies for cyber risk disclosure. Specifically, I first read sufficient breach firms' 10-Ks to summarize a list of keywords. The list includes "data?breach*", "cyber?attack*", "cyber?security*", "cyber?terrorism (terrorists)*", "cyber?threat*", "data?risk*", "data?security*", "hack*", "malware*", "privacy?risk*", and "spyware*".[14] Then, I use a Python code to search above keywords in parsed 10-Ks text files and generate a total count number of these characters for our sample firms.[15] Finally, I generate a variable, *Cyber risk discussion*, which equals one if cyber risk related words appear at least once in firms 10-k filings in a given year.

[14] "?" and "*" are used as wildcards in the search when encounters different formats of word. Specifically, "?" denotes symbol for the meaning of "or" when the search encounters a space or a "-" (e.g., "cyberattack", "cyber-attach", and "cyber attack"). "*" denotes the trailing characters so that plurals (e.g. data breaches) are picked up in the search.

[15] The parsed 10-X texts are public available from The Notre Dame Software Repository for Accounting and Finance at https://sraf.nd.edu/.

We interact *Risk position* and *Cyber risk dicussion* measured in the previous year with *Breach peer* respectively, and add to Equation (1). The results reported in Table 13 show that as predicted, the interaction terms are significantly positive, consistent with our prediction that peer firms with higher risk management awareness are more responsive to hacking events in the same industry by increasing IT investment.

---Insert Table 13 Here---

4.8 The moderating effect of SEC cybersecurity disclosure guidance in 2011

As I brief in Introduction, in light of the growing cyber risk, the division of corporate finance of the SEC issued a guidance to strengthen the disclosure of cyber risk among public firms in October 2011. Prior studies on the effect of this guidance have generated mixed evidence. On the one hand, the guidance has led to a rapid increase in cybersecurity disclosures in breach firms' 10-K reports, suggesting that the guidance has an effect (Berkman et al., 2018). On the other hand, Hilary et al., (2016) argue that the guidance is largely a boilerplate due to the vague definition of material risk narrated in the guidance. In this section, I explore whether the 2011 SEC guidance on cyber security disclosure enhances the peer effects in IT investment or not.

We define *Post-SEC guidance 2011* as a dummy that equals one for the period after 2011 and zero otherwise, I then interact it with *Breach peer*. The result from the triple-difference analysis is reported in Table 14. I find that the interaction term has a positive and significant coefficient, suggesting that the 2011 SEC guidance enhances peer firms' response in IT investment to hacking events; whereas before the issuance of the 2011 SEC

guidance, the response is insignificant. Therefore, the effect of 2011 SEC guidance is not a boilerplate as peer firms exhibit more significant responses in IT investment to hacking events after the issuance of the 2011 cyber risk disclosure guidance.

---Insert Table 14 Here---

In Appendix Table A5, I also directly examine whether peer firms increase their disclosure of the cybersecurity risk in 10-K filings as a response to hacking events occurred to other firms in their industry. The DID estimation results show that peer firms increase the disclosure of cybersecurity risk by about 5% following a hacking attack in other firms in the same industry. This can be a channel for our baseline results on peer firms' increase in IT investment or peer firms simultaneously use both increased disclosure and IT investment as complementary precautionary actions to manage the potential hacking risk, litigation and/or reputation risk.

In Appendix Table A6, I examine whether the increase in IT investment has a spillover effect on other corporate investment decisions (acquisition, CAPEX excluding IT spending, R&D, advertising). It appears that firms increase acquisitions and R&D, but cut non-IT related fixed asset investment.

5. Conclusion

Exploiting a list of malicious hacking against U.S. public firms from 2008 to 2018 and using a difference-in-difference analysis, I find that that peer firms (i.e., firms with the same four-digit SIC code as hacked firms) significantly increase their IT investment

following a hacking in their industry. The results hold for several robustness checks including using matched control groups and alternative identification specifications. The effect is more pronounced for firms that are more exposed to cyberattacks, have higher risk management awareness (denoted by having a designated risk officer or a risk committee, or having previously discussed cyber risk in corporate filings), and for the period after the SEC's issuance of the guidance on more disclosure on cyber risk in 2011. I find that board interlocking with breach firms is a channel for the relevant information on the effects of hacking attack and the entailing corporate responses to be transmitted from data breach firms to peer firms.

Overall, our results suggest that peer firms respond to hackings in the same industry by taking precautionary actions to manage the potential cyberattack risk facing the company, litigation and/or reputation risk facing the board. In this way, a negative event (i.e., a hacking attack) may result in positive externalities to the industry and its stakeholders.

By providing the first piece evidence on data breach events engendering a peer effects in real economic activities, our evidence contributes to the understanding and corporate risk management of cybersecurity risk - an emerging and increasingly important operational risk. Our findings have important implications for investors, corporates, regulators, and other stakeholders.

References

Aguilar, Luis A. 2014. Speech on "Boards of Directors, Corporate Governance and Cyber-Risks: Sharpening the Focus." Available at https://www.sec.gov/news/speech/2014-spch061014laa

Akey, Pat, Stefan Lewellen, and Inessa Liskovich. 2018. "Hacking Corporate Reputations." *SSRN Electronic Journal* 1(814).

Armstrong, Christopher S., Stephen Glaeser, and John D. Kepler. 2019. "Strategic Reactions in Corporate Tax Planning." *Journal of Accounting and Economics* 68(1): 101-232.

Beatty, Anne, Scott Liao, and Jeff Jiewei Yu. 2013. "The Spillover Effect of Fraudulent Financial Reporting on Peer Firms' Investments." *Journal of Accounting and Economics* 55(2–3): 183–205.

Berkman, Henk, Jonathan Jona, Gladys Lee, and Naomi Soderstrom. 2018. "Cybersecurity Awareness and Market Valuations." *Journal of Accounting and Public Policy* 37(6): 508–526.

Bloom, Nicholas, Luis Garicano, Raffaella Sadun, and John Van Reenen. 2014. "The Distinct Effects of Information Technology and Communication Technology on Firm Organization." *Management Science* 60(12): 2859–2885.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer. 2012 "Salience theory of choice under risk." *The Quarterly journal of economics* 127(3): 1243-1285.

Brynjolfsson, Erik, and Lorin M. Hitt. 2003. "Computing Productivity: Firm-Level Evidence." *Review of Economics and Statistics* 85(4): 793–808.

Cheng, Shijun, Robert Felix, and Raffi Indjejikian. 2019. "Spillover Effects of Internal Control Weakness Disclosures: The Role of Audit Committees and Board Connections." *Contemporary Accounting Research* 36(2): 934–957.

Ettredge, Michael L., and Vernon J. Richardson. 2003. "Discussion of Information Transfer among Internet Firms: The Case of Hacker Attacks." *Journal of Information Systems* 17(2): 83–86.

Dimmock, Stephen G., William C. Gerken, and Nathaniel P. Graham. 2018. "Is Fraud Contagious? Coworker Influence on Misconduct by Financial Advisors." *Journal of Finance* 73(3): 1417–1450.

Fang, Vivian W., Xuan Tian, and Sheri Tice. "Does stock liquidity enhance or impede firm innovation?." *Journal of Finance* 69.5 (2014): 2085-2125.

Files, Rebecca, and Umit G. Gurun. 2018. "Lenders' Response to Peer and Customer Restatements." *Contemporary Accounting Research* 35(1): 464–493.

Gande, Amar, and Craig M. Lewis. 2009. "Shareholder-Initiated Class Action Lawsuits: Shareholder Wealth Effects and Industry Spillovers." *Journal of Financial and Quantitative Analysis* 44(4): 823–850.

Gleason, Cristi, Nicole Jenkins, and W. Bruce Johnson. 2005. "Financial Statement Credibility: The Contagion Effects of Accounting Restatements." *Accounting Review* 83(1): 83–110.

Grennan, Jillian. 2019. "Dividend Payments as a Response to Peer Influence." *Journal of Financial Economics* 131(3): 549–570.

Hilary, Gilles, Benjamin Segal, and May H. Zhang. 2016. "Cyber-Risk Disclosure: Who Cares?" *SSRN Electronic Journal*: 1–59.

Huang, Henry H., and Wang, Chong. 2020 "Do Banks Price Firms' Data Breaches?" *Accounting Review, forthcoming*.

Kaiser, Zachary. (2018). Data breaches versus cyber liabilities: Are you protected? September 26, https://www.mcclone.com/blog/data-breaches-vs-cyber-liability-are-you-protected

Kamiya, S., Kang, J.K., Kim, J., Milidonis, A. and Stulz, R.M., 2020. Risk management, firm reputation, and the impact of successful cyberattacks on target firms. *Journal of Financial Economics, forthcoming*.

Kedia, Simi, Kevin Koh, and Shivaram Rajgopal. 2015. "Evidence on Contagion in Earnings Management." *Accounting Review* 90(6): 2337–2373.

Leary, Mark T., and Michael R. Roberts. 2014. "Do Peer Firms Affect Corporate Financial Policy?" *Journal of Finance* 69(1): 139–178.

Lin, Chen, Sibo Liu, and Gustavo Manso. 2020. "Shareholder Litigation and Corporate Innovation." *Management Science* (October).

Lin, Yupeng, Ying Mao, and Zheng Wang. 2018. "Institutional Ownership, Peer Pressure, and Voluntary Disclosures." *Accounting Review* 93(4): 283–308.

Messing, Maren J., and James Zucke. 2015. "Bennek v. Home Depot and the future of Cybersecurity-related Derivative Suits." September 22, 2015. Available at https://www.pbwt.com/data-security-law-blog/bennek-v-home-depot-cyber-security-related-derivative-suits-against-officers-and-directors/

Makridis, Christos, and Benjamin Dean. 2018. "Measuring the Economic Effects of Data Breaches on Firm Outcomes: Challenges and Opportunities." *Journal of Economic and Social Measurement* 43(1–2): 59–83.

Manski, Charles F. 1993. "Identification of Endogenous Social Effects the Reflection Problem." *Review of Economic Studies* 60(3): 531–542.

Masulis, Ronald W., Sichen Shen, and Hong Zou. 2020. "Director Liability Protection and the Quality of Outside Directors" European Corporate Governance Institute – Finance *Working Paper* No. 672/2020.

Nordlund, James. 2019. "The Disclosure of Cybersecurity Risk." *Working Paper*.

PricewaterhouseCoopers. (2014). "Economic impact of trade secret theft: A framework for companies to safeguard trade secrets and mitigate potential threats."

Tuzel, Selale, and Miao Ben Zhang. 2017. "Local Risk, Local Factors, and Asset Prices." *Journal of Finance* 72(1): 325–370.

Whited, Toni M., and Guojun Wu. 2006. "Financial Constraints Risk." *Review of Financial Studies* 19(2): 531–559.

Youn, Soo. (2019). "The Capital One data breach is alarming, but these are the 5 worst corporate hacks." https://abcnews.go.com/Technology/marriotts-data-breach-largelargest-worst-corporate-hacks/story?id=59520391

Figure 1 Number of data breach events

This figure reports the number of data breach events in PRC from 2005 to 2018. The dot line represents the fitted values.

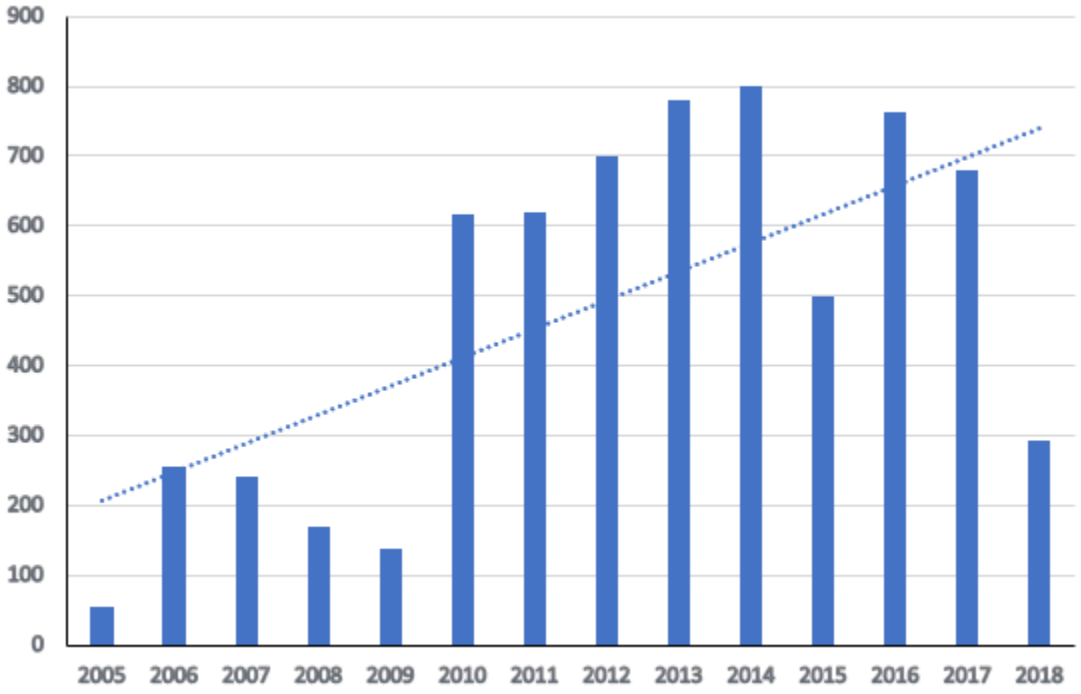

Figure 2 Logged number of total loss of data records in data breaches

This figure reports the logged number of total loss records from data breach events in PRC from 2005 to 2018. The dot line represents the fitted values.

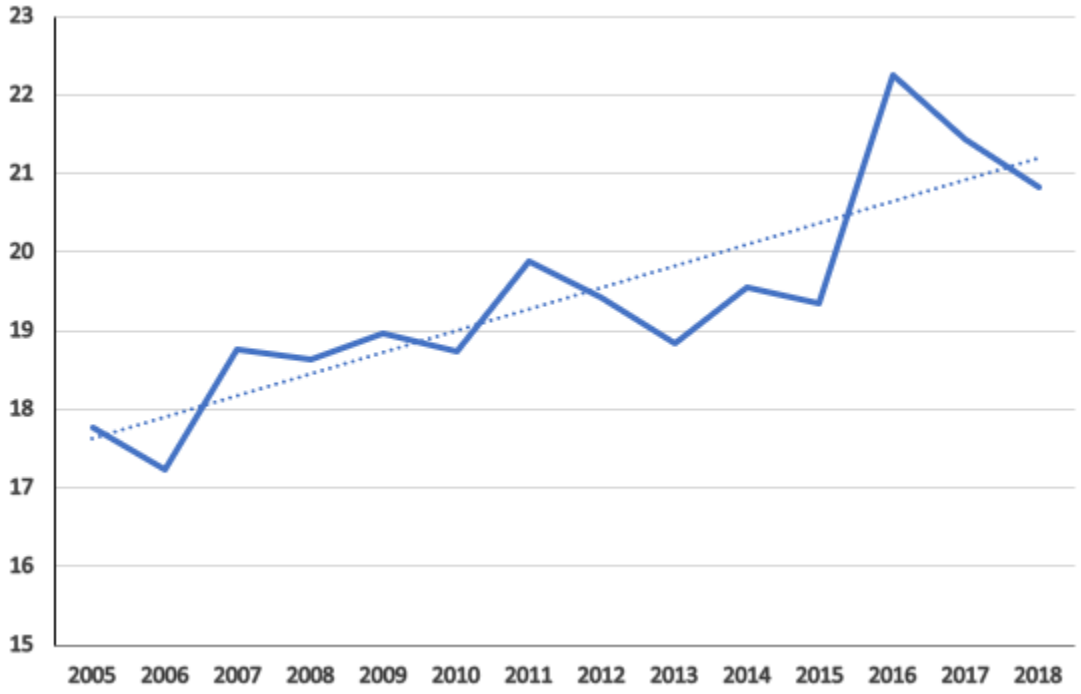

Figure 3 Geographic distribution of data breaches

This figure shows the geographic distribution of all data breaches in PRC from 2005 to 2018. The darker(lighter) blue indicates that the states have experienced more(fewer) data breaches. The bottom right displays the legend showing the scale for different colors in the figure.

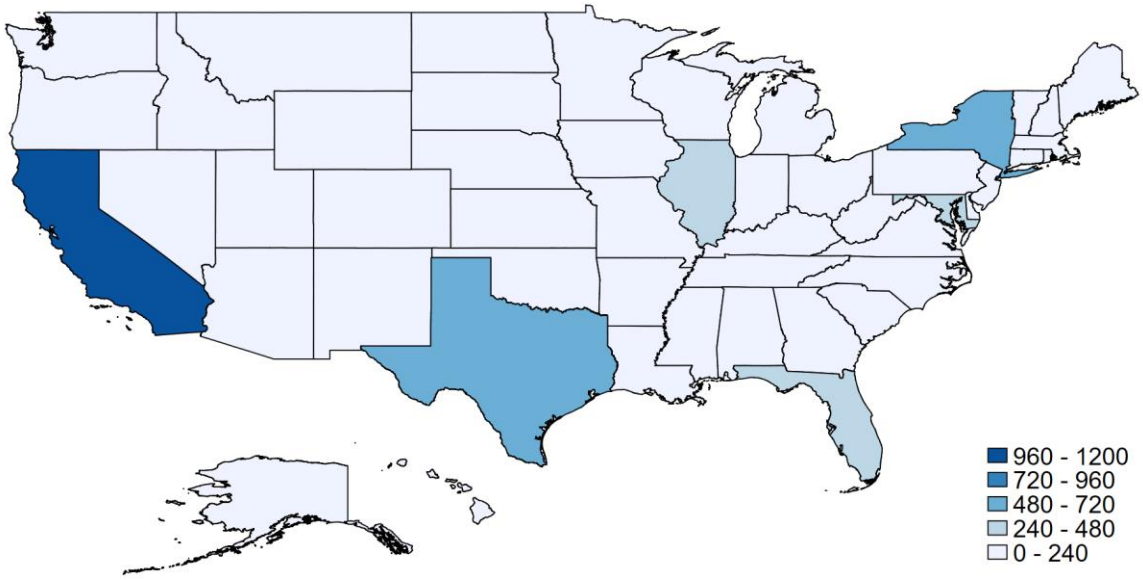

Table 1 Industry distribution of hacking events used in our sample

This table descries the industries of hacking events used in our sample. I include 129 hacking events resulting in customer and/or employee/ record loss in 72 unique four-digit SIC industries from 2008 to 2018. For brevity, this table reports the distribution of hacking events at the two-digit SIC industry level.

two-digit SIC code	Description	N	two-digit SIC code	Description	N
13	Oil and Gas Extraction	1	57	Furniture and Home furnishings Stores	2
23	Apparel and Other Textile Products	2	58	Eating and Drinking Places	2
27	Printing and Publishing	2	59	Miscellaneous Retail	9
28	Chemical and Allied Products	3	60	Depository Institutions	7
29	Petroleum and Coal Products	2	61	Non depository Institutions	2
30	Rubber and Misc. Plastics Products	2	62	Security and Commodity Brokers	3
35	Industrial Machinery and Equipment	5	63	Insurance Carriers	10
36	Electronic and Other Electric Equipment	5	70	Hotels and Other Lodging Places	1
38	Instruments and Related Products	7	73	Business Services	27
45	Transportation by Air	1	78	Motion Pictures	1
48	Communications	9	79	Amusement and Recreation Services	2
49	Electric, Gas, and Sanitary Services	1	80	Health Services	6
50	Wholesale Trade – Durable Goods	2	82	Educational Services	1
51	Wholesale Trade – Nondurable Goods	2	87	Engineering and Management Services	4
53	General Merchandise Stores	5	99	Non-Classifiable Establishments	1
55	Automotive Dealers and Service Stations	2	Total		129

Table 2 Summary statistics

This table reports the descriptive statistics of the main variables used in the analysis. Panel A reports statistics of the main variables used in the analysis. Panel B presents average values of the variables for peer and control firms one year before hacking events. The unit of total IT spending used in the analysis is one dollar, while the statistics here are reported in 1,000 dollars for brevity. I provide all variable definitions in Appendix Table A1. I winsorize all continuous variables at the 1st/99th percentiles. The significance is based on the *p*-values for *t*-tests and *z*-values for Wilcoxon signed-rank tests that the mean and the median difference of variables between peer firms and control firms are equal to zero. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

Panel A: Summary statistics

Variables	Mean	SD	P25	Median	P75	N
Total IT spending ($1,000)	24,100.00	73,600.00	523.32	2,690.00	12,400.00	25,283
Size	7.24	1.90	6.01	7.24	8.45	25,283
Leverage	0.24	0.21	0.06	0.19	0.36	25,283
M/B	2.46	3.78	1.06	1.69	2.91	25,283
ROA	0.01	0.12	0.00	0.02	0.06	25,283
Institutional block ownership	0.32	0.19	0.19	0.32	0.44	21,677
Intangibility	0.79	0.24	0.69	0.89	0.97	25,283
Cash holdings	0.13	0.15	0.03	0.07	0.18	25,283

Panel B: Comparison of firm characteristics between the peer firms and control firms in the year before hacking events

	Peer firms		Control firms		Test of difference	
Variables	Mean	Median	Mean	Median	Mean	Median
Ln (1+Total IT spending)	15.00	14.94	14.80	14.99	-0.21	0.05
Growth rate in total IT spending	0.022	0.017	0.025	0.019	0.003	0.002
Size	7.41	7.29	7.11	7.15	-0.30	-0.14
Leverage	0.18	0.14	0.27	0.25	0.09***	0.11***
M/B	2.27	1.76	7.86	1.93	5.60	0.17
ROA	-0.03	0.02	0.03	0.04	0.03**	0.02***
Institutional block ownership	0.30	0.30	0.33	0.32	0.03	0.02
Intangibility	0.90	0.95	0.75	0.83	-0.15***	-0.12***
Cash holdings	0.17	0.12	0.12	0.07	0.04**	-0.05**

Table 3 Cumulative abnormal returns (CARs) for peer firms around hacking events

This table reports the cumulative abnormal returns (CARs) for peer firms around hacking events. I exclude peer firms that have confounding announcements within the 11-day (-5,5) window centered at the public date of hacking events: (1) release of financial reports, such as 10-Qs and 10-Ks. (2) release of merger-relate 8-K filings. The procedure yields 580 events with available stock price data for the event study. The abnormal stock returns are calculated using the market model. The market model parameters are estimated over the period (−280, −61) with the CRSP value-weighted return as the market index. The numbers in parentheses are *p*-values for *t*-tests and *z*-values for Wilcoxon signed-rank tests that the mean CAR and the median CAR are equal to zero. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

	Mean	*Median*
CAR (%)	(1)	(2)
CAR (-1,1)	-0.285*	-0.172**
	(0.069)	(0.031)
CAR (-3,3)	-0.673***	-0.614***
	(0.001)	(0.000)
CAR (-5,5)	-0.779***	-0.574***
	(0.000)	(0.000)

Table 4 DID analysis of hacking events on peer firms' IT investments

This table reports the results firm difference-in-differences (DID) analysis of hacking events on peer firms' IT investments. The independent variable is *Breach Peer*, which equals one if a firm belongs to the four-digit SIC industry where another firm in the same four-digit SIC industry experiences a hacking attack in the previous 12 months and zero otherwise. The dependent variable is the natural logarithm of one plus *Total IT spending*. I provide all variable definitions in Appendix Table A1. I winsorize all continuous variables at the 1st/99th percentiles. I include firm, year, state-by-year, and industry (two-digit SIC code)-by-year fixed effects as indicated in different columns. I report the *t*-statistics in parentheses with robust standard errors clustered at four-digit SIC code industry level. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

Y=	*Ln (1+Total IT spending)*					
	(1)	(2)	(3)	(4)	(5)	(6)
Breach peer	0.164***	0.164***	0.073**	0.169***	0.171***	0.096**
	(3.13)	(3.09)	(2.03)	(3.26)	(3.26)	(2.17)
Size				0.303***	0.298***	0.328***
				(5.82)	(5.91)	(6.73)
Leverage				-0.369	-0.415**	-0.172
				(-1.64)	(-1.98)	(-1.22)
M/B				0.002	0.002	0.004
				(0.41)	(0.41)	(1.08)
ROA				-0.405***	-0.362**	-0.184
				(-2.83)	(-2.54)	(-1.64)
Institutional block ownership				0.193***	0.163**	0.063
				(2.93)	(2.39)	(1.12)
Intangibility				-0.425	-0.343	-0.259
				(-1.53)	(-1.20)	(-1.06)
Cash holdings				0.040	0.025	-0.077
				(0.21)	(0.14)	(-0.45)
Observations	25,159	25,159	25,159	21,642	21,642	21,642
Adj. R-squared	0.832	0.833	0.863	0.840	0.841	0.871
Firm FE	YES	YES	YES	YES	YES	YES
Year FE	YES			YES		
State-Year FE		YES	YES		YES	YES
Industry-Year FE			YES			YES

Table 5 DID dynamic analysis of hacking events on peer firms' IT investments

This table reports the results firm difference-in-differences (DID) dynamics analysis of hacking events on peer firms' IT investments. *Breach peer (-2)* is a dummy variable that equals one if there is a hacking two years later. *Breach peer (-1)* is a dummy variable that equals one if there is a hacking one year later. *Breach peer (0)* is a dummy variable that equals one if the hacking event happened during the past 12 months. *Breach peer (1)* is a dummy variable that equals one if the hacking event happened one year ago. *Breach peer (2)* is a dummy variable that equals one if the hacking event happened two years ago. *Breach peer (3)* is a dummy variable that equals one if the hacking event happened three years ago. *Breach peer (4+)* is a dummy variable that equals one if the hacking event happened at least four years ago. The omitted group is years that are at least two years before the hacking event. The dependent variable is the natural logarithm of one plus *Total IT spending*. I provide all variable definitions in Appendix Table A1. I winsorize all continuous variables at the 1st/99th percentiles. I report the *t*-statistics in parentheses with robust standard errors clustered at four-digit SIC code industry level. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

Y=	*Ln (1+Total IT spending)*	
	(1)	(2)
Breach peer (-2)	-0.054	-0.064
	(-0.76)	(-0.78)
Breach peer (-1)	0.061	0.063
	(1.52)	(1.51)
Breach peer (0)	0.212**	0.230***
	(2.38)	(2.66)
Breach peer (1)	0.225**	0.220**
	(2.22)	(2.24)
Breach peer (2)	0.283*	0.296**
	(1.95)	(2.05)
Breach peer (3)	0.078	0.114
	(0.78)	(1.30)
Breach peer (4+)	-0.031	-0.011
	(-0.43)	(-0.16)
Size		0.286***
		(5.62)
Leverage		-0.202
		(-1.30)
M/B		0.000
		(0.18)
ROA		0.029***
		(11.34)
Institutional block ownership		0.186***
		(2.82)
Intangibility		-0.504*
		(-1.83)
Cash holdings		-0.039
		(-0.22)
Observations	25,159	21,642
Adj. R-squared	0.833	0.842
Firm FE	YES	YES
Year FE	YES	YES

Table 6 Robustness of DID analysis of hacking events on peer firms' IT investments: Using a matched sample

This table reports the results of robustness for difference-in-differences (DID) analysis of hacking events on peer firms' IT investments, using a matched sample. Specifically, I match a control firm (with replacement) to a peer firm one year prior to a hacking event based on the nearest propensity score estimated from logit regressions. As suggested by Kamiya et al. (2020), I include variables that could affect the probability of firms being hacked: *Size*, *M/B*, *ROA*, *Institutional block ownership*, *Intangibility*, and *Cash holdings*. Following Fang et al. (2014), I also include the level and the annual growth of IT investment in the matching process so that treatment and matched control firms not only have similar IT spending but also a parallel trend before hacking events. The procedure yields 898 matched pairs of peer firms and control firms for this analysis. Panel A reports the covariate balance check between peer firms and control firms after matching. Panel B reports the DID results. The independent variable is *Breach Peer*, which equals one if a firm belongs to the four-digit SIC industry where another firm in the same four-digit SIC industry experiences a hacking attack in the previous 12 months and zero otherwise. The dependent variable is the natural logarithm of one plus *Total IT spending*. I provide all variable definitions in Appendix Table A1. I winsorize all continuous variables at the 1st/99th percentiles. I include firm, year, state-by-year, and industry (two-digit SIC code)-by-year fixed effects as indicated in different columns. I report the *t*-statistics in parentheses with robust standard errors clustered at four-digit SIC code industry level. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

Panel A: Covariate balance check

	Peer firms	Control firms	Test of difference
Ln (1+Total IT spending)	15.32	15.49	-0.17
Growth rate in total IT spending	0.15	0.12	0.03
Size	7.61	7.69	-0.08
M/B	2.43	2.42	0.01
ROA	0.04	-0.01	0.05
Institutional block ownership	0.31	0.32	-0.01
Intangibility	0.91	0.90	0.01
Cash holdings	0.14	0.15	-0.01

Panel B: DID using the matched sample

Y=	*Ln (1+Total IT spending)*					
	(1)	(2)	(3)	(4)	(5)	(6)
Breach peer	0.135***	0.127***	0.121***	0.125***	0.118***	0.125***
	(3.40)	(3.11)	(2.95)	(2.93)	(2.71)	(2.71)
Size				0.475***	0.479***	0.490***
				(6.92)	(6.54)	(6.28)
Leverage				-0.565**	-0.435*	-0.292
				(-2.28)	(-1.97)	(-1.27)
M/B				0.002	0.003	0.002
				(0.22)	(0.33)	(0.25)
ROA				-0.718***	-0.636***	-0.447*
				(-3.23)	(-2.84)	(-1.87)
Institutional block ownership				0.321**	0.248*	0.237*
				(2.47)	(1.78)	(1.90)
Intangibility				-0.991*	-0.818	-0.984
				(-1.93)	(-1.47)	(-1.49)
Cash holdings				-0.117	-0.091	0.086
				(-0.47)	(-0.33)	(0.38)
Observations	7,111	7,111	7,111	6,805	6,805	6,805
Adj. R-squared	0.852	0.855	0.874	0.859	0.860	0.879
Firm FE	YES	YES	YES	YES	YES	YES
Year FE	YES			YES		
State-Year FE		YES	YES		YES	YES
Industry-Year FE			YES			YES

Table 7 Robustness of DID analysis of hacking events on peer firms' IT investments: Using the first hacking event as an alternative identification strategy

This table reports the results of robustness for difference-in-differences (DID) analysis of hacking events on peer firms' IT investments, using the first hacking event of each industry as an alternative identification. Specifically, for peer firms, I keep three years before and three years after the first hacking event in their respective industry. Observations falling in the same window for control firms operating in a different four-digit SIC industry that has never encountered a hacking in the sample period are also included in the analysis. *Breach peer* equals one if by year *t* a firm's four-digit SIC industry has experienced a hacking event and zero for otherwise. This setting is similar to studies investigating the effects of staggered law adoption, such as Bank Combination Law (Bertrand and Mullainathan, 2003) and Universal Demand Law (Lin et al. 2020). The dependent variable is the natural logarithm of one plus *Total IT spending*. I provide all variable definitions in Appendix Table A1. I winsorize all continuous variables at the 1st/99th percentiles. I include firm, year, state-by-year, and industry (two-digit SIC code)-by-year fixed effects as indicated in different columns. I report the *t*-statistics in parentheses with robust standard errors clustered at four-digit SIC code industry level. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

Y=	*Ln (1+Total IT spending)*					
	(1)	(2)	(3)	(4)	(5)	(6)
Breach peer	0.252***	0.237***	0.159***	0.275***	0.251***	0.180***
	(3.84)	(3.68)	(3.26)	(3.88)	(3.63)	(3.41)
Size				0.335***	0.330***	0.340***
				(5.87)	(5.73)	(6.14)
Leverage				-0.423**	-0.474**	-0.244*
				(-2.10)	(-2.39)	(-1.71)
M/B				0.001	0.001	0.003
				(0.26)	(0.33)	(0.75)
ROA				-0.423***	-0.383**	-0.214
				(-2.97)	(-2.54)	(-1.51)
Institutional block ownership				0.240***	0.209**	0.086
				(2.97)	(2.42)	(1.19)
Intangibility				-0.393	-0.250	-0.192
				(-1.31)	(-0.81)	(-0.74)
Cash holdings				-0.125	-0.126	-0.112
				(-0.71)	(-0.71)	(-0.65)
Observations	20,410	20,410	20,410	17,642	17,642	17,642
Adj. R-squared	0.847	0.848	0.873	0.850	0.851	0.876
Firm FE	YES	YES	YES	YES	YES	YES
Year FE	YES			YES		
State-Year FE		YES	YES		YES	YES
Industry-Year FE			YES			YES

Table 8 DID analysis of hacking events on peer firms' different components of IT investments

This table reports the results from difference-in-differences (DID) analysis of hacking events on peer firms' different components of IT investments. The independent variable is *Breach Peer*, which equals one if a firm belongs to the four-digit SIC industry where another firm in the same four-digit SIC industry experiences a hacking attack in the previous 12 months and zero otherwise. The dependent variable includes the four components of total IT investment: Communications, Hardware, Software, and Services. I use the natural logarithm of one plus each variable as dependent variables as indicated in different columns. I provide all variable definitions in Appendix Table A1. I winsorize all continuous variables at the 1st/99th percentiles. I report the *t*-statistics in parentheses with robust standard errors clustered at four-digit SIC code industry level. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

Y=	*Commnunications*	*Hardware*	*Softwares*	*Services*
	(1)	(2)	(3)	(4)
Breach peer	0.167**	0.131***	0.172***	0.145***
	(2.27)	(2.85)	(2.86)	(3.19)
Size	0.337***	0.281***	0.330***	0.283***
	(6.52)	(5.34)	(6.07)	(5.42)
Leverage	-0.445*	-0.344	-0.438*	-0.277
	(-1.87)	(-1.56)	(-1.83)	(-1.32)
M/B	0.002	0.002	0.001	0.001
	(0.63)	(0.50)	(0.23)	(0.30)
ROA	-0.423***	-0.435***	-0.385**	-0.386***
	(-2.98)	(-2.91)	(-2.39)	(-2.81)
Institutional block ownership	0.177**	0.187***	0.215***	0.185***
	(2.33)	(2.69)	(3.09)	(2.84)
Intangibility	-0.868**	-0.565**	-0.481	-0.363
	(-2.55)	(-2.00)	(-1.56)	(-1.35)
Cash holdings	0.110	0.129	0.135	0.035
	(0.56)	(0.68)	(0.67)	(0.19)
Observations	21,642	21,642	21,642	21,642
Adj. R-squared	0.822	0.811	0.853	0.873
Firm FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

Table 9 A placebo test of hacking events on peer firms' IT investments: Using non-hacking data breaches

This table reports the results of a placebo test for difference-in-differences (DID) analysis of hacking events on peer firms' IT investments, using non-hacking data breaches. I include data breaches with a loss of records other than type "HACK" and reconstruct our sample using the same specification to estimate. The types of data breach in this analysis include "Payment card fraud (CARD)", "Insider taking advantage of company system (INSD)", "Physical loss, stolen, or discarded (PHYS)", "Loss, Stolen, or Discarded Portable Device (PORT)", "Lost, stolen, or discarded electronic devices not designed to be moved (STAT)", "Unintended/accidental disclosure (DISC)", and "Unknown all other types (UNKN)". The independent variable is *Breach Peer*, which equals one if a firm belongs to the four-digit SIC industry where another firm in the same four-digit SIC industry experiences a hacking attack in the previous 12 months and zero otherwise. The dependent variable is the natural logarithm of one plus *Total IT spending*. I provide all variable definitions in Appendix Table A1. I winsorize all continuous variables at the 1st/99th percentiles. I include firm, year, state-by-year, and industry (two-digit SIC code)-by-year fixed effects as indicated in different columns. I report the *t*-statistics in parentheses with robust standard errors clustered at four-digit SIC code industry level. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

Y=	*Ln (1+Total IT spending)*					
	(1)	(2)	(3)	(4)	(5)	(6)
Breach peer	0.025	0.023	-0.046	0.040	0.035	-0.047
	(0.85)	(0.74)	(-1.56)	(1.36)	(1.18)	(-1.41)
Size				0.306***	0.304***	0.326***
				(5.76)	(5.90)	(6.51)
Leverage				-0.395	-0.429*	-0.145
				(-1.62)	(-1.89)	(-1.01)
M/B				0.001	0.002	0.005
				(0.30)	(0.46)	(1.22)
ROA				-0.422***	-0.380***	-0.180
				(-2.91)	(-2.61)	(-1.60)
Institutional block ownership				0.218***	0.192***	0.085
				(3.36)	(2.89)	(1.54)
Intangibility				-0.461*	-0.377	-0.302
				(-1.69)	(-1.35)	(-1.26)
Cash holdings				0.082	0.041	-0.089
				(0.73)	(0.38)	(-0.88)
Observations	25,043	25,043	25,043	21,539	21,539	21,539
Adj. R-squared	0.830	0.831	0.862	0.838	0.839	0.870
Firm FE	YES	YES	YES	YES	YES	YES
Year FE	YES			YES		
State-Year FE		YES	YES		YES	YES
Industry-Year FE			YES			YES

Table 10 An information transmission channel for the peer effect: board interlocking

This table reports the triple difference estimates of the effect of hacking events on peer firms' IT investments when the board of peer firms is interlocked with the board of breach firms. The dependent variable is *Total IT spending*, defined as the logarithm of one plus the total IT spending. The independent variable is *Breach Peer*, which equals one if a firm belongs to the four-digit SIC industry where another firm in the same four-digit SIC industry experiences a hacking attack in the previous 12 months and zero otherwise. *Interlock board* equals one if a director of a peer firm also serves on the board of the breached firm in the year when a hacking event occurs. I use data from Institutional Shareholder Services (ISS)-Directors to construct *Interlock board*. I provide all variable definitions in Appendix Table A1. I winsorize all continuous variables at the 1st/99th percentiles. I include firm, year, state-by-year, and industry (two-digit SIC code)-by-year fixed effects as indicated in different columns. I report the *t*-statistics in parentheses with robust standard errors clustered at four-digit SIC code industry level. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

Y=	*Ln (1+Total IT spending)*					
	(1)	(2)	(3)	(4)	(5)	(6)
Breach peer (β_1)	0.158	0.085	0.133	0.174**	0.122	0.138
	(1.59)	(1.13)	(1.59)	(1.99)	(1.53)	(1.59)
*Breach peer * Interlock board* (β_2)	0.157**	0.169**	0.122*	0.141*	0.137**	0.112*
	(1.96)	(2.45)	(1.90)	(1.81)	(1.96)	(1.73)
Interlock board	-0.013	0.020	0.003	-0.014	0.013	0.008
	(-0.23)	(0.44)	(0.07)	(-0.24)	(0.27)	(0.16)
Size				0.202***	0.164**	0.164**
				(3.10)	(2.55)	(2.42)
Leverage				-0.023	0.300	0.206
				(-0.08)	(1.32)	(0.85)
M/B				0.004	0.005	0.005
				(0.62)	(0.74)	(0.79)
ROA				-0.197	-0.052	-0.045
				(-0.96)	(-0.25)	(-0.20)
Institutional block ownership				-0.013	-0.143	-0.170
				(-0.09)	(-1.14)	(-1.31)
Intangibility				-1.712***	-0.958**	-1.053***
				(-3.86)	(-2.50)	(-2.78)
Cash holdings				0.556**	0.453*	0.390*
				(2.45)	(1.96)	(1.73)
Observations	11,833	11,833	11,833	10,894	10,894	10,894
Adj. R-squared	0.974	0.980	0.980	0.976	0.981	0.981
$\beta_1+\beta_2$	0.315	0.254	0.255	0.315	0.259	0.250
p-value ($\beta_1+\beta_2=0$)	0.000	0.000	0.000	0.000	0.000	0.000
Firm FE	YES	YES	YES	YES	YES	YES
Year FE	YES			YES		
State-Year FE		YES	YES		YES	YES
Industry-Year FE			YES			YES

Table 11 Heterogeneities in the effects of hacking events on peer firms' IT investments: The role of exposures

This table reports the triple difference estimates of the effect of hacking events on peer firms' IT investments conditional on hacking event, firm, and industry heterogeneities. The independent variable is *Breach Peer*, which equals one if a firm belongs to the four-digit SIC industry where another firm in the same four-digit SIC industry experiences a hacking attack in the previous 12 months and zero otherwise. *High record loss* equals one if the loss of record from a hacking event is in the top tercile of record loss among 129 hacking events included in our analysis and zero otherwise. For control groups, I set the value to zero. *Big firm* equals one if a firm's size is above the top tercile in a given year and zero otherwise. *High valuation* equals one if a firm's Tobin's Q is above the top tercile in a given year and zero otherwise. *High ROA* equals one if a firm's return on assets is above the top tercile in a given year and zero otherwise. *Less constrained* equals one if a firm's Whited and Wu (2006)'s index is below the bottom tercile in a given year and zero otherwise. *High advertising* equals one if a firm's advertising expenses scaled by total assets is above the top tercile in a given year and zero otherwise. For *Big firm*, *High valuations*, *High ROA*, *Less constrained*, and *High advertising*, I use the one-year lagged values in the analysis. *High user data IND* equals one if a firm is in the industries that have a higher amount of customer data, and such industries include wholesale and retail (one-digit SIC code equals 1), financials (one-digit SIC code equals 6), hotels (two-digit SIC code equals 70), and communications (two-digit SIC code equals 48). The standalone *High user data IND* is omitted due to the presence of firm FE. I provide all variable definitions in Appendix Table A1. I winsorize all continuous variables at the 1st/99th percentiles. I report the *t*-statistics in parentheses with robust standard errors clustered at four-digit SIC code industry level. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

Y=	*Ln (1+Total IT spending)*						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Breach peer (β_1)	0.121**	0.155**	0.208**	0.215**	0.249***	0.217**	0.150***
	(2.56)	(2.08)	(2.48)	(2.53)	(3.22)	(2.43)	(2.64)
*Breach peer * High record loss (β_2)*	0.203**						
	(2.36)						
*Breach peer * Big firm(β_2)*		0.263***					
		(5.57)					
*Breach peer * High valuation (β_2)*			0.130***				
			(3.93)				
*Breach peer * High ROA (β_2)*				0.104***			
				(2.83)			
*Breach peer * Less constrained (β_2)*					0.257**		
					(2.43)		
*Breach peer * High advertising (β_2)*						0.117**	
						(2.03)	
*Breach peer * High user data IND (β_2)*							0.475**
							(2.11)
High record loss	0.031						
	(0.41)						
Big firm		0.149***					
		(2.85)					
High valuations			-0.080***				
			(-3.31)				
High ROA				-0.033			
				(-1.60)			
Less constrained					-0.241***		
					(-7.35)		
High advertising						-0.041	
						(-1.27)	
Observations	25,159	21,109	21,109	21,109	21,109	21,109	25,159
Adj. R-squared	0.833	0.856	0.856	0.856	0.857	0.839	0.832
β_1+ β_2	0.324	0.418	0.338	0.319	0.506	0.334	0.625
p-value (β_1+ β_2=0)	0.000	0.000	0.000	0.000	0.000		0.000
Firm FE	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES

Table 12 Heterogeneities in the effects of hacking events on peer firms' IT investments: The role of risk management awareness

This table reports the triple difference estimates of the effect of hacking events on peer firms' IT investments conditional on whether firms have a designated risk officer or a board risk committee overseeing risk management or not and conditional on whether firms discuss cyber risk in 10-K filings before a hacking event. The dependent variable is *Total IT spending*, defined as the logarithm of one plus the total IT spending. The independent variable is *Breach Peer*, which equals one if a firm belongs to the four-digit SIC industry where another firm in the same four-digit SIC industry experiences a hacking attack in the previous 12 months and zero otherwise. *Risk position* equals one if a firm has an executive with a risk title or has a risk committee. Other titles include chief enterprise risk officer, chief risk & compliance officer, chief risk management officer, chief operations credit & risk officer, vice president of risk management, vice president of risk and investments, etc. *Cyber risk discussion* equals one if cyber risk word appears at least once in firms 10-k filings in a given year. I use data from Execucomp to construct *Risk position*. For *Cyber risk* and *Risk position*, I use the one-year lagged values in the analysis. I provide all variable definitions in Appendix Table A1. I winsorize all continuous variables at the 1st/99th percentiles. I report the *t*-statistics in parentheses with robust standard errors clustered at four-digit SIC code industry level. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

Y=	*Ln (1+Total IT spending)*			
	(1)	(2)	(3)	(4)
Breach peer (β_1)	0.178***	0.167**	0.241***	0.239***
	(2.66)	(2.51)	(3.05)	(3.22)
*Breach peer * Risk position (β_2)*	0.207**	0.238**		
	(2.23)	(2.34)		
Risk position	0.173*	0.179		
	(1.93)	(1.50)		
*Breach peer * Cyber risk discussion (β_2)*			0.088**	0.066*
			(2.46)	(1.91)
Cyber risk discussion			0.129***	0.097***
			(3.61)	(2.60)
Size		0.227***		0.215***
		(3.22)		(4.12)
Leverage		-0.177		-0.130
		(-0.76)		(-0.79)
M/B		0.003		-0.003
		(0.58)		(-0.83)
ROA		-0.507***		-0.400**
		(-2.80)		(-2.53)
Institutional block ownership		0.128		0.213***
		(0.97)		(3.02)
Intangibility		-0.973**		-0.585**
		(-2.56)		(-2.12)
Cash holdings		-0.060		-0.020
		(-0.29)		(-0.11)
Observations	14,905	12,977	21,109	18,345
Adj. R-squared	0.829	0.840	0.856	0.862
$\beta_1+\beta_2$	0.385	0.454	0.217	0.163
p-value ($\beta_1+\beta_2=0$)	0.000	0.000	0.000	0.000
Firm FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

Table 13 Moderating effects of SEC cybersecurity disclosure guidance in 2011

This table reports the triple difference estimates of the effect of hacking events on peer firms' IT investments conditional on the release of SEC guidance on cyber security disclosure in 2011. The dependent variable is *Total IT spending*, defined as the logarithm of one plus the total IT spending. The independent variable is *Breach Peer*, which equals one if a firm belongs to the four-digit SIC industry where another firm in the same four-digit SIC industry experiences a hacking attack in the previous 12 months and zero otherwise. *SEC guidance 2011* equals one if after year 2011 and zero otherwise. I provide all variable definitions in Appendix Table A1. I winsorize all continuous variables at the 1st/99th percentiles. I include firm, state-by-year, and industry (two-digit SIC code)-by-year fixed effects as indicated in different columns. I report the *t*-statistics in parentheses with robust standard errors clustered at four-digit SIC code industry level. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

Y=	*Ln (1+Total IT spending)*			
	(1)	(2)	(3)	(4)
Breach peer (β_1)	-0.066	-0.053	-0.133	-0.099
	(-0.87)	(-0.69)	(-1.62)	(-1.25)
*Breach peer * SEC guidance 2011 (β_2)*	0.171**	0.155*	0.241**	0.205**
	(2.06)	(1.87)	(2.55)	(2.24)
Size			0.291***	0.284***
			(6.52)	(6.51)
Leverage			-0.064	-0.107
			(-0.49)	(-0.82)
M/B			0.001	0.001
			(0.36)	(0.42)
ROA			-0.249**	-0.243**
			(-2.24)	(-2.16)
Institutional block ownership			0.128**	0.114**
			(2.23)	(1.98)
Intangibility			-0.218	-0.187
			(-0.99)	(-0.83)
Cash holdings			-0.151	-0.170
			(-1.01)	(-1.17)
Observations	25,159	25,159	21,014	21,014
Adj. R-squared	0.863	0.863	0.888	0.888
$\beta_1+\beta_2$	0.105	0.102	0.108	0.106
p-value ($\beta_1+\beta_2=0$)	0.004	0.004	0.003	0.004
Firm FE	YES	YES	YES	YES
State-Year FE	YES	YES	YES	YES
Industry-Year FE		YES		YES

Appendix A. Examples of costs for data breaches

1. *"Equifax expects to pay out another $100 million for data breach - States it believes this will be the last of the payouts" (Housing Wire, Feb.14, 2020)*
https://www.housingwire.com/articles/equifax-expects-to-pay-out-another-100-million-for-data-breach/

In 2019, Equifax agreed to pay out nearly $700 million to settle numerous federal and state investigations. The company set aside another $99.6 million in its fourth-quarter earnings report for "certain legal proceedings and government investigations related to the 2017 cybersecurity incident." Beyond that, the company spent an additional $337 million in 2019 on technology and data security, legal and investigative fees, and product liability for the breach. In total, the breach cost Equifax $1.14 billion in 2019 alone.

"Overall, the data breach in 2017 has cost Equifax more than $1.7 billion since it was first disclosed in 2017. The company also cautions that despite its current belief that this $100 million will cover all its "remaining liabilities," it is possible that its financial punishment is not over yet."

2. "ICO fines Marriott 18.4 million pounds for failing to secure customer data" (Reuters, Oct.30, 2020)

https://www.reuters.com/article/us-marriott-intnl-ico/ico-fines-marriott-18-4-million-pounds-for-failing-to-secure-customer-data-idUSKBN27F1LH

U.S. hotel group Marriott International was fined 18.4 million pounds ($23.98 million) by the UK's data watchdog, The Information Commissioner's Office (ICO) at the end of October in 2020, for its data breach that began in 2014 but got public until 2018. In the second quarter of 2019 alone, Marriott has incurred a charge of $126 million for its data breach in 2018. The company is also facing a London class action by millions of former guests demanding compensation.

3. "*Yahoo reaches $117.5 million settlement in huge data breach*" (CBS News, Apr.10, 2019)

https://www.cbsnews.com/news/yahoo-data-breach-117-5-million-settlement-reached/

Yahoo has spent 152.5 million dollars in lawsuit settlement with Yahoo users and fines to regulators due to its huge data breach in 2013 and 2014. The amount includes the $35 million paid for regulatory fine to resolve federal regulators' charges that the online pioneer deceived investors by failing to disclose those breaches in April 2018.

"In April 2018, Yahoo agree to pay a $35 million fine to resolve federal regulators' charges that the online pioneer deceived investors by failing to disclose those breaches, among the biggest in internet history. The SEC alleged that although Yahoo senior managers and attorneys were told about the breach, the company failed to fully investigate it. The agency said the breach wasn't disclosed to investors until more than two years later, when Yahoo was working on closing Verizon's acquisition of its operating business in 2016."

Appendix B. Anecdotal Evidence of Peers' Reaction to Data Breach in the Same Industry

1. Walmart's responses to Target's data breach in 2013

In December 2013, the Target experienced the largest data breach in retail history, credit/debit card and contact information involved more than 110 million Target consumers has been leaked. On April 28 at the InformationWeek Conference, the chief information officer of Walmart, one of the biggest competitors of Target, Karenann Terrell, presented a keynote speech and focused on how to organize IT to deliver on business objectives in the post-Target breach era, "What Target taught the entire industry was that you can't have any single point of failure…What we learned is we have to have white-hat testing capability on staff for continual testing". She also mentioned that although the Target breach was not caused by a malicious insider, malicious insiders are extremely difficult to identify, and data analytics can play a big role. (*Information Week*, April 29, 2015, see https://www.informationweek.com/strategic-cio/executive-insights-and-innovation/wal-mart-cio-karenann-terrell-data-analysis-key-to-customer-insights/d/d-id/1320198)

In Walmart's 2014 annual report, Walmart discloses a brand-new section called "Data and Privacy Risks", and specifically discusses data breach risk the company is facing and the significant cost when a cyber-attack happens.

"Any failure to maintain the security of the information relating to our company, customers, associates and vendors that we hold, whether as a result of cybersecurity attacks on our information systems or otherwise, could damage our reputation with customers, associates, vendors and others, could cause us to incur substantial additional costs and to become subject to litigation, and could materially adversely affect our operating results … financial condition and liquidity, could result in the release to the public of confidential information about our operations and financial condition and performance and could result in litigation against us … Moreover, a security breach could require us to devote significant management resources to address the problems created by the security breach and to expend significant additional resources to upgrade further the security measures that we employ to guard such important personal information against cyberattacks and other attempts to access such information and could result in a disruption of our operations, particularly our digital retail operations."

(Extracted from Data and Privacy Risks Section, Walmart 2014 annual report)

2. TransUnion response to Equifax's data breach in 2017

On September 7, 2017, one of the biggest credit rating companies, Equifax, revealed that months-long illegitimate access to its credit-report databases had led to the breach of personally identifiable information of over 143 million people, nearly all in the U.S. The total number grew through March 2018 to over 148 million affected. The company waited six weeks to disclose the breach. Records varyingly included credit-card, driver's license, and Social Security numbers, date of birth, phone numbers, and email addresses. The major

competitor of Equifax, TransUnion, discusses the reasons and possible consequences of the event in its "Risk Factor" section in 2018 annual report. Below is an excerpt.

"While recent, highly publicized cybersecurity incidents, including the data incident announced by Equifax on September 7, 2017, have heightened consumer awareness of cybersecurity risks, they have also emboldened individuals or groups to target our systems even more aggressively. The preventive actions we take to address cybersecurity risk, including protection of our systems and networks, may be insufficient to repel or mitigate the effects of cyberattacks in the future as it may not always be possible to anticipate, detect or recognize threats to our systems, or to implement effective preventive measures against all cybersecurity risks."

"The extent of a particular cybersecurity incident and the steps that we may need to take to investigate it may not be immediately clear, and it may take a significant amount of time before such an investigation can be completed and full and reliable information about the incident is known. While such an investigation is ongoing, we may not necessarily know the extent of the harm or how best to remediate it, and certain errors or actions could be repeated or compounded before they are discovered and remediated, any or all of which could further increase the costs and consequences of a cybersecurity incident."

(Extracted from the Risk Factor, TransUnion 2018 annual report)

Appendix C. Examples of failures of shareholder derivative suits

1. "*Judge OK's Target Breach Settlement"* (Mathew J. Schwartz, March 19, 2015)

https://www.bankinfosecurity.com/target-eyes-10m-breach-settlement-a-8031

Due to its massive data breach in 2013, The Target has settled a consumer class-action lawsuit by paying $10 million to victims, as well as to reimburse plaintiffs' attorneys' fees and expenses up to $6.75 million Target has also agreed on a $19 million settlement with MasterCard on behalf of affected card issuers.

2. *"Shareholder Derivative Suit Following Data Breach Misses Target."* (James Thompson, July 15, 2016)

https://blogs.orrick.com/securities-litigation/2016/07/15/shareholder-derivative-suit-following-data-breach-misses-target/

Shareholder brought derivative suits listing 14 individual officers and directors as defendants alleging that "Target's (1) failed to properly provide for and oversee an information security program; and (2) failed to give customers prompt and accurate information in disclosing the breach. Based on these allegations, plaintiffs asserted claims for breach of fiduciary duty, gross mismanagement, waste of corporate assets and abuse of control." Target's board rejected the lawsuit demand, and motions to dismiss is granted by Judge Paul A. Magnuson of the United States District Court for the District of Minnesota on July 7, 2016.

Appendix Table A1 Variable definitions

Variables	Definitions
Breach peer	One if a firm belongs to the four-digit SIC industry where another firm in the same four-digit SIC industry experiences a hacking attack in the previous 12 months and zero otherwise.
Total IT spending	The sum of hardware spending, software spending, communication service spending, and service spending.
Growth rate in total IT spending	(Total IT spending- lagged Total IT spending) /lagged Total IT spending
Hardware spending	The spending for hardware that IT departments buy and support, regardless of whether the IT department itself operates that equipment (such as servers) or oversees the use of this equipment by employees (such as PCs). Hardware includes PCs (laptops, desktops, and tablets); servers/mainframes, peripherals: monitors, terminals, printers, keyboards, mice, USB devices, etc.; other hardware specific to the industry, such as point-of-sales equipment based on PCs, smart cards, embedded computer chips, etc.; storage devices.
Software spending	The spending for software, which is defined as software from third parties, whether that software is packaged or semi-packaged software delivered on CD and installed within the company, hosted by a third party, offered on a SaaS basis from a multitenant shared-instance server accessible by a browser, or custom-created for a company by third-party contractors or consultants. Software includes all categories of middleware software, such as license and subscriptions; application software, such as risk and payment management software; vertical industry applications, such as banking management systems and security trading systems; computer operating systems software.
Communication service spending	The spending for communication service, which is defined as the network equipment that companies operate to support their communications needs. Communication service includes routers, carrier line equipment, fiber optic equipment, switches, private branch exchanges (PBXes), radio and TV transmitters, Wi-Fi transmitters, desktop telephone sets; wide-area network (WAN) and local-area network (LAN) equipment, videoconferencing and telepresence equipment, cable boxes, other network equipment, end-client mobile devices such as cell phones/iPhones that are bought by individuals.
Service spending	The spending for service, which is defined as project-based consulting or systems integration services that vendors provide to businesses and Governments, whether on or off-site. Service includes contractors, consulting services for IT strategy, security assessments and process change; systems integration; project services; mainframe outsourcing, desktop support outsourcing, distributed systems outsourcing, network outsourcing, application hosting, application management outsourcing and application testing; computer hardware support and maintenance

	services.
Size	Natural logarithm of total assets (*at*).
Leverage	(Long-term debt (*dltt*) + debt in current liabilities(*dlc*))/ total assets(*at*).
Tobin's Q	(Total assets (*at*) – common/ordinary equity (*ceq*) + market value of equity (*prcc_f* × *csho*))/ total assets (*at*)
M/B	market value of equity (*prcc_f* × *csho*)/ total equities (*ceq)*.
ROA	Net income(*ni)*/ total assets (*at*).
Institutional block ownership	Number of shares held by institutional shareholders that own more than 5%
	of a firm's equity scaled by the total number of shares outstanding.
Intangibility	1 – total property, plant, and equipment (*ppent*)/ total assets (*at*)
Cash holdings	Cash and cash equivalents(*che*)/ total assets (*at*)
Interlock board	One if a director of the peer firm also serves on the board of breached firms.
Risk position	One if a firm has an executive with a risk title or has a risk committee and zero otherwise.
Cyber risk word	One if cyber risk word from the word list appears at least once in firms 10-k filings in a given year. The list of keywords includes "data?breach*", "cyber?attack*", "cyber?security*", "cyber?terrorism (terrorists)*", "cyber?threat*", "data?risk*", "data?security*", "hack*", "malware*", "privacy?risk*", and "spyware*".
Cyber risk disclosure	Number of cyber risk word from the word list appears in firms 10-k filings in a given year. The list of keywords includes "data?breach*", "cyber?attack*", "cyber?security*", "cyber?terrorism (terrorists)*", "cyber?threat*", "data?risk*", "data?security*", "hack*", "malware*", "privacy?risk*", and "spyware*".

Appendix Table A2 Effects of hacking events on hacked firms' IT investment

This table reports the change in the level of IT spending for hacked firms around a hacking event in our sample over the window (-3, 3) excluding the event year (year 0). For each hacked firm, I take the difference between the post-hacking-period mean (median) IT spending and the pre-hacking-period mean (median) IT spending, and then test whether the mean (median) change across different hacked firms is statistically significant or not. The significance is based on the *p*-values for *t*-tests and *z*-values for Wilcoxon signed-rank tests that the mean and the median differences of variables between pre- and post-hacking period equal zero. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively. Firms that suffered a data breach on average increases their IT investment by about 18.5% in the three-year window around the data breach.

	Pre-hacking period (3 years)		Post-hacking period (3 years)		Test of difference	
	Mean	Median	Mean	Median	Mean	Median
Ln (1+ Total IT spending)	14.676	15.958	17.390	17.182	2.714***	1.224***
					(7.16)	(6.14)

Appendix Table A3 Robustness of DID analysis of hacking events on peer firms' IT investments: Using an alternative control group

This table reports the results of robustness for difference-in-differences (DID) analysis of hacking events on peer firms' IT investments, using an alternative control group. Specifically, I define control group as firms in the same three-digit, but in a different four-digit SIC industry that have never encountered a hacking event over our sample period. The independent variable is *Breach Peer*, which equals one if a firm belongs to the four-digit SIC industry where another firm in the same four-digit SIC industry experiences a hacking attack in the previous 12 months and zero otherwise. The dependent variable is the natural logarithm of one plus *Total IT spending*. I provide all variable definitions in Appendix Table A1. I provide all variable definitions in Appendix Table A1. I winsorize all continuous variables at the 1st/99th percentiles. I include firm, year, state-by-year, and industry (two-digit SIC code)-by-year fixed effects as indicated in different columns. I report the *t*-statistics in parentheses with robust standard errors clustered at four-digit SIC code industry level. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

Y=	*Ln (1+Total IT spending)*					
	(1)	(2)	(3)	(4)	(5)	(6)
Breach peer	0.167***	0.168***	0.103*	0.152***	0.158***	0.146*
	(3.50)	(3.07)	(1.73)	(3.46)	(3.26)	(1.75)
Size				0.318***	0.298***	0.364***
				(4.50)	(5.45)	(6.65)
Leverage				-0.086	-0.106	0.079
				(-0.28)	(-0.35)	(0.26)
M/B				-0.003	-0.000	0.003
				(-0.46)	(-0.03)	(0.56)
ROA				-0.266	-0.205	-0.158
				(-1.28)	(-1.05)	(-0.93)
Institutional block ownership				0.247**	0.221**	0.147*
				(2.58)	(2.36)	(1.94)
Intangibility				-0.086	0.005	-0.057
				(-0.16)	(0.01)	(-0.12)
Cash holdings				-0.005	-0.053	-0.165
				(-0.02)	(-0.29)	(-0.97)
Observations	9,652	9,652	9,652	8,394	8,394	8,394
Adj. R-squared	0.842	0.845	0.869	0.855	0.857	0.880
Firm FE	YES	YES	YES	YES	YES	YES
Year FE	YES			YES		
State-Year FE		YES	YES		YES	YES
Industry-Year FE			YES			YES

Appendix Table A4 Robustness of DID analysis of hacking events on peer firms' IT investments: Excluding the financial crisis period

This table reports the results of robustness for difference-in-differences (DID) analysis of hacking events on peer firms' IT investments, excluding the financial crisis period (i.e., year 2008 and year 2009). The independent variable is *Breach Peer*, which equals one if a firm belongs to the four-digit SIC industry where another firm in the same four-digit SIC industry experiences a hacking attack in the previous 12 months and zero otherwise. The dependent variable is the natural logarithm of one plus *Total IT spending*. I provide all variable definitions in Appendix Table A1. I winsorize all continuous variables at the 1st/99th percentiles. I include firm, year, state-by-year, and industry (two-digit SIC code)-by-year fixed effects as indicated in different columns. I report the *t*-statistics in parentheses with robust standard errors clustered at four-digit SIC code industry level. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

Y=	*Ln (1+Total IT spending)*					
	(1)	(2)	(3)	(4)	(5)	(6)
Breach peer	0.151***	0.152***	0.065*	0.158***	0.158***	0.089**
	(3.83)	(3.71)	(1.79)	(3.93)	(3.83)	(1.99)
Size				0.274***	0.271***	0.320***
				(5.03)	(5.14)	(6.24)
Leverage				-0.299	-0.351*	-0.154
				(-1.43)	(-1.74)	(-1.06)
M/B				0.002	0.001	0.004
				(0.41)	(0.34)	(1.05)
ROA				-0.359**	-0.340**	-0.179
				(-2.14)	(-2.03)	(-1.32)
Institutional block ownership				0.204***	0.176**	0.074
				(2.99)	(2.48)	(1.24)
Intangibility				-0.344	-0.272	-0.279
				(-1.26)	(-0.99)	(-1.13)
Cash holdings				0.088	0.083	-0.032
				(0.47)	(0.45)	(-0.19)
Observations	22,714	22,714	22,714	19,623	19,623	19,623
Adj. R-squared	0.837	0.837	0.867	0.844	0.845	0.874
Firm FE	YES	YES	YES	YES	YES	YES
Year FE	YES			YES		
State-Year FE		YES	YES		YES	YES
Industry-Year FE			YES			YES

Appendix Table A5 DID analysis of hacking events on peer firm's cyber risk disclosure

This table reports the results firm difference-in-differences (DID) analysis of hacking events on peer firms' IT investments. The dependent variable is *Cyber risk disclosure*, defined as the total word count of cyber risk related keywords from our data library searched in firms' 10-Ks. The independent variable is *Breach Peer*, which equals one if a firm belongs to the four-digit SIC industry where another firm in the same four-digit SIC industry experiences a hacking attack in the previous 12 months and zero otherwise. I provide all variable definitions in Appendix Table A1. I winsorize all continuous variables at the 1st/99th percentiles. I report the *t*-statistics in parentheses with robust standard errors clustered at four-digit SIC code industry level. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

Y=	*Ln (1+Cyber risk word)*	
	(1)	(2)
Breach peer	0.061***	0.052**
	(3.51)	(2.57)
Size		0.044
		(1.38)
Leverage		0.209**
		(2.08)
M/B		0.002
		(0.74)
ROA		0.075
		(0.89)
Institutional block ownership		0.026
		(0.70)
Intangibility		0.031
		(0.23)
Cash holdings		0.035
		(0.65)
Observations	25,283	21,677
Adj. R-squared	0.619	0.625
Year FE	YES	YES
Firm FE	YES	YES

Appendix Table A6 DID analysis of hacking events on peer firms' financial decisions

This table reports the difference-in-differences (DID) estimates of the effect of hacking events on peer firms' financial decisions. The independent variable is *Breach Peer*, which equals one if a firm belongs to the four-digit SIC industry where another firm in the same four-digit SIC industry experiences a hacking attack in the previous 12 months and zero otherwise. *Acquiring* equals acquiring expenditures scaled by lagged total assets. *CAPX* equals capital expenditures minus IT investment, scaled by lagged total assets. *Repurchase* equals purchase of common and preferred stock scaled by lagged total assets. *R&D* equals R&D expenditures scaled by lagged total assets. *Advertising* equals advertising expenditures scaled by lagged total assets. I provide all variable definitions in Appendix Table A1. I winsorize all continuous variables at the 1st/99th percentiles. I report the *t*-statistics in parentheses with robust standard errors clustered at four-digit SIC code industry level. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

Y=	*Acquiring*	*CAPX-IT*	*Repurchase*	*R&D*	*Advertising*
	(1)	(2)	(3)	(4)	(5)
Breach peer	0.005**	-0.003**	-0.002	0.001*	-0.000
	(2.25)	(-2.50)	(-1.39)	(1.83)	(-0.35)
Size	0.042***	0.017***	-0.008***	-0.005***	-0.002*
	(8.27)	(8.99)	(-3.82)	(-3.15)	(-1.94)
Leverage	0.116***	-0.024***	0.014	0.001	-0.008***
	(7.98)	(-3.50)	(1.56)	(0.35)	(-4.11)
M/B	0.000	-0.000***	-0.000	-0.000	-0.000
	(1.11)	(-7.74)	(-0.99)	(-0.98)	(-0.60)
ROA	0.000*	-0.000	0.000	0.000	0.008***
	(1.68)	(-0.19)	(0.54)	(0.59)	(2.84)
Institutional block ownership	-0.008	-0.010***	-0.005*	-0.002	-0.000
	(-1.29)	(-2.96)	(-1.91)	(-1.46)	(-0.02)
Intangibility	0.169***	-0.123***	-0.009	-0.002	-0.014**
	(6.93)	(-10.40)	(-0.83)	(-0.54)	(-2.35)
Cash holdings	-0.188***	-0.013**	0.002	-0.014***	-0.007*
	(-11.58)	(-2.31)	(0.19)	(-3.64)	(-1.70)
Observations	17,723	15,568	17,616	18,184	8,633
Adj. R-squared	0.228	0.729	0.535	0.949	0.943
Firm FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES