european corporate governance institute

Hacking Corporate Reputations

Finance Working Paper N° 948/2024
January 2024

Pat Akey
INSEAD, University of Toronto and ECGI
Stefan Lewellen
Pennsylvania State University and ECGI
Inessa Liskovich
AirBnB
Christoph M. Schiller
Arizona State University

© Pat Akey, Stefan Lewellen, Inessa Liskovich and Christoph M. Schiller 2024. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

This paper can be downloaded without charge from:
https://doi.org/10.67929/ecgi-fin-948-2024

european corporate governance institute

ECGI Working Paper Series in Finance

Hacking Corporate Reputations

Working Paper N° 948/2024

January 2024

Pat Akey
Stefan Lewellen
Inessa Liskovich
Christoph M. Schiller

We thank Emilio Bisetti, Peter Cziraki, Alexander Dyck, Craig Doidge, Caroline Flammer, Xavier Giroud, Avi Goldfarb, Nico Inostroza, Steve Karolyi, Yrjo Koskinen, Souad Lajili Jarjir, Karl Lins, Kristina McElheran, Thomas Schmid, Duane Seppi, Henri Servaes, Laura Starks, Tracy Yue Wang, seminar participants at Alberta, Australian National University, Cambridge, Carnegie Mellon, Toronto, and UT Austin, and participants in the 2017 LBS Summer Symposium early ideas session, the 2018 NBER Summer Institute IT & Digitization poster session, the 2018 NFA Meetings, the 2018 Conference on CSR, the Economy, and Financial Markets, the 2019 AFA Meetings, the 2019 CUHK-RCFS Conference, Finance, Organizations, and Markets Virtual Seminar Series, and the 2019 UNPRI Conference for helpful comments. We also thank Hannah Ji, Avi Schiff, Zachary Stack, and Haruka Tatagi for excellent research assistance.

© Pat Akey, Stefan Lewellen, Inessa Liskovich and Christoph M. Schiller 2024. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Abstract

We exploit unexpected corporate data breaches to study the loss and repair of corporate reputation. Reputation loss decreases equity value and brand value, increases customer churn and prompts more negative media coverage. Firms repair their reputation by increasing their charitable donations (a novel measure of CSR investment), political contributions, employee wages and investment in IT. These actions are targeted to stakeholders that are particularly important or in situations that are particularly salient to their stakeholders. We observe similar dynamics of reputation loss and repair following the release of negative news about firms' social behaviors.

Keywords: Corporate Social Responsibility, CSR, Corporate Reputation, Cyberattack, Data Breach

JEL Classifications: G32, G31, G14, M14, D22, D23, L21, L25, L86, G38, G39

Pat Akey
Associate Professor of Finance
University of Toronto
105 St George Street
Toronto, Ontario M5S 3E6 M5S1S4, Canada
e-mail: pat.akey@rotman.utoronto.ca

Stefan Lewellen*
Assistant Professor of Finance
Pennsylvania State University
360 Business Building
State College, PA 16802, USA
e-mail: lewellen@psu.edu

Inessa Liskovich
Assistant Professor
University of Texas at Austin
1 University Station, B6600
Austin, TX 78712, USA
e-mail: inessa.liskoivch@mccombs.utexas.edu

Christoph M. Schiller
Assistant Professor
Arizona State University, W.P. Carey School of Business
300 E Lemon St,
Tempe, AZ 85287, USA
phone: +1 480-727-1981
e-mail: Christoph.Schiller@asu.edu

*Corresponding Author

Hacking Corporate Reputations*

Pat Akey† INSEAD/ Toronto

Stefan Lewellen‡ Penn State

Inessa Liskovich§ AirBnB

Christoph Schiller¶ Arizona State

December 31, 2023

Abstract

We exploit unexpected corporate data breaches to study the loss and repair of corporate reputation. Reputation loss decreases equity value and brand value, increases customer churn and prompts more negative media coverage. Firms repair their reputation by increasing their charitable donations (a novel measure of CSR investment), political contributions, employee wages and investment in IT. These actions are targeted to stakeholders that are particularly important or in situations that are particularly salient to their stakeholders. We observe similar dynamics of reputation loss and repair following the release of negative news about firms' social behaviors.

JEL Codes: G32, M14, L86

*We thank Emilio Bisetti, Peter Cziraki, Alexander Dyck, Craig Doidge, Caroline Flammer, Xavier Giroud, Avi Goldfarb, Nico Inostroza, Steve Karolyi, Yrjo Koskinen, Souad Lajili Jarjir, Karl Lins, Kristina McElheran, Thomas Schmid, Duane Seppi, Henri Servaes, Laura Starks, Tracy Yue Wang, seminar participants at Alberta, Australian National University, Cambridge, Carnegie Mellon, Toronto, and UT Austin, and participants in the 2017 LBS Summer Symposium early ideas session, the 2018 NBER Summer Institute IT & Digitization poster session, the 2018 NFA Meetings, the 2018 Conference on CSR, the Economy, and Financial Markets, the 2019 AFA Meetings, the 2019 CUHK-RCFS Conference, Finance, Organizations, and Markets Virtual Seminar Series, and the 2019 UN PRI Conference for helpful comments. We also thank Hannah Ji, Avi Schiff, Zachary Stack, and Haruka Tatagi for excellent research assistance.

†INSEAD and University of Toronto. Phone: +1 (647) 545-7800, Email: pat.akey@rotman.utoronto.ca

‡Pennsylvania State University. Phone: +1 (814) 867-5848, Email: lewellen@psu.edu

§AirBnB. Email: inessal@gmail.com

¶Arizona State University. Phone: +1 (480) 727-1981, Email: christoph.schiller@asu.edu

Electronic copy available at: https://ssrn.com/abstract=3143740

How do firms respond to the destruction of intangible capital? Recent research has highlighted the importance of intangible capital in production (e.g., Belo et al., 2014; Crouzet and Eberly, 2018; Corhay et al., 2020; Belo et al., 2022) and in the economy more generally (Corrado and Hulten, 2010), but aside from human capital (e.g., Jäger, 2016), there is little work on how firms respond to the destruction of intangible capital. As Noe (2012) writes in *The Oxford Handbook of Corporate Reputation*, "an interesting avenue for future research would be to extend the boundaries of the economic reputation model to encompass repair, both superficial and substantive."[1] In the same handbook, Karpoff (2012) lists "how and when do firms rebuild damaged reputations?" as one of six questions deserving further research attention in the area of corporate reputations, writing "other than ... anecdotes, we do not know whether firms tend to reinvest in reputation following a reputational loss, under what conditions they do so, and whether the reinvestment is successful."

Our paper aims to fill this gap in the literature. We ask three different questions related to the loss and potential repair of corporate reputations. First, which stakeholders respond to events that impair a firm's reputation? Second, which actions might a firm take to repair its reputation? Third, do firms tailor their responses in particular situations or to cater to particular stakeholders? We answer these questions in the context of two distinct empirical settings: data breaches and a broader set of reputation shocks from the data provider RepRisk.

We study data breaches because these events arguably affect corporate reputations while being plausibly unrelated to firms' product quality or financial condition. In addition, data breaches are largely idiosyncratic, 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. For example, it is hard to imagine that the 2014 data breach of employee records at Coca-Cola by a disgruntled employee would affect the taste or smell of Coca-Cola's products. Moreover, data breaches impact nearly all sectors of the economy and companies of every size and profile and executives frequently list firm reputation as one likely casualty of a data breach.[2] However, these events also represent

[1]While Rhee and Kim (2012) offer a behavioral model of reputation repair, formal models of reputation (e.g., Klein and Leffler, 1981; Kreps and Wilson, 1982; Milgrom and Roberts, 1982; Fudenberg and Levine, 1989; Maksimovic and Titman, 1991; Noe et al., 2012; Board and Meyer-ter-Vehn, 2013; Marinovic et al., 2018; Huang et al., 2020; Levine, 2021) do not model the reputation repair process.

[2]For example, a 2016 *Economist Intelligence Unit* survey found that C-level executives from 16 countries and various industries listed corporate reputation as the single most important company asset requiring

Electronic copy available at: https://ssrn.com/abstract=3143740

a very specific notion of lost reputation. In order to ensure that our results have external validity, we complement our analysis with 2,700 events where firms experience large, discrete increases in reputation risk. We find that stakeholders respond similarly to data breaches and RepRisk events and that firms generally take similar actions following both types of events to repair their reputations.

While there are many different definitions of reputation in the literature, we define reputation as the set of value-relevant firm characteristics that affect stakeholders' perceptions about the firm.[3] In the spirit of Hayek (1948), we argue that firms compete with other firms along many dimensions on the basis of their reputations, and thus, any value-relevant characteristic that firms compete over should be included when defining the firm's reputation. We also argue that firms can possess different reputations among different sets of stakeholders; for example, Amazon likely has different reputations with its employees and investors. Thus, understanding how stakeholders and firms respond to reputation-impairing events requires us to examine many different types of corporate stakeholders.

We begin by identifying a number of value-relevant stakeholders and study how they respond to reputation-impairing events. Building on work by Freeman (2015), we examine whether capital market participants, consumers, and traditional and social media respond to data breaches and other reputation-impairing events. We find that these events prompt a wide range of long and short term effects among stakeholders that would spur managers to attempt to rebuild their reputations. For example, we find negative equity returns on the order of 1–1.5% around both types of events, which prompt managers to discuss "reputation" and other CSR-related words more frequently in their conference calls with financial analysts, compared to peer firms that did not experience a reputation-impairing event. We find that consumers' perceptions of a firm or of its brands are impacted by data breaches—survey-measures of brand strength are lower following a data breach and "customer churn" (Baker et al., 2023) is higher following a RepRisk event. Moreover, we find that various measures of media coverage become more negative after firms experience a negative shock to their reputation. For example, we find that press coverage, in particular the coverage of local media outlets, becomes more negative following data breaches and that social media "buzz" around firms that suffer a RepRisk event is higher and substantially more negative.These results suggest

protection from cyberattacks. See http://tinyurl.com/59fwvu5c.

[3]This definition is similar but distinct to that of Noe (2012), who defines reputation as "an assessment of economic agents regarding the characteristics of a firm."

Electronic copy available at: https://ssrn.com/abstract=3143740

that a variety of stakeholders, including those whose attitudes are more difficult to measure such as local communities or the government, respond negatively to reputation-impairing events. To the extent that a "good" reputation enables firms to more easily contract with important stakeholders, lost reputation should increase the difficulty for a firm to operate efficiently, which should manifest in reduced values over the long term. Our final result in this section confirms this intuition. We find that firms' long-term valuations (i.e., market-to-book ratios) are lower for up to five years following a data breach or RepRisk event.

We answer our second question by identifying a set of actions and investments that firms may undertake to repair their reputations with various stakeholders. Intuitively, negative reputation shocks should reduce the value of a firm's intangible capital, thereby leaving the firm below its optimal level. Under the assumption of decreasing returns, a reduction in the value of the firm's existing stock of intangible capital should lead to an increase in the marginal benefit of additional investment. Since marginal costs are likely to remain unchanged, firms should therefore increase intangible investment following negative reputation events. While the universe of potential firm responses is very wide, we focus on five types of responses that are in line with some of the stakeholders typically highlighted by the literature: advertising expenditures, charitable contributions and CSR scores to repair reputation with consumers or local communities, political contributions to repair reputation with the government, increasing salaries to repair reputation with employees, and – specific to IT-related threats such as data breaches – investments in IT and cybersecurity.

We find that firms respond on many of these margins. As one might expect, firms seem to increase their investment in IT security following a data breach, but not following a RepRisk event. Firms are 12 percentage points more likely to discuss IT security in their annual reports. Firms also respond along a variety of margins that are not directly related to their operations. For example, firms' spending on charitable contributions is \$700,000–\$1.49 million per year higher relative to unaffected firms in the four years following the disclosure of a data breach and \$980,000–\$1.85 million per year following a RepRisk event, and they are up to 28 percentage points more likely to have a charitable foundation. Increased "investment" in socially responsible actions translates into increased CSR scores. We find that CSR scores are 0.3–0.65 standard deviations higher in the years following either type of reputation-impairing event. We find that political contributions are higher after both types of reputation shocks—specifically contributions of firms' Political Action Committees (PAC)

Electronic copy available at: https://ssrn.com/abstract=3143740

are $100,000–$250,000 higher in the election cycle following a data breach or RepRisk event. We find that firms increase annual employee wages following data breaches by $600–$1,200 (per employee) following data breaches but not following RepRisk events. However, firms do not seem to respond on *all* margins. Perhaps surprisingly, we find no evidence that they increase advertising expenditures following a reputation-impairing event.

In our final set of tests we examine our third question: do firms tailor their responses in particular situations or to cater to particular stakeholders? Intuitively, one or two categories of stakeholders might be particularly important for a firm's bottom line. In such a case firms should prioritize repairing reputation to these stakeholders or might prioritize reputation rebuilding activities following shocks that are particularly salient to a group of stakeholders. We first show that firms that are more consumer-facing increase their CSR-related investments more than firms that are not, consistent with theories of CSR as an instrument to differentiate firms in the product market (Albuquerque et al., 2019). Our next two tests examine political contributions. We show that firms with major government contracts increase their political contributions more than other types of firms that are politically active after both data breaches and RepRisk events, and that firms increase their political contributions more sharply after RepRisk events related to regulatory "violations" compared to, for example, social or product-related events. Recent work suggests that governments monitor the social responsibility of firms (Flammer, 2018; Gantchev et al., 2022) and our results suggest firms that depend on the government must respond particularly strongly to repair their reputation with this stakeholder. Our final test reexamines our finding that firms increase wages after a data breach. We exploit the fact that some data breaches affect customer records while others affect employee records and find that firms only increase employee wages after a data breach affects employee records, consistent with them targeting their response to repair their reputation to stakeholders that were affected.

Our results are robust to concerns regarding selection biases, omitted variables, and our choice of empirical specifications. For example, any narrative about selection on unobservables must explain why cyberattackers particularly choose firms that are predetermined to experience long-term value declines and long-term investments in CSR relative to their industries precisely at the same times (unanticipated by the markets) when the attacks take place, which we believe is highly unlikely. Our results also survive propensity-score matching and the coefficient stability tests proposed by Oster (2019). In addition, our main

Electronic copy available at: https://ssrn.com/abstract=3143740

specifications include firm fixed effects, industry-by-year fixed effects, and time-varying firm characteristics, which should absorb most latent factors. We also verify that our results are robust to differences-in-difference models that explicitly account for staggered timing in the treatment (see e.g., Sun and Abraham, 2021; Gardner, 2022). Controlling for the firm-level corporate governance or risk management factors examined in Kamiya et al. (2021) does not change any of our main results.

Our paper makes five primary contributions to the existing literature. First, an oft-considered theoretical motivation for CSR investment is that CSR guards against reputation risk, all else equal (Heal, 2005; Elfenbein et al., 2012; Kitzmueller and Shimshack, 2012; Albuquerque et al., 2019).[4] However, we are unaware of any empirical studies that specifically attempt to isolate corporate reputation as a motivation for firms' investment in CSR.[5] To our knowledge, this is the first to document direct investment in CSR as a response to a negative reputation shock. More generally, our paper is among the first to examine whether firms attempt to rebuild their stock of intangible capital following a negative shock.

Second, the idea that CSR can provide "insurance" against negative shocks has been the subject of many empirical studies (Godfrey et al., 2009; Hong and Kacperczyk, 2009; Vanhamme and Grobben, 2009; Flammer, 2013; Barrage et al., 2020; Hong and Liskovich, 2014; Lins et al., 2017; Albuquerque et al., 2020) and has drawn indirect theoretical interest (Albuquerque et al., 2019). However, these studies examine whether a firm's *existing* stock of CSR can help the firm when it experiences a negative shock, such as an oil spill, misconduct, or regulatory actions. In contrast, we exploit a negative shock to firms' reputations and examine whether firms subsequently *replenish* their stock of CSR in response to the event using a novel measure of CSR investment. Relative to the studies above, our paper also exploits a setting in which it is less likely that a negative reputation shock is contaminated with other fundamental news about the firm's current or future business prospects.

Third, the literature on reputation in economics has largely focused on games where firms can make unobservable investments to improve reputation through higher product quality (Klein and Leffler, 1981; Kreps and Wilson, 1982; Milgrom and Roberts, 1982; Fudenberg and Levine, 1989; Maksimovic and Titman, 1991; Holmstrom, 1999; Mailath and Samuelson, 2001;

[4] A large literature has also established a positive empirical link between financial performance and CSR (Margolis et al., 2009; Edmans, 2011; Flammer, 2015). Researchers have also introduced and investigated a number of theories for why firms may choose to engage in strategic CSR (see, e.g., Bénabou and Tirole, 2010).

[5] Servaes and Tamayo (2013) find that the benefits of CSR a concentrated among consumer-focused firms, but do not examine reputation.

Electronic copy available at: https://ssrn.com/abstract=3143740

Board and Meyer-ter-Vehn, 2013; Marinovic et al., 2018; Huang et al., 2020; Levine, 2021). Most of these theories differ primarily in their assumptions about information structure and timing. We instead show that firms also respond to reputation shocks *not* related to product quality by making *observable* investments to repair their reputations, thus offering new facts that can help guide future theoretical work. Our findings also emphasize that the firm's overall reputation is an aggregation of its (differing) reputations among distinct stakeholders groups, in a similar spirit to Tirole (1996). In addition, our paper specifically focuses on how firms *respond* to a plausibly exogenous shock to reputation, which is new to the literature.

Fourth, the existing empirical literature on corporate reputations has focused on negative reputation shocks such as financial misconduct (Armour et al., 2017; Karpoff et al., 2008; Murphy and Shrieves, 2009; Chakravarthy et al., 2014), environmental violations (Karpoff et al., 2005) and product recalls (Jarrell and Peltzman, 1985; Liu and Shankar, 2015; Dai et al., 2020). The evidence from these papers suggests that reputation losses are large and in some cases overshadow direct legal penalties. Our paper adds to this literature by examining reputation *repair* as opposed to documenting the magnitude of reputation losses. Our findings also inform theory and can help future researchers interested in examining models of reputation repair.

Finally, our paper contributes to the growing finance literature on corporate data breaches.[6] Florackis et al. (2023) develop text-based measures of cybersecurity risk and show that their measures predict cyberattacks and affect stock prices. We complement their paper by documenting the long-term negative firm value effects of data breaches and by providing evidence of firms' investments in intangible capital following such breaches. Makridis and Dean (2018) also show that firm sales, capital, and TFP fall following successful cyberattacks. We complement their paper by linking data breaches to firms' reputations and by focusing on whether firms invest in intangible capital following such events.

Two other data breach papers are closely related to ours. Lending et al. (2018) link large data breaches to future changes in firms' CSR scores, but they do not provide any analysis attempting to explain this link. In addition, they do not focus on reputation and do not measure firms' direct investment in intangible capital following data breaches. Second, using a distinct sample of large cyberattacks, Kamiya et al. (2021) examine the effects of

[6] In the computer science literature, Campbell et al. (2003), Acquisti et al. (2006), and Spanos and Angelis (2016) document significant negative short-term stock market reactions to corporate data breaches. Aldasoro et al. (2020) also analyze the characteristics of firms affected by cyberattacks (a common type of data breach).

Electronic copy available at: https://ssrn.com/abstract=3143740

cyberattacks on short-term stock prices, firm financial condition, risk management, and CEO turnover. They find that firms affected by cyberattacks subsequently increase their investment in risk management, which they theoretically attribute to firms responding to reputational losses. In contrast, we use a broader sample of data breaches to study a different research question (intangible investment following negative-reputation events) and complement our analysis with a broad set of reputation-impairing events. Unlike Kamiya et al. (2021), we present direct evidence that firm reputations suffer as a result of data breaches and provide direct evidence that firms take specific actions to repair their reputations, particularly by investing in intangible capital. Our paper complements Kamiya et al. (2021) by providing direct evidence of how reputation is lost and (potentially) repaired, which provide motivation for firms to invest in risk management, though we show that neither risk management nor governance changes are driving our results.

1 Reputation Investment and Repair

1.1 Defining reputation

Economists have long recognized that reputations can be valuable for firms. In *The Meaning of Competition*, Hayek (1948) notes that, contrary to theories of perfect competition, "[i]n actual life ... competition is in a large measure competition for reputation or good will." Over the years, the literature in economics has come to regard reputation as a measure of *contract commitment*: for example, as noted by Klein and Leffler (1981), "economists have long considered reputations or brand names to be private devices which provide incentives that assure contract performance in the absence of any third-party enforcer." Modern definitions largely follow this approach; in *The Oxford Handbook of Corporate Reputation*, Karpoff (2012) defines reputation as "the present value of the cash flows earned when an individual or firm eschews opportunism and performs as promised on explicit and implicit contracts."

Most definitions of reputation categorize reputation in terms of relationships with customers. For example, after defining reputation, Klein and Leffler (1981) write that the "private-contract enforcement mechanism relies upon the value to the firm of repeat sales to satisfied customers as a means of preventing nonperformance." More recently, Board and Meyer-ter-Vehn (2013) define reputation as "the market's belief about product quality" and directly model reputation as a function of customers' beliefs about quality. However, as noted by Bénabou and Tirole (2010), many other types of "stakeholders" (not just

Electronic copy available at: https://ssrn.com/abstract=3143740

customers) can have contractual or quasi-contractual relationships with firms, including, as defined by Freeman (2015) in the *Wiley Encyclopedia of Management*, "suppliers, customers, stockholders, employees, communities, political groups, governments, [and] media." It is natural to think that the concept of firm reputation may extend to these other stakeholders as well. As noted by Noe (2012), given that firm reputation is "an assessment of economic agents regarding the characteristics of a firm," it follows that "reputation is not a single property that a firm does or does not have; rather, the firm has a reputation with specific stakeholders regarding specific [firm] characteristics." For example, ExxonMobil may have a reputation for high-quality gasoline among consumers, but may carry a reputation for poor environmental performance among community groups. Thus, firms may have many different reputations among many different sets of stakeholders.

In this paper, we return to the spirit of the general definition put forth by Hayek (1948). We define reputation as the set of *value-relevant* firm characteristics that affect stakeholders' perceptions about the firm. Our view, following Hayek (1948), is that "stakeholders" here should reflect all groups of agents the firm may reasonably compete for on the basis of its reputation or goodwill. This arguably yields a quite broad set of reputation-relevant stakeholders. For example, since firms compete in labor markets for employees, employees are a relevant stakeholder. Since firms compete for investors, investors are also a relevant stakeholder. Since firms compete for government contracts, the government is a relevant stakeholder, and so on. One can make the case that firms have value-relevant reputations with most (if not all) of the stakeholder groups listed in Freeman (2015)'s definition.

1.2 Reputational investment

Implicit in Hayek (1948)'s (and our) definition of reputation is the idea that firms may need to *invest* in order to compete with other firms on the basis of reputation. Such investment may take many forms and may be *stakeholder-specific* – see, e.g., Board and Meyer-ter-Vehn (2013), where firms invest in improving product quality to affect customer reputation, or Lins et al. (2017), where firms invest in CSR as a form of insurance against unexpected negative future shocks.[7] Moreover, firms' individual reputational investments may affect their reputations with many stakeholders at once (in the spirit of Tirole, 1996).

Another key distinction between our definition and other definitions of reputation is

[7]Most papers about reputation in the economics literature do not specifically model firm investment. Exceptions include Board and Meyer-ter-Vehn (2013) (cited above) and Noe et al. (2012).

Electronic copy available at: https://ssrn.com/abstract=3143740

that our definition can apply to *firms* as a whole and not just to products or brands. For example, General Motors, Inc. (GM) sells automobiles through many different brands – each of which has its own distinct reputation with consumers – yet GM's reputation among union and non-union workers is largely driven by decisions made at the company level. In some cases a firm's "company-level" reputation may flow through to the firm's product reputations, while in other cases product-level reputations may spill over to affect the reputation of the company as a whole. For example, GM's environmental reputation will likely flow through to each of the company's brands, whereas a series of battery fires affecting one model of car may spill over to affect the reputations of other brands or even GM as a whole.

These types of hierarchies and interactions are largely ignored by the existing literature on reputation, yet they are arguably salient factors for firm decision-makers. Intuitively, firms can make at least three types of investment at the firm level or the brand level that directly or indirectly affect the firm's reputation. First, firms can invest in physical or related capital. Such product- or brand-level investments can allow the firm to improve its products or services, driving reputation gains and/or revenue/market share growth among various stakeholder groups (see, e.g., Board and Meyer-ter-Vehn, 2013). Second, the firm can invest in CSR activities to insure the firm against negative shocks (see, e.g., Lins et al., 2017). Anecdotal evidence suggests that this type of ex-ante reputation investment largely occurs at the firm level, though it is possible some firms invest in brand-level reputation "insurance" or invest ex-ante to insure the firm's reputation among specific stakeholders. Finally, the firm can invest in CSR activities *after* a negative shock occurs in order to rebuild the firm's stock of reputational capital. This third type of investment is the main focus of our paper.

Ex-post investment in reputational capital has two potential advantages. First, ex-post investment can target both specific stakeholders (e.g., community members) *and* can target the specific reputational damage the firm has suffered (e.g., a fine for environmental damages), which is extremely difficult to achieve through ex-ante investment in physical or reputational capital. Second, because ex-post investment by definition occurs later than ex-ante investment, the present value of the investment cost is lower – that is, the time value of money can also play a role in firms' reputational capital investment choices. These potential advantages suggest that ex-post investment in reputation repair may be optimal for at least some firms, a proposition that we consider more formally below.

Electronic copy available at: https://ssrn.com/abstract=3143740

1.3 Economic framework

To better motivate our focus on reputational repair, we construct a parsimonious model of firm investment in which a representative firm is subject to the possibility of a negative shock and can invest in physical capital, in (ex-ante) reputation-building activities, and in (ex-post) reputation repair following the realization of a negative shock. The model is static and features two periods. In the first period, the firm chooses how much to invest in physical capital, how much to invest in ex-ante reputation "insurance," and how much to invest in the following period in ex-post reputation repair activities (should a negative event occur), subject to a budget constraint, in line with Section 1.2. All firm decisions are made at time $t = 1$. Nature then determines whether the firm faces an exogenous reputation-destroying event and the firm responds by investing in reputation repair as decided previously. Payoffs are then realized at $t = 2$.

To keep the model parsimonious, we assume that a firm's revenue is a direct function of the flow investment in (physical) capital, which we denote K. Revenue also depends on a firm's reputation value, which we normalize to one. A negative reputation shock (for example, a cyberattack) happens in the second period with exogenous and known probability p. If such an event occurs, we assume that the firm's reputation value will be reduced to zero absent reputational investment.

Firms can respond to a potential negative reputation shock in two ways: by investing in a technology today that will reduce the size of the reputational loss in case the firm experiences a negative reputation shock in the second period, and by investing *after* a negative reputation shock has occurred to repair the firm's reputation. We model these investments very parsimoniously. We let I denote the firm's investment in ex-ante reputation insurance, and R denote the firm's investment in ex-post reputation repair. We assume that both technologies have linear benefits. For example, a firm investing I in reputation insurance will have a reputation value of I should the negative shock happen. We allow the efficiency of reputational repair to vary from the efficiency of the other two types of investment: investing R earns the firm a benefit of θR, where $\theta = 1$ would make the efficiency of investment in reputational repair equal to the efficiency of investment in physical capital or reputation insurance. Note that the firm only invests in R if the negative shock occurs.[8] We assume that

[8]If a negative shock does not occur, we assume that the firm is not required to invest the money it set aside in the second period (since doing so would yield no benefit); effectively, the money just goes unspent

Electronic copy available at: https://ssrn.com/abstract=3143740

the cost of investing in physical capital is K^2, the cost of investing in reputation insurance is I^2, and the cost of investing in reputation repair is θR^2. Second-period cash flows are discounted using the discount factor $\rho < 1$.

The firm's problem therefore has three components: how much to invest in physical capital (K), how much to invest in reputation insurance (I), and how much to potentially invest in ex-post reputation repair (R). All three decisions must be made in the first period and must satisfy the budget constraint that (WLOG) $K + I + \rho R \leq 1$. Putting everything together, the firm's objective as of $t = 1$ is:

$$\begin{aligned} \max_{K,I,R} E[\pi] &= \rho\left[K + ((1-p) + pI) + p\theta R\right] - K^2 - I^2 - \rho p\theta R^2, \\ \text{s. t.} \qquad & K + I + \rho R \leq 1 \ . \end{aligned}$$

Taking FOCs and solving yields the model's main result:

Proposition 1. *Suppose $0 < \rho < 1, 0 < p \leq 1$, and $\theta \geq 0$. Then equilibrium levels of K^*, I^*, and R^* are strictly positive.*

Proof. See appendix. ■

Proposition 1 states that even when reputation "insurance" is available and is efficient at protecting firms from harm, the firm will still optimally invest in ex-post reputation repair. The intuition for this result is best seen by examining the comparative statics of optimal investment to parameters such as the discount rate ρ and the efficiency parameter θ.

Proposition 2. *Suppose p is sufficiently large such that $\rho(2 + p) > 2$. Then:*

1. *Increasing the discount rate (i.e., lowering the discount factor) increases investment in ex-post repair (R^*).*
2. *Increasing the efficiency of ex-post investment leads to higher R^* and lower K^* and I^*.*
3. *Increasing the probability of a negative shock leads to higher I^* and lower K^* and R^*.*

Proof. See appendix. ■

This proposition highlights the main intuition behind the model. First, increasing the discount rate means investing in R^* will grow larger in the following period, so the firm will optimally increase investment in R^*. This helps explain why the firm does not invest solely in

and there is an opportunity cost associated with not spending it.

Electronic copy available at: https://ssrn.com/abstract=3143740

ex-ante reputation insurance. Second, increasing the efficiency of ex-post investment makes it more appealing to invest in R^* relative to the other investment types. For example, unlike ex-ante reputation insurance, firms can specifically target ex-post reputation repair to target the specific stakeholders whose reputation beliefs are most affected by the negative event. All else equal, this should make ex-post investment more efficient than ex-ante investment, leading to a higher θ and thus higher investment in reputation repair. Finally, a higher probability of a negative event leads to more investment in ex-ante insurance and lower ex-post investments (mild) and capital investments (stark). This occurs in the model because I is directly attached to p; it directly reduces the harm caused by an attack. As such, firms will respond to an increase in p by increasing I relative to the other two forms of investment.

Overall, the framework above suggests a role for both ex-ante and ex-post reputational investment within firms' overall investment schedules. The framework also highlights some of the potential benefits of ex-post reputational investments; namely, the time value of the money and firms' ability to target this type of investment to specific stakeholders.

1.4 Isolating reputation-related investments

While all three types of investments (K^*, I^*, R^*) can improve firms' stock of reputational capital, some of these investments – and in particular, K^* – may be undertaken for other reasons as well. For example, a firm investing in better product quality may primarily hope to increase demand and market share while also hoping to improve the product's long-run reputation. As such, it can be difficult to isolate investments undertaken for reputation-related purposes from investments undertaken for different economic rationales.

Empirically, this logic suggests that measures of firms' reputation investment should focus on investments that do not *directly* affect a firm's products or services. One example of such an investment is corporate charitable contributions – if Goldman Sachs donates money to the opera, this donation may well affect Goldman's reputation among consumers, employees, community members, and other stakeholders, but the donation should *not* directly affect any of the products or services offered by Goldman Sachs. Hence, one can plausibly infer that such charitable contributions should largely represent investments in reputation rather than investments in the firm's products or services.

Electronic copy available at: https://ssrn.com/abstract=3143740

2 Data

We obtain data from 18 different data sources. While many of these are commonly used in the literature, some are less common, and some are altogether novel. This section briefly describes each data source, with a focus on datasets that are less common to the literature. A detailed description of each data source is provided in Appendix IA.II. We provide summary statistics in Appendix Tables IA.1 and IA.2.

2.1 Corporate reputation shocks

We use two distinct types of corporate reputation shocks: corporate data breaches and a broader set of reputation shocks from the data provider RepRisk. We obtain data on data breaches at public corporations from the Privacy Rights Clearinghouse (PRC) website.[9] Our dataset begins in 1999 and ends in 2015. The data contains 287 data breaches, of which the vast majority involve customer records (66%) or employee records (33%).

We also obtain reputation shocks from RepRisk, a private company recording daily news events affecting corporate reputations across 28 distinct CSR and reputation risk issues such as air pollution, product controversies, discrimination, and labor practices, as well as broader scandals including violations of national legislation or international standards. RepRisk screens over 80,000 public sources in 20 languages, including print, online and social media, government bodies, regulators, think tanks, and newsletters to construct this data. RepRisk's analysts further classify each news item according to its novelty, severity, and reach. We retain only newly-reported, i.e., "novel", events with a high level of reach, credibility and influence – for example, issues that were reported by international news organizations. In total, we identify 2,700 negative reputation events for the firms in our sample spanning 2007 to 2015 (see, e.g., Dai et al., 2020 and Gantchev et al., 2020).

2.2 Corporate social responsibility

We obtain data on firms' CSR investment and CSR scores from a variety of sources. For CSR investment, we obtain data on corporate charitable contributions from Foundation Directory Online (FDO), supplemented with data from the Urban Institute's National Center for Charitable Statistics (NCCS) database.[10] The sample period for our donations data is

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

[10] Among others, Masulis and Reza (2015), and Cai et al. (2021) use FDO data. Bertrand et al. (2018), Ahn et al. (2020), and Bertrand et al. (2020) use data from NCCS.

Electronic copy available at: https://ssrn.com/abstract=3143740

2003 to 2014. By supplementing FDO data with NCCS data, we ensure the broadest and deepest possible coverage for our charitable contributions data.[11]

We measure CSR scores using the MSCI ESG KLD Stats ("KLD") index. A firm's KLD score is an index that equals the number of CSR "strengths" minus the number of CSR "concerns." Since KLD changed their computation methodology repeatedly over our sample period, we create a time-consistent KLD index, following Hong et al. (2019), as detailed in Appendix IA.II. For ease of interpretation, we normalize our final measure to have a mean of 0 and a standard deviation of 1 throughout the sample and refer to it as "Norm CSR." We also examine the robustness of our CSR findings using ESG scores from Asset4.

2.3 Conference call transcripts

We further construct measures to capture reputation-related content from quarterly corporate earnings call transcripts sourced from the data provider Streetevents from 2004 to 2014. For analyses conducted at the firm-year level, we collapse the quarterly data at the annual frequency by summing over the occurrence of individual words in our content dictionaries, which are described in detail in Appendix IA.II. We also distinguish between statements made during the management presentation and the Q&A section of each earnings call and construct indicator and count variables that capture whether issues related to reputation, CSR, data security, and data breaches are discussed.

2.4 Media sentiment

We obtain two measures of corporate media sentiment. First, we construct measures of national and regional news media sentiment using data from Ravenpack (RP) Edge for the period from 2000 to 2016. The RP sentiment measure is based on individual news item such as TV segments, radio features, blog posts, and newspaper articles from thousands of sources and is distributed between -1 and 1, indicating highly negative to highly positive news media sentiment. Our process for creating firm-year measure of media sentiment is described in more detail in Appendix IA.II.

Second, we obtain social media data based on Twitter (now known as X) content from the data provider Social Market Analytics (SMA) (now known as Context Analytics) at the firm-by-day frequency, including daily Twitter sentiment and 'buzz', which measures the cross-sectionally adjusted, abnormal Tweet volume. Sentiment scores have an average of zero

[11] We confirm that our results are similar using only FDO or NCCS data, respectively.

Electronic copy available at: https://ssrn.com/abstract=3143740

and a standard deviation of one by construction. The sample period for this data is 2011 (shortly after Twitter was launched) to 2021.

2.5 Consumer preferences

We also obtain two measures of consumers' time-varying preferences towards firms. First, following Larkin (2013), we obtain data on brand values from Brand Asset Valuator (BAV), a brand-level valuation model produced by a subsidiary of the advertising firm Young & Rubicam. The BAV model relies on consumer surveys (not accounting or market data) and is organized at the brand level, e.g., 'Fanta', 'Diet Coke', and 'Sprite' for the Coca Cola Company. The sample period is 2001 to 2011. Following Larkin (2013), our main metric of consumer perceptions is 'brand strength' (between 0 and 100), which captures how much regard and loyalty consumers have towards a given brand.

Second, we obtain firm-level customer churn data from Baker et al. (2023). Using individual-level credit card and checking account data, Baker et al. (2023) compute customer churn as the difference between the share of firm f's revenue coming from individual i going from period $t-k$ to t, averaged across individuals $i = 1, ..., I$. We use the loss of existing customers as our main measure of customer churn.[12] By construction, churn is distributed between 0 and 1. The sample period is 2011 to 2015.

2.6 IT security and investment

We measure firms' investments in IT security and infrastructure using a novel textual analysis of corporate 10-K filings. We begin by scanning each company's 10-K filing for keywords related to 'IT security' and count the number of occurrences. To construct a proxy for *investment* in IT security, we also scan each 10-K filing containing at least one IT Security reference for keywords related to 'investment' that are located within 100 characters before and after an 'IT security' reference. We remove instances where firms mention the terms 'data breach', 'hack', and 'hacking' to avoid confounding effects. We also construct control variables of the length and vocabulary complexity of each 10-K filing, following Loughran and McDonald (2014). Our 10-K based IT security and IT investment measures are new to the literature and complement the 10-K based textual analysis measures developed by Saunders and Tambe (2015) to capture firms' "data assets."

[12] Baker et al. (2023) construct overall churn, and churn by new and by existing customers. We focus on the latter measure because we are interested in how shocks to firms' reputations affect (existing) stakeholders.

Electronic copy available at: https://ssrn.com/abstract=3143740

2.7 Employee salaries

Further, we obtain individual-level occupation, job title, work status, and salary data from Glassdoor.com. The sample period for this data is 2006 to 2016. In addition to salary information, we also collect data from Glassdoor.com on individuals' work experience, occupation group, location (at the metropolitan area level), gender, and education.

2.8 Corporate political contributions

We obtain political contributions data from the Federal Election Commission. We manually match firms' political action committees (PACs) to their Compustat identifiers and use detailed PAC contribution data to calculate the total dollar amount of contributions to each candidate for U.S. Congress at the firm-candidate-election cycle frequency.

2.9 Other firm outcome and control variables

We obtain data on firm returns from CRSP and firm fundamentals from Compustat. Following Kamiya et al. (2021), we use data from BoardEx to construct a variable indicating if the firm has a board committee with a name that includes the word "risk", and a variable indicating if the CEO holds a dual role as the chairman of the board. We further collect indicators of S&P 500 membership for our sample firms from CRSP. Some variables in our analysis, such as firms' market-to-book (M/B) ratios, are computed by merging data from both CRSP and Compustat. In addition, we use data from the Compustat Segment Files to identify major suppliers and customers (i.e., corporate and government customers). Finally, we use Thomson Reuters' Institutional Holdings (13f) database to obtain the proportion of shares held by institutional investors who own at least 5% of a firm's outstanding equity.

3 Empirical approach

3.1 Main specification

We run a series of difference-in-differences tests to identify the effects of data breaches and other types of reputation-reducing events on corporate outcomes and subsequent firm responses. Our main sample is an unbalanced annual panel data set spanning the years 1999 to 2015. Our main regressions take the form:

$$y_{it} = \alpha + \gamma Post_{it} + \beta x_{it} + f_{j,t} + f_i + \varepsilon_{it} \ ,$$

Electronic copy available at: https://ssrn.com/abstract=3143740

where y_{it} captures outcomes such as the charitable contributions or CSR score for firm i in industry j in year t, x_{it} captures time-varying firm characteristics such as ln(Assets), ln(Assets)^2, and market leverage, f_{jt} represents industry-by-year fixed effects, and f_i represents firm fixed effects. Our main variable of interest is $Post_{ijt}$, which identifies firm-year observations following the disclosure of a reputation-reducing event such as a data breach. We use two different definitions of $Post_{ijt}$. Our first definition sets $Post_{ijt}$ equal to one for firms disclosing an event in either the current year or the previous year, while our second definition extends the post-event window from two years to four years by setting $Post_{ijt}$ equal to one for firms disclosing an event in either the current year or the previous three years.

Recent research has shown that two-way fixed effect (TWFE) estimators can be biased in differences-in-differences designs with staggered treatment (e.g., Goodman-Bacon, 2021; Sun and Abraham, 2021; Gardner, 2022; De Chaisemartin and d'Haultfoeuille, 2020). However, there is (to our knowledge) little research on how to compare the effects across different treatment groups using these types of estimators. Much of our analysis studies firms' attempts to tailor their reputation repair activities towards different stakeholders or reputation events of different salience, akin to a triple-difference analysis. As such, our main analysis retains the two-way fixed effects structure described above. However, we show in Section 7 that our main conclusions are robust to using these new estimation techniques.

3.2 Identification

Our empirical approach examines reputation loss and repair following two types of events: data breaches and RepRisk events. We focus on two distinct samples (data breaches and RepRisk) because both samples differ ex ante in their internal and external validity. RepRisk events arguably have high external validity because they cover many different categories of reputation-destroying events, but RepRisk events may have limited internal validity for causal inference because many of the event types covered by RepRisk (e.g., labor exploitation) are firm decision variables. On the other hand, a wealth of anecdotal evidence suggests that data breaches are likely outside of firms' direct control and are generally unrelated to a firm's products or services (e.g., breaches of customers' credit card data), but firms' responses to data breaches may potentially differ from their responses to another type of reputation shocks. Thus, relative to RepRisk events, data breach events likely have greater internal validity for causal inference, but potentially weaker external validity. Given these trade-offs, we include

Electronic copy available at: https://ssrn.com/abstract=3143740

both event types in our analysis.

3.2.1 Data breaches

There are two main identifying assumptions underlying our data breach tests. First, data breaches are assumed to be at least partially unpredictable, such that realizations of data breaches are not a direct function of observable or unobservable firm characteristics, *conditional* on controlling for fixed effects and time-varying observable controls. Second, data breach realizations, while affecting firms' reputations, are assumed to convey little information about firms' actual product qualities, management acumen, and other related variables. To the extent that these assumptions are true, data breaches would then arrive in a plausibly exogenous fashion and would affect firms' reputation among stakeholders without directly impacting stakeholders' beliefs about the firm's underlying products or the quality of firms' governance structures or managerial talent.

We document strong support for both assumptions. First, while firm characteristics are weakly associated (economically) with data breach realizations in the absence of fixed effects and controls, we show that *within-firm* data breach propensities are largely unpredictable over time. This makes intuitive sense: while (say) Target and Walmart may have slightly different data breach propensities in the cross-section due to different product offerings or different risk management practices, the realization of a data breach at Target instead of Walmart in a given time period, after accounting for both time-invariant and time-varying observable differences between the two firms, should largely be random. Anecdotal evidence strongly supports this argument: for example, Bassett et al. (2020) report that the top three causes of data breaches in their sample, i.e., phishing, stolen login credentials, and mis-delivered emails, are due to human errors. The timing of these events is plausibly exogenous, and even an infinite amount of IT spending arguably could not prevent such breaches. For our first identifying assumption to be violated, it would thus have to be the case that some unrelated event happened at Target (but not Walmart) in a given quarter or year that affected both the propensity of Target to experience a realized data breach, relative to Walmart, *and* similarly affected all of the various reputation loss and reputation repair variables we study at Target, relative to Walmart. We believe this possibility is highly unlikely to be true.

Nevertheless, we perform a number of tests that further cast doubt upon this potential alternative explanation. First, we examine the ability of time-varying firm characteristics (i.e., those that are most likely to be under the control of a firm) to predict the incidence of a

Electronic copy available at: https://ssrn.com/abstract=3143740

data breach versus the ability of time-fixed or firm-fixed factors, which we are able to control for in our analysis. Panel 1a of Table 1 presents the results of these tests, where the outcome variable indicates whether a firm is subject to a data breach with at least 1,000 compromised records in a given year. Column (1) utilizes our full sample and includes no fixed effects, whereas column (2) includes the same variables but adds industry-year and firm fixed effects. We add CSR-scores in column (3), governance indices in column (4), and additional controls for risk management, corporate governance, and IT expenditure in column (5). In columns (6) and (7) we add a host of additional firm characteristics found in Kamiya et al. (2021). All firm-level covariates are measured as of the year prior to the data breach.

Consistent with Kamiya et al. (2021), we find that some firm characteristics are statistically associated with data breaches. However, observable firm characteristics do not have substantial predictive power for data breach incidents. For example, the R^2 of the model without fixed effects is below 0.01, suggesting that typical "fundamental" variables have no meaningful ability to predict data breaches in the cross-section. Across specifications, the only variable in Panel 1a that reliably predicts data breaches is the M/B ratio, but the predictive value of this variable is economically small. While the overall R^2 increases meaningfully when including fixed effects, the within-R^2 remains below 0.1% across all columns. Other firm characteristics such as profitability, corporate social responsibility (i.e., normalized CSR scores), or corporate governance do not reliably predict data breaches. Similarly, none of the additional firm characteristics from Kamiya et al. (2021) in column (6) and (7) predict future data breaches in a statistically or economically meaningful way. Hence, time-varying observable firm characteristics do not appear to be systematically related to actual data breach incidents in our sample.[13] In sum, once we control for the cross-sectional factors that are most predictive of data breaches, data breaches seem relatively unpredictable.

Our second identifying assumption, which is that firms' responses to data breaches are driven by (expectations of) shocks to firm reputation rather than (say) direct shocks to product quality or managerial skill, is also highly intuitive. For example, it is unlikely that

[13]Indeed, while the total R^2 of our most saturated model is 0.209, essentially all of the model's explanatory power is from fixed effects that we are able to include as controls—the within-R^2 is only 0.0054. Moreover, in Appendix Table IA.3, we successfully replicate the results of Florackis et al. (2023), who find predictability of data breaches in the cross-section using a new measure of cybersecurity risk. However, similar to our overlap with Kamiya et al. (2021), once we include firm fixed effects in the regressions, the predictability of data breaches disappears. We do not include the cybersecurity measure of Florackis et al. (2023) in Table 1 as it significantly reduces the sample size due to low overlap with our other data sources.

Electronic copy available at: https://ssrn.com/abstract=3143740

consumers change their beliefs about the quality of (say) laundry detergent offered at Target because the firm experienced a data breach, and this is particularly true given the sources of most data breaches (phishing, stolen credentials, and so on), which are entirely unrelated to the firm's products or services. Similarly, given the ubiquitous nature of the data breach threat faced by most large corporations (which are the focus of our study), we believe it is highly unlikely that a material fraction of firm stakeholders view data breach *realizations* as a negative signal about a firm's governance or management acumen. As explained by Dave DeWalt, the former CEO of cybersecurity firms FireEye and McAfee and member of the US president's National Security Telecommuncations Advisory Council, data breaches are inevitable for nearly every company: "Even the strongest banks in the world; banks like JPMorgan, retailers like Home Depot, retailers like Target can't spend enough money or hire enough people to solve this problem." In a similar vein, the consulting firm Oliver Wyman states that "[c]yberattacks have become a permanent and persistent threat to organizations across commercial and government sectors. The question organizations are facing is not *if* a cyberattack will happen, but *when*. The difference between the winners and losers in a cyberattack, is how effectively the organization handles the response. The degree of loss and reputational damage (impact on brand value and customer loyalty) from a major cyberattack can be severe and irrevocable."

Of course, if data breaches and other reputation-impairing events are largely the result of bad luck, one may ask why such events would tarnish a firm's reputation with its stakeholders. However, since reputations ultimately reflect *perceptions* by outside stakeholders, all it takes to generate reputation loss following a data breach is for *some* consumers or other stakeholders to be concerned about the firm, even if data breaches are completely random and signal nothing about the firm's management team or future prospects. Indeed, the Oliver Wyman quote above argues both that cyberattacks are inevitable *and* that firms will experience reputational damage as a result of these (inevitable) attacks. This argument is illogical unless at least some stakeholders react (perhaps irrationally) to purely random data breach disclosures. For example, even sophisticated individuals are well known to have trouble evaluating base rates (see, e.g., the well known "Linda" experiments by Tversky and Kahneman, 1983), so even if all firms are subject to the same risk of a data breach, the first *realization* of a breach may lead some stakeholders to make (potentially irrational) negative inferences about the firm. Yet if some stakeholders *do* react to data breach disclosures, whether rational or not, firms

Electronic copy available at: https://ssrn.com/abstract=3143740

have every reason to respond.

3.2.2 RepRisk events

The identifying assumptions for our RepRisk events are similar to the assumptions for our data breach sample. On the one hand, since RepRisk events span a much wider array of negative corporate shocks – many involving what are most likely firm-level decisions – one might think that RepRisk events are more predictable than data breaches. However, similar to data breaches, Panel 1b shows that firm-level covariates do a relatively poor job of predicting RepRisk events, suggesting that these events are not, on average, direct consequences of observable firm characteristics. With few exceptions, none of the firm characteristics significantly explain the occurrence of future RepRisk events, and the effects of significant predictors such as profitability (ROA) and firm size are economically small. The within-R^2 is below 0.5% across all specifications, indicating that the collective ability of time-varying firm characteristics to predict negative reputation events is very small. Thus, despite the concerns noted above about RepRisk's internal validity, RepRisk events seem to be fairly unexpected given time-invariant and time-varying firm-level characteristics.

Similarly, while one might be concerned that RepRisk events contain negative information about firms' products, managers, or other related variables, Section 5 shows that firms respond similarly to RepRisk events and data breaches, suggesting that stakeholder-level beliefs following negative-reputation events are unlikely to be driven by concerns regarding such omitted variables.

4 How do firms lose reputation?

We answer our first research question by examining how a variety of stakeholders react after firms experience an event that impairs their reputation. We begin by studying reactions of financial market stakeholders', the media', and consumers' reactions to data breaches and conclude by studying how these stakeholders react to RepRisk events.

4.1 Short-term stock market value

We begin by confirming that firms' stock prices drop significantly after disclosing a data breach. We do so to show that these are important events that managers and stakeholders plausibly respond to. Figure 1a plots breached firms' average cumulative abnormal returns

Electronic copy available at: https://ssrn.com/abstract=3143740

(CARs) over the $[-10; 30]$ day event window.[14] As in Kamiya et al. (2021), the figure shows that CARs are consistently negative and account for a decline in firm value of 1.5% in the 30 days after a data breach. As shown in Figure 1b, we also find a negative market reaction of similar in magnitude around RepRisk events. The average CAR over the $[-10; 30]$ window is -1.02% for affected firms, and this return is more precisely estimated since there are many more RepRisk events than data breaches.

4.2 Equity Analysts

We next present evidence that data breaches and RepRisk events are associated with reductions in firms' reputations. We begin by using transcripts of firms' quarterly earnings calls to examine if firm executives and equity analysts discuss reputation and CSR-related topics more frequently following a reputation shock, as this would suggest that firms (or analysts) perceive a link between such events and firm reputations. To examine this idea, we construct variables indicating if firms mention data breaches, reputation-, or CSR-related terms, respectively, in conference calls, as described in Section 2.3. To distinguish between management's prepared remarks and interactions with equity analysts, we separately construct variables for management presentations and Q&A.

Table 2 presents results for the management presentation (Panel 2a) and Q&A-portion (Panel 2b) of conference calls following data breaches. We find that management presentations are more likely to mention data breaches (columns 1–4), reputation (columns 5–8), and CSR-related words (columns 9–12) following data breach events. The magnitudes are sizable. For example, in the two years following a data breach, managers are 3.01 pp more likely to talk about a data breach (column 2), 6.32 pp more likely to talk about reputation (column 6), and 7.86 pp more likely to mention CSR-related terms, which correspond to 750%, 39.7%, and 40.9% of the corresponding unconditional probabilities, respectively. We also find that analysts are more likely to ask about CSR-related issues in the aftermath of a data breach, both in the short and long term. For example, CSR terms are 6.2 (8.18) pp more likely to be discussed in the two (four) years after a data breach, which is a large increases relative to the unconditional probability of 26.6%.

As summarized in Panels 3a and 3b of Table 3, we find similar evidence in the years after a RepRisk event. For example, a discussion of 'reputation' is 3.65 pp more likely and

[14] We measure cumulative abnormal returns (CARs) using a 100-day estimation window that ends 50 days before a breach is publicly disclosed, using the Fama-French three-factor model.

Electronic copy available at: https://ssrn.com/abstract=3143740

CSR-related terms are 11.6 pp more likely to appear in the Q&A portion of a conference call (columns (4) and (8) of Panel 3b) up to four years after the occurrence of a RepRisk event. This effect can be compared to the unconditional probabilities of 13% and 26%, respectively. These large magnitudes underscore that data breaches and RepRisk events are material events to managers, while also suggesting that (1) reputation losses are important in the eyes of financial market stakeholders, (2) managers believe an association exists between data breaches and firm reputations, and (3) managers appear to link together reputation *loss* and reputation *repair* in their presentations to financial markets following a data breach.[15]

4.3 News Media

Data breaches and RepRisk scandals are likely to generate intense negative press for affected firms. Such negative media attention has been linked to reduced corporate reputations (see, e.g., Wartick, 1992), and one would expect a stronger, more persistent reaction in regions where the firm has a notable presence (Bénabou and Tirole, 2010), as the event is likely to have a particularly high salience for stakeholders in the local community. We therefore examine how sentiment towards affected firms in regional and national news media evolves following data breaches and RepRisk events.

We use the sentiment score from Ravenpack Edge to measure the tone of media reporting, as detailed in Section 2.4. We consider only "business" articles to isolate media sentiment about firms' economic decisions and remove articles related to the broader economy or issues that Ravenpack classifies as "society". We further also exclude news items related to cyberattacks to ensure that potential changes in media tone are capturing changes in firm perceptions broadly and not only narrowly related to data breach events.

Panel 2c of Table 2 presents the results of this analysis. Focusing on regional news outlets in the left panel, we find that media coverage becomes significantly more negative after a firm experiences a data breach. For example, in columns (3) and (6) we find that news sentiment decreases by 0.0326 (0.0212) in the two (four) years following a data breach, which corresponds to 17% (11%) of the unconditional sample mean. For national news outlets, we find a weaker, less precisely estimated effect. The negative coefficient estimates drop by about half compared to regional news outlets and are not significant at conventional levels. These findings indicate that data breaches lead to a long-term reduction in media sentiment

[15]These results also validate the increasing use of RepRisk in academic research (e.g., Derrien et al., 2021; Duchin et al., 2022).

Electronic copy available at: https://ssrn.com/abstract=3143740

about the firm in the local communities where the firm operates.

As documented in Panel 3c of Table 3, we find broadly similar results for the effect of RepRisk events on news media sentiment. The left panel focusing on regional news outlets reports lower news sentiment in the two years following a RepRisk event. The estimated effect is smaller and less precisely estimated compared to our data breach results.

4.4 Social media

We next examine the social media response to RepRisk scandals using measures of Twitter sentiment and "buzz" as detailed in Section IA.II.[16] Numerous firm stakeholder groups such as consumers, investors, the local community, and the overall public are consumers of and producers of social media content.

Twitter users typically react to ongoing events almost immediately, allowing us to estimate the social media reaction to reputation shocks at a high frequency. To reflect this high-frequency nature, we conduct our analysis at the firm-by-*day* level and plot the average daily social media sentiment (Fig. 2a) and "buzz" (Fig. 2b) for the $[-10; 15]$ event window around RepRisk events in Figure 2. Figure 2a shows that average social media sentiment drops below zero around the event day, remains depressed for about 5 days after the event, and subsequently returns to pre-event levels around zero.[17] Similarly, "buzz" (i.e., Tweet volume normalized in the daily cross-section), spikes on day $t = 0$ (Figure 2b), indicating that social media users are paying close attention to the reputation shocks.[18]

4.5 Consumers

Following Larkin (2013), we next use brand-level data from BAV (as detailed in Section 2.5) to examine whether consumers' perception of firms' brand strength changes following data breaches. A large literature in marketing and related fields has shown that brand (and firm) reputations are important inputs into consumer purchase decisions. If data breaches lead to reputation losses, firms' brand perceptions may suffer as a result. Indeed, Panel 2d of Table 2 documents that consumer perceptions of firms' brands are significantly lower up to four years after a data breach. Again, the economic magnitudes are sizeable. For example, as shown in

[16] Due to the emergence of social media in later part of our sample period, we are not able to implement a similar analysis for data breaches.

[17] This dynamic is consistent with the quickly changing, short-lived nature of Twitter discussions, where the main topics of conversation change on a daily basis.

[18] We note that both sentiment and buzz begin to change a few days before the event day, consistent with the stock market reaction shown in Figure 1b.

Electronic copy available at: https://ssrn.com/abstract=3143740

column (3), brand strength in the two years after a data breach declines by 2.59 points, or about 5.6% relative to the sample mean (and 9.8% of a standard deviation).

In Appendix Table IA.4 we further present supporting evidence that this finding is not due product-specific reputation effects but rather affects firm reputations more broadly. Specifically, this table shows that the effect of data breaches on brand strength is of similar magnitude for firms with a high number of brands compared to single-brand firms.

Since BAV data is only available for the early part of our sample, there is insufficient sample-overlap to implement similar tests for RepRisk events. Instead, we use data on customer churn at the quarterly frequency from Baker et al. (2023) to study the effect of negative reputation shocks on consumers, a key group of stakeholders. As detailed in Section IA.II, this measure of customer churn captures the spending-share-adjusted change of pre-existing customers from the previous to the current firm-quarter.[19]

The results are presented in Panel 3d of Table 3. In the specifications with firm fixed effects in columns (2), (4), and (6), we find a statistically significant, positive, albeit moderate effect of RepRisk events on customer churn for up to 16 quarters after the event. For example, the coefficient estimate of 0.0122 is equivalent to about 5% of a standard deviation. The effect is less precisely estimated in specifications without firm fixed effects. These results are consistent with recent evidence showing a negative effect of severe RepRisk events on consumer demand using grocery store scanner data (Christensen et al., 2023; Houston et al., 2023; Meier et al., 2023) and foot traffic data (Dube et al., 2023; Xiao et al., 2023).

4.6 Long-term market value

Finally, we examine whether lost reputation affects firm value over the long run. If negative reputation shocks were short-lived, managers would have substantially less motivation to respond to them. Panel 2e of Table 2 shows that firms' market-to-book ratios decline by at least .49 units (column 3) in the two years after the breach, or 17% of a standard deviation. We find that the decline in M/B is smaller but remains statistically significant up to four years after the data breach: columns (4)–(6) show that M/B declines by at least .27 units or 9% of a standard deviation in the four years following a breach.

We find qualitatively and quantitatively similar declines in market-to-book ratios for treated relative to control firms following the occurrence of RepRisk events, as shown in

[19]For example, churn would take the value of 0.5, if half of the consumers who shopped at firm i in the previous quarter purchased from the same firm again in the current quarter.

Electronic copy available at: https://ssrn.com/abstract=3143740

Panel 3e of Table 3. M/B values decline by at least 0.21 units in the two years after an event (column 3) and by at least .18 in the four years after an event (column 6), which corresponds to 7% and 6.2% of a standard deviation, respectively.[20]

Collectively, the results of this section demonstrate that corporate reputation is important to many types of firm stakeholders. We find that data breaches and other negative reputation shocks are followed by an increase in discussion of corporate reputation in earnings conference calls, negative press coverage both in traditional and social media, decreased brand value, and increased customer churn. These events have large negative effects on firm value, both in the short and long term, and are strikingly similar in magnitude across our two (very different) empirical settings.

5 How do firms respond to lost reputation?

Our results thus far suggest that a wide variety of firm stakeholders respond negatively to data breaches and other types of corporate scandals. We now examine how firms respond to these negative reputation shocks. While the universe of potential firm responses is essentially unlimited, we focus our analysis on five natural response types that line up with the types of stakeholders the literature has shown are often relevant for firm decision-making (see, e.g., Freeman, 2015): advertising expenditures, charitable contributions and CSR scores, employee salaries, political contributions and – specific to IT-related threats such as cyberattacks – investment in IT and cybersecurity.

5.1 CSR and charitable contributions

We begin by hypothesizing that firms will increase CSR investment in an attempt to rebuild their reputations among affected stakeholder groups following the realization of a negative reputation shock. While other work has linked CSR with responses to negative shocks, this work has focused on showing that firms investing in CSR *before* negative events benefit from less adverse reactions to the realization of the event. As discussed in Section 1.1, there are numerous reasons why it may also be optimal for firms to increase CSR investment *after* the realization of a negative shock, even in the presence of an ex-ante investment option. Existing work ties CSR to customer perceptions of a firm (e.g., Albuquerque et al., 2019, 2020; Rehman et al., 2020).

To test this hypothesis, we would ideally study a measure of firms' total investment in

[20] Appendix Table IA.5 reports similar findings across various measures of accounting performance.

Electronic copy available at: https://ssrn.com/abstract=3143740

CSR. However, CSR scores such as those constructed by KLD and Asset4 are measures of CSR "stocks" or levels rather than CSR investment, and are often measured with a significant lag. While CSR levels are likely correlated with investment, a more direct measure of CSR investment is desirable in this context. A key innovation of our paper is to introduce a direct measure of CSR investment: corporate charitable contributions. Corporate charitable contributions represent actual dollars spent during a given time period on CSR activities. As such, they are a direct (albeit partial) measure of CSR investment.

We first examine if corporate charitable contributions increase following data breaches and RepRisk events, respectively. The answer, as shown in Table 5, is yes. As shown in Panel 5a, donations from firms' charitable foundations increase by approximately $700,000 to $1.49 million per year in the two years following data breaches, and by similar amounts per year in the four years following data breaches.[21] We present results for charitable donations after RepRisk events in Panel 5b and find very similar results: charitable contributions increase by approximately $1 to $1.85 million per year in the two years following the event, and by similar amounts in the following years. These magnitudes are economically large; the average annual donation amount for the firms in our sample is only around $378,000 across all firms, and $3.12 million per year conditional on making any charitable contributions at all.

Panels 5c and 5d of Table 5 show that firms without charitable foundations are also more likely to set up a foundation following a data breach or RepRisk event. Depending on the specification, firms are 9 to 21 percentage points more likely to have a foundation after experiencing a data breach, relative to the control group of unaffected firms in similar six-digit GICS industry classifications. This represents roughly a doubling of the base rate of firms with charitable foundations (14.6%) across our sample. We find similar if not somewhat stronger results for foundation formation after a RepRisk event. Firms are 17–27 percentage points more likely to form a foundation in the two years after a RepRisk event, and 17–27 percentage points more likely to do so in the five years after the event.

We next examine whether firms' CSR investments translate into higher CSR scores. Understanding CSR scores is important for three reasons. First, while CSR investment data is available to us only for charitable contributions, firms may also invest in other types of CSR activities. CSR scores allows us to observe the long-run effects of these otherwise

[21] Tables IA.6 and IA.7 in the Appendix confirm that using alternative measures of charitable contributions from FDO and NCCS only, log-transformed donations, or charitable donations scaled by revenue as the dependent variable, respectively, yields similar overall results.

Electronic copy available at: https://ssrn.com/abstract=3143740

unobservable investments. Second, while the lag between investment and outcomes can often by substantial, CSR scores provide an external measure of the stakeholder benefits from firms' CSR investments, and thus indicate if firms' CSR investments ultimately increase the level of the firm's CSR stock (and in turn, the firm's reputation).Finally, CSR scores are commonly used by external parties such as institutional investors for decision-making purposes.

We present the results of this analysis in Panels 5e and 5f. We find that CSR scores are higher in the years following a data breach or RepRisk event. As shown in Panel 5f, CSR scores increase 15 to 32 percent of a standard deviation in the two years after the data breach, but statistical significance is somewhat muted. By the end of the five-year window, CSR scores increase by 39 to 52 percent of a standard deviation, an effect that is statistically significant at the one-percent level in all specifications. We find a sharper increase in the CSR scores of firms that suffer a RepRisk event (Panel 5f). CSR scores increase by 41 to 66 percent of a standard deviation in the two years following an event, and by 47 to 65 percent of a standard deviation in the five years following an event. While it is not possible to conclusively know why the CSR scores increase at a faster rate in the RepRisk sample, we theorize that this is due to the larger probability that firms form charitable foundations after a RepRisk event. Charitable foundations, in particular those that support local causes, are a component of how KLD assesses the social impact of a firm. Hence, it makes intuitive sense that CSR scores evolve faster when firms have a higher probability of taking actions that are included in the score calculation. We believe that these results further show that charitable activities are a useful measure of "investment" in CSR that the literature can use going forward. Finally, we present dynamics plots in Figure 3 for all CSR-related variables that we study. These plots confirm our main post-breach findings graphically and suggest that our results generally follow parallel trends prior to the data breach or RepRisk event.

5.2 Political contributions

We next examine whether firms increase their political contributions after they experience an event that impairs their reputation. A large literature has shown that the government is an important stakeholder for many firms, and while not all firms are politically active, many firms facilitate relations with the government by making political contributions (e.g., Cooper et al., 2010; Akey, 2015; Brogaard et al., 2021). We conjecture that firms may increase their political contributions following a reputation-impairing event in order to manage their

Electronic copy available at: https://ssrn.com/abstract=3143740

relationship with this important stakeholder.

Data on political contributions data is available to us at the firm-by-election-cycle level. Since election cycles are two years long, we choose to collapse our firm-year panel data at the firm-cycle level by averaging across all variables for each two-year period, and estimate difference-in-difference models as before at the firm-cycle level. Our main dependent variable is the dollar amount of political contributions in a given cycle.[22]

We present results for data breaches in Panel 6a of Table 6. We find evidence that firms increase their political contributions after they experience a data breach. In the two years following a breach (i.e., approximately one election cycle), political contributions increase by \$107–\$258 thousand, and by \$133–\$198 thousand in the four years after the breach, with most results statistically significant at the five-percent level. These figures represent about 29–71 percent of a standard deviation. Moreover, since PAC contributions to politicians are capped at \$10,000 per election cycle, this can be compared to contributing the maximum amount in an election cycle to 10 to 25 new politicians.

We find similar results for RepRisk events in Panel 6b of Table 6. As shown in columns (1)–(3), firms increase their political contributions by \$113–\$232 thousand in the two years after an event, and by \$110–\$193 thousand in the four years following a breach (columns 4–6). These results are statistically significant at the one-percent level and of comparable economic magnitudes to our results following data breaches.

5.3 Wages

We next examine if firms increase employee wages after they experience an event that impairs their reputation. Employees are an important stakeholder group, and both theoretical and empirical research suggests that firm reputation can be an important component in how firms and their employees contract for labor (e.g., Carmichael, 1984; Hales and Williamson, 2010; Rice and Schiller, 2023). It is possible that firms respond to lost reputation by increasing the wages of their employees in order to repair their reputation with this important stakeholder. We use wage data from Glassdoor as detailed in Section 2.7. Unlike the previous analyses, this data is collected at the employee-level (often from employees working in different occupations at the same firm) rather than the firm level, which allow us to control for more granular

[22]We also implement this test at the annual frequency, use log-transformed political contributions, and with governance controls following Kamiya et al. (2021) in Appendix Table IA.8, and find similar results.

Electronic copy available at: https://ssrn.com/abstract=3143740

person-level fixed effects in addition to the firm-level fixed effects included elsewhere.[23]

We present results for data breaches in Panel 7a of Table 7. The dependent variable is annual salary (in $K). All of our specifications include firm-level controls, along with firm, industry-year, metro-area, and occupation fixed effects. Standard errors are clustered at the unit of treatment, i.e., at the firm-level.

Across all specifications, we find that salaries increase after a firm suffers a data breach. In the two years following a breach, annual salaries of employees at treated firms increase by about $650, and by about $1,000 in the four years following the breach, relative to control firms in the same years. Relative to the unconditional sample mean of $73,000, this represents an increase of approximately 0.89% and 1.36%, respectively. This magnitude can also be compared to the effect of one additional year of work experience, for which we find an average increase in annual salary of about $1,800. The result are similar when we include additional fixed effects for employee education and gender in columns (2) and (4).

In contrast with our findings for data breaches, we do not find a robust response of salaries to RepRisk events in Panel 7b. While most of the point estimates are positive, they are substantially smaller than the point estimates that we obtain for data breaches and not statistically significant. The differences in responses for the two event categories is interesting, and provides some evidence that firms tailor their responses to different types of events. We return to this observation below, where we provide evidence that firms increase wages only when employees themselves are affected by a reputation-impairing event.

5.4 IT spending and advertising

We finally study whether firms increase their IT investment or advertising following adverse reputation shocks. We begin with IT investment. Firms experiencing a data breach might naturally respond by increasing their investment in IT, and it is possible (though less likely) that IT investment would increase following other types of negative reputation events as well. However, since IT investment is rarely disclosed in firms' financial statements, prior studies have been limited to reporting IT investment trends at the industry level (Kennedy and Stratopoulos, 2017; Aldasoro et al., 2020). To overcome this issue, we introduce a novel, firm-level measure of IT investment by examining the frequency with which firms discuss IT

[23]We choose to implement these tests at the individual person-level rather than collapsing at the firm-level to avoid concerns that any findings may be due to within-firm composition effects.

Electronic copy available at: https://ssrn.com/abstract=3143740

spending in their annual 10-K filings.[24]

It is also possible that firms seek to repair their reputations through increased advertising. For example, Cohen and Gurun (2018) find that firms attempt to positively influence the outcome of legal actions by purchasing advertising in the locations where they are subject to legal actions. Similarly, one could imagine firms purchasing either localized or nationwide advertising in the wake of a data breach or another corporate scandal.

Panel 4a of Table 4 presents results for IT investment and advertising expenditure following data breaches. Columns (1)–(6) examine IT investment, while columns (7)–(12) examine advertising expenditure scaled by total assets. We find that firms are more likely to discuss "IT security" (and related terms) in their 10-K filings following a data breach. Specifically, firms are between 3.8 and 6.9 percentage points more likely to discuss IT security in 10-K filings in the two years after the breach and 9.2–12.6 percentage points more likely to do so in the four years after the breach. These estimates are generally statistically significant, particularly at the longer horizon, and are highly economically significant. The unconditional likelihood that a firm discusses "IT security" and related terms is 13.9 percent, suggesting that over the longer horizon, firms nearly double the likelihood that they discuss these items in their 10-Ks. In contrast, we find no evidence that firms increase their advertising expenditures after a data breach. Indeed, as shown in columns (7)–(12), we find that the coefficients of interest are both economically and statistically insignificant. Hence, while it is possible that firms update the *content* of their advertisements in response to negative reputation shocks, they do not seem to purchase more advertisements following a shock.

Panel 4b presents the corresponding results for RepRisk events. Unlike data breaches, it is less natural for firms to increase IT investment in response to (say) a child labor or toxic emissions scandal. Hence, this test can be thought of as a type of placebo test: if firms do not respond to (say) animal mistreatment by increasing IT spending, it would suggest that firms appear to carefully tailor their reputation investments to the specific reputation-impairing event. Panel 4b shows that firms by and large do not increase their IT investment or advertising expenses following RepRisk events. While most of the point estimates on IT investments are positive, they are insignificant after firm-level controls are included, and the point estimates are substantially smaller than the responses to data breaches. We find similar

[24]We remove instances where firms appear to be discussing the data breach itself by excluding the terms 'data breach', 'hack', and 'hacking' to avoid confounding effects.

Electronic copy available at: https://ssrn.com/abstract=3143740

(non)-results for advertising expenses for RepRisk events as for data breaches.

6 Do firms tailor their responses to negative events?

Our results so far suggest that firms take a variety of actions to repair their reputation and their relations with key stakeholders following data breaches and other types of corporate scandals. We next examine whether firms tailor their reputation repair activities towards particularly important stakeholders or in response to events that are particularly salient to a specific group of stakeholders.

6.1 CSR investment and consumer-facing firms

We begin by examining whether consumer-facing firms increase CSR investment more than business-facing firms following negative events. Existing work has suggested that CSR activities aid in customer perceptions of product differentiation (e.g., Albuquerque et al., 2019, 2020) and in building customer loyalty (e.g., Rehman et al., 2020). We conjecture that consumer-facing firms therefore have a stronger incentive to increase CSR investment following negative shocks relative to business-facing firms. We test this conjecture by redefining our event variables to separately measure reputation-impairing events (e.g., data breaches) that affect consumer-facing firms or business-facing firms to directly examine the magnitude of responses across firm types. We delineate consumer-facing and business-facing firms at the industry-level, using data from the Compustat Segments database as detailed in Section 2.9. Specifically, we define a firm as business-facing (consumer-facing) if the percentage of firms in its industry with major corporate customers is above (below) the median. [25] Our specifications are otherwise the same as in Section 5.

We present these results in Table 8. Panel 8a presents our results for data breaches. In all but one specification, we find that the point estimate of the CSR response is larger for consumer-facing firms than for business-facing firms (i.e., the Post coefficients that are interacted with "HasCCus = Low"). These results are most stark in columns (1)–(4), which examine charitable donations. Specifically, we find that *only* consumer facing firms respond to data breaches by increasing charitable donations. For example, in columns (3) and (4) we find that consumer-facing firms increase their charitable contributions by $1.64 and $1.06

[25] Our results are qualitatively similar if we use industry-level advertising expenditures to define consumer- and business-facing firms, although some of the coefficient differences are measured with less statistical precision, as shown in Appendix Table IA.9.

Electronic copy available at: https://ssrn.com/abstract=3143740

million, respectively, while business-facing firms increase donations by an insignificant $87 thousand dollars. These differences in responses between consumer- and business-facing firms are statistically and economically significant.

We find similar patterns for charitable contributions in consumer- and business-facing industries following RepRisk events in Panel 8b. The point estimates are systematically higher in consumer-facing industries for both the amount of charitable charitable contributions (columns 1–4) and for the probability of starting a foundation (columns 5–8) and generally statistically different from each other at conventional levels. We find mixed results for KLD scores, without clear patterns in economic magnitudes or statistical significance as to whether external CSR scores are increasing more for consumer- or business-facing firms. However, it is possible that KLD scores are increasing similarly for business-facing firms due to investments in other types of CSR activities that are not captured by our charitable contributions data.

6.2 Political contributions and government suppliers

We next examine whether firms for which the government is a major customer increase their political contributions more than other politically active firms. Recent research suggests that governments monitor the ESG performance of potential contractors (Flammer, 2018; Gantchev et al., 2022).[26] Hence, we conjecture that government contractors are likely to differentially increase their political contributions after they experience events that impair their reputation since the government is a particularly important stakeholder. We define government contractors as firms reporting at least one government entity (e.g., state government, federal department, federal agency, etc.) as a principal customer in the Compustat Segment files, and separately identify changes in political contributions of major government contractors and other politically active firms after a data breach or RepRisk event.

We present these results in Table 9. Focusing on data breaches, we find that the point estimates for both government contractors and non-contractors are positive for both the short- and long-term horizons, but that the economic magnitudes and statistical significance are substantially larger for government contractors (Panel 9a). For example, as shown in column (4), government contractors increase their political contributions by $471 thousand in the four years following the data breach (significant at the one percent level), while

[26] A large literature shows that firms without government contracts also form political connections, for example to mitigate policy uncertainty (e.g., Hassan et al., 2019; Akey and Lewellen, 2020) or acquire policy-relevant information (e.g., Ovtchinnikov et al., 2020), among many others.

Electronic copy available at: https://ssrn.com/abstract=3143740

non-government contractors increase their contributions by only $73 thousand (significant at the 10 percent level). The difference between these responses is itself statistically significant at the five-percent level, as shown in the chi-squared test at the bottom of the table.

We find analogous results for RepRisk events in Panel 9b. Government contractors increase their political contributions by more than non-contractors at all horizons following RepRisk events. While contributions by both firm types are positive and statistically significant, we find consistent evidence that the economic magnitudes are substantially larger for government contractors. Turning again to column (4), we find that government contractors increase their political contributions by $340 thousand after a RepRisk event, while non-contractors increase their contributions by only $69 thousand, with both results statistically significant at the one percent level. As with data breaches, the difference between these coefficients is itself statistically significant at the one percent level.

6.3 Political contributions and legal violations

In our final two tests we study how firms tailor their response to lost reputation following shocks that disproportionately affect a specific group of stakeholders. To this end, we first examine whether firms increase their political contributions more sharply for specific types of RepRisk events. RepRisk classifies negative reputation events into five broad categories: 'environmental', 'social', 'governance', 'product', and *'violations'*, which indicate violations of domestic or international standards, regulations, treaties, or sanctions. This allows us to test if firms adjust their political contributions differently after violations of government policy as opposed to other types of RepRisk event, as firms might find it more costly to repair their relationship with politicians if the lost reputation is related to government policy.

We split our variable of interest to account for RepRisk events that are classified as "violations" by the data provider and those that are not, and separately estimate the effect on political contributions across the two groups in Table 10. While we find that political contributions rise after all types of RepRisk events, RepRisk events that involve violations are associated with a stronger increases in campaign contributions. More specifically, we find that the "Post" variables interacted with "RRI-V=Yes" have larger magnitudes than those interacted with "RRI-V=No." For example, in column (2) we find that the within-firm increase in political contributions is $133 thousand in the two years after a RepRisk event that relates to a violation, but only $79 thousand for a RepRisk event involving other

Electronic copy available at: https://ssrn.com/abstract=3143740

categories, a difference that is statistically significant at the 10 percent level. Hence, firms may attempt to repair their reputations with the same stakeholder differently based on the *type* of reputation-damaging event that occurs.

6.4 Wages, employee-focused data breaches, and labor mobility

Last, we examine whether firms increase employee salaries more sharply following a data breach that directly affects employee records. To the extent that firms' reputations matter to employees, we might expect that employee salaries increase the most when a data breach affects them personally.

We present the results of this analysis in Table 11. As in our other tests of tailored firm-responses, we redefine our shock variables to identify those data breaches that impact employee records and those data breaches that do not. We find that employee salaries only increase after data breaches that involve employee records. Across all post-breach horizons, we find that the "Post" coefficients interacted with "Empl. Hack=Yes" are positive and statistically significant at the one percent level, and represent an annual increase in employee salaries of roughly $1,100 to $1,500 per year. In contrast, we find substantially smaller coefficients for other data breaches (i.e., the 'post' coefficients interacted with "Empl. Hack=No"). While all of the estimates are positive, only one coefficient is statistically significant and the economic magnitudes are smaller and statistically different from the coefficients estimated for breaches that affect employee records.

Additionally, we study if employee salaries change more strongly in firms where employees have higher labor mobility, as firms may find it more costly to repair their reputation with employees when employees have greater outside options in the labor market. We use labor redeployability (i.e., the percentage of workers in a given firm with 'common' job titles), and the average education level and average salary within a firm, respectively, to delineate firms with a high- and low labor mobility workforce, and estimate similar tests as before. The results, summarized in Table IA.10, broadly show a stronger effect of data breaches on employee salaries for firms with higher labor mobility, especially at the one-year horizon. Collectively, our results suggest that not only do firms take a variety of different actions to repair their reputation, but they target their responses to key stakeholders and events that are particularly salient for their stakeholders.

Electronic copy available at: https://ssrn.com/abstract=3143740

7 Additional robustness

We perform many additional tests to ensure that our results are not driven by changes in risk management or corporate governance and are robust to alternative econometric specifications, alternative definitions of outcome variables, and a variety of potential identification concerns.

7.1 Risk management and corporate governance

In closely related work, Kamiya et al. (2021) find that firms increase their focus on risk management following data breaches. This raises a concern that the post-breach increases in CSR we document in Table 5 could be capturing an increased focus on risk management rather than firms' attempts to repair their reputations. It is also possible that, even if firms *are* actively working to rebuild their reputations following data breaches, they are doing so because of an enhanced focus on risk management. Either of these narratives would raise serious questions about whether the channel we document in Table 5 is truly distinct from the risk management channel documented by Kamiya et al. (2021).

To examine this question, we first rerun our main tests from Section 5 (Tables 5 to 7) after including one of the key risk management variables used in Kamiya et al. (2021), i.e., a dummy indicating that the firm has a board-level committee with "risk" in the name. We also include an indicator variable for whether the CEO has a dual role as chairperson, and a variable indicating the fraction of shares owned by institutional blockholders, as well as controls for corporate governance using the governance indices proposed by Gompers et al. (2003) and Bebchuk et al. (2009). These variables should capture broader changes in the firm's governance structure or ownership base that could in turn affect firms' risk management or reputation repair activities.

Internet Appendix Tables IA.8, IA.11, and IA.12 contain the results of these tests using CSR scores, political contributions, and employee salaries, respectively, as dependent variables. Panel (a) of each table presents data breach results while panel (b) presents results for RepRisk events. Our results are qualitatively unchanged with the inclusion of these controls. A small minority of estimates lose statistical significance, but the inclusion of these additional variables reduces the sample size of some of the tests substantially, which we believe explains the reduced statistical significance of the tests.

While these results suggest that our findings are unlikely to be driven by an increased focus on risk management or governance, we re-run our risk management robustness tests

Electronic copy available at: https://ssrn.com/abstract=3143740

for CSR activity after allowing for an interaction effect between the post-breach indicators (which arguably capture firms' reputation repair responses) and the main risk management variable from Kamiya et al. (2021). If our post-breach variables are capturing firms' risk management responses rather than firms' reputation repair activities, we would expect our main effects to be small and insignificant, while the interaction effects should be large and significant. Internet Appendix Table IA.13 shows that this is not the case.

7.2 Alternative estimation of differences-in-differences

Recent work has shown that differences-in-differences (DiD) estimation with staggered treatments may be biased in two-way fixed effects (TWFE) estimation (e.g., De Chaisemartin and d'Haultfoeuille, 2020; Sun and Abraham, 2021; Gardner, 2022). We verify that our main results on firm responses to reputation losses are robust to this concern by implementing our analysis using the estimators of Sun and Abraham (2021) and Gardner (2022), as these methods explicitly account for staggered treatment in difference-in-differences settings.

Appendix Table IA.14 presents results that examine CSR activity, political contributions, and wages following both data breaches and RepRisk events, respectively. Panels (a) and (c) report estimates using the Sun and Abraham (2021) estimator, while panels (b) and (d) report estimates using the Gardner (2022) estimator. As in our main analysis, we report average treatments effects on the treated (ATTs) estimated from specifications that omit or include firm fixed effects. Since the Sun and Abraham (2021) estimator requires selecting a "normalization" period in models with firm fixed effects, we provide estimates using either $t = 0$ or $t = -1$ relative to the reputation-impairing event as the normalization period. Our results are qualitatively similar using these techniques, suggesting that our main results do not suffer from a large degree of estimation bias.

We do not perform similar analyses for our tests examining heterogeneous firm responses to reputation events as, to the best of our knowledge, the literature does not yet have an established set of analogous techniques to estimate triple-differences or difference-in-coefficients tests across different groups. More generally, a recent survey of the econometrics literature on the topic (De Chaisemartin and d'Haultfoeuille, 2023) notes that it is not clear that researchers should systematically abandon two-way fixed effects estimation. For example, De Chaisemartin and d'Haultfoeuille (2023) cite the case of multiple, repeated treatments as one reason not to abandon TWFE. This applies especially to our setting using RepRisk

Electronic copy available at: https://ssrn.com/abstract=3143740

events, where several firms experience up to 8 or 9 treatments throughout the sample period.

7.3 Alternative measurement of CSR

We next examine the robustness of our results to alternative measures of CSR scores using Thompson Reuters' Asset4 database (now known as Refinitiv ESG) of corporate CSR scores, which has been extensively used in the literature (e.g., Ferrell et al., 2016; Liang and Renneboog, 2017). Appendix Table IA.15 shows how Asset4 scores change following data breaches and negative reputation events from RepRisk, respectively. In Columns (1) and (2), we provide estimates of our most saturated regression models using the composite Asset4 ESG score, while columns (3)–(8) present the coefficients for each sub-component of Asset4. The results show that CSR scores increase following negative reputation events even when using the Asset4 measure of CSR, which suggests that our results are not driven by our choice of KLD as our primary measure of CSR.[27]

7.4 Other omitted variables

We examine how robust our estimates are to selection biases related to unobservable covariates using the bounding procedure proposed by Oster (2019). This method estimates how large an omitted variable or selection bias would need to be in order to change the sign of an estimated coefficient, by considering how the incremental addition of control variables changes the magnitude of the coefficient of interest and the R^2 of the regression model, under the assumption that selection on unobservables is proportionate to selection on observables. The procedure estimates a test statistic, δ^*, that indicates the impact of introducing covariates. A finding of $\delta^* < 0$ indicates that the introduction of covariates causes the coefficient of interest to *increase* in magnitude, while $\delta^* \geq 0$ indicates that the introduction of covariates causes the coefficient of interest to fall. At $\delta^* = 1$, the suggested critical value for this test, selection on unobservables would need to be as large as selection on observables to render a point estimate equal to zero. Appendix Table IA.17 presents the results from this analysis for all of our major dependent variables of interest.[28] All estimates of δ^* are either negative,

[27] We also confirm that our CSR results following data breaches are robust to removing changes in CSR scores that are related to potential changes in IT security, governance, and risk management (see Table IA.16). Our results are also not driven by changes in governance or risk management CSR scores; the KLD factors associated with these issues are not time-consistent and are hence dropped in the process of constructing the normalized CSR scores we use in our main tests.

[28] Due to the computational complexity of implementing the Oster (2019) with high-dimensional fixed effects, we are unable to execute this test for wages at the employee level.

Electronic copy available at: https://ssrn.com/abstract=3143740

indicating that the introduction of relevant control variables strengthens our results, or are greater than 3, suggesting that selection on unobservables would need to be at least three times larger than selection on observables to reduce our point estimates to zero.

7.5 Matching estimator

Finally, one of our identifying assumptions is that all firms in a given six-digit GICS industry classification have similar probabilities of experiencing a data breach at a given point in time after controlling for observable characteristics. As shown in Table 1, within a given industry-year pair, observable characteristics cannot reliably predict which firms will experience data breaches, and the combination of controls, fixed effects, and the Oster (2019) bound test should largely eliminate concerns about correlated observable or unobservable variables driving our results. Nevertheless, to reduce any remaining concerns about covariate balance, we re-estimate our main tests after matching each treated firm with 10 industry peers using propensity scores based on observable variables such as (lagged) firm value, CSR scores, governance, size, leverage, and profitability. Appendix Table IA.18 shows that all of our results hold qualitatively (and most hold quantitatively) using these propensity-score matched control groups despite an 80% reduction in sample size, providing further validity for the identifying assumptions behind our main tests.

8 Conclusions

We exploit unexpected corporate data breaches and other negative events to study how firms respond to the destruction of reputational capital. We focus on data breaches because they represent negative reputation shocks that are plausibly uncorrelated with firms' underlying product quality or financial performance. Existing research has studied investment actions designed to insure against future negative shocks such as CSR investment, but little is known about how firms respond to *rebuild* reputation once this intangible asset is impaired.

We propose a definition of reputation that harkens back to Hayek (1948)—the set of *value-relevant* firm characteristics that affect stakeholders' perceptions about the firm—where "stakeholders" are the (large) set of agents the firm may reasonably compete for on the basis of its reputation or goodwill. We find that negative shocks to reputation matter to a variety of such stakeholders. These events prompt large declines in equity values, more negative news coverage, both in traditional media and on social media, as well as worsened consumer perceptions of the affected firms. Managers discuss reputation and CSR more frequently in

Electronic copy available at: https://ssrn.com/abstract=3143740

their communications with capital market stakeholders.

Next, we examine a variety of actions that firms may take to repair their reputation with these stakeholders. We find that firms, in addition to increased IT security investments, respond along a variety of margins that are not directly related to their operations: For example, firms' spend more on charitable contributions after shocks to their reputation, which subsequently translates into higher CSR scores. We also find that political contributions are higher after reputation shocks and that firms increase annual employee wages following data breaches. However, firms do not seem to respond on *all* margins. Perhaps surprisingly, we do not find increases in advertising expenditures after reputation-impairing events.

Finally, we provide evidence that firms' responses are tailored to specific stakeholders and types of reputation-impairing events. We find that consumer-facing firms increase their CSR-related investments more strongly than other firms, consistent with theories of CSR as an instrument of product differentiation. We show that firms with major government contracts increase their political contributions more than other types of firms, and that firms increase their political contributions more sharply after events classified as "violations" of government policies and regulations. Finally, we find that firms only increase employee wages after a data breach affects employee records. Collectively, these results suggest that firms rationally target their responses to repair reputations to prioritize particularly important stakeholders or events.

To our knowledge, our paper is the first to present direct evidence of *tangible* corporate investments in *intangible* capital following negative corporate reputation shocks. The burgeoning literature on CSR has largely overlooked the factors that cause firms to increase direct investment in CSR, while the growing literature on intangible capital has largely overlooked the role of CSR as an effective form of intangible investment. By describing how firms respond to the destruction of intangible capital, our paper links these two literatures and helps to improve our understanding of the within-firm catalysts that drive intangible investment decisions. Our paper also presents a novel measure of CSR investment, novel measures of corporate IT spending, and new evidence regarding the long-term value destruction caused by negative corporate reputation shocks. Benjamin Franklin may have been right when he said that it only takes one bad deed to lose a good reputation, but our results suggest that firms can take tangible steps to repair their intangible mis-steps.

Electronic copy available at: https://ssrn.com/abstract=3143740

References

Acquisti, A., Friedman, A., and Telang, R. (2006). Is There a Cost to Privacy Breaches? An Event Study. In *Proceedings of the Twenty-Seventh International Conference on Information Systems*.

Ahn, C., Houston, J. F., and Kim, S. (2020). Hidden in plain sight: The role of corporate board of directors in public charity lobbying. *Working paper, University of Florida*.

Akey, P. (2015). Valuing Changes in Political Networks: Evidence from Campaign Contributions to Candidates in Close Congressional Elections. *Review of Financial Studies*, 28:3188–3223.

Akey, P. and Lewellen, S. (2020). Policy Uncertainty, Political Capital, and Firm Risk-Taking. Unpublished working paper.

Albuquerque, R., Koskinen, Y., Yang, S., and Zhang, C. (2020). Resiliency of environmental and social stocks: An analysis of the exogenous covid-19 market crash. *The Review of Corporate Finance Studies*, 9(3):593–621.

Albuquerque, R., Koskinen, Y., and Zhang, C. (2019). Corporate Social Responsibility and Firm Risk: Theory and Empirical Evidence. *Management Science*, 65(10):4451–4469.

Aldasoro, I., Gambacorta, L., Guidici, P., and Leach, T. (2020). The Drivers of Cyber Risk. CEPR Discussion Paper 14805.

Armour, J., Mayer, C., and Polo, A. (2017). Regulatory Sanctions and Reputational Damage in Financial Markets. *Journal of Financial and Quantitative Analysis*, 52(4):1429–1448.

Baker, S. R., Baugh, B., and Sammon, M. (2023). Customer churn and intangible capital. *Journal of Political Economy - Macroeconomics*, 1(3):000–000.

Bansal, R., Wu, D., and Yaron, A. (2022). Socially responsible investing in good and bad times. *The Review of Financial Studies*, 35(4):2067–2099.

Barrage, L., Chyn, E., and Hastings, J. (2020). Advertising, reputation and environmental stewardship: Evidence from the BP oil spill. *American Economic Journal: Economic Policy*, 12:33–61.

Bassett, G., Hylander, C. D., Langlois, P., Pinto, A., and Widup, S. (2020). *2020 Data Breach Investigations Report*. Verizon, Inc.

Bebchuk, L. A., Cohen, A., and Ferrell, A. (2009). What Matters in Corporate Governance? *Review of Financial Studies*, 22:783–827.

Belo, F., Gala, V. D., Salomao, J., and Vitorino, M. A. (2022). Decomposing firm value. *Journal of Financial Economics*, 143(2):619–639.

Belo, F., Lin, X., and Vitorino, M. A. (2014). Brand capital and firm value. *Review of Economic Dynamics*, 17(1):150–169.

Bénabou, R. and Tirole, J. (2010). Individual and Corporate Social Responsibility. *Economica*, 77:1–19.

Bertrand, M., Bombardini, M., Fisman, R., Hackinen, B., and Trebbi, F. (2018). Hall of Mirrors: Corporate Philanthropy and Strategic Advisory. Working Paper, University of Chicago.

Bertrand, M., Bombardini, M., Fisman, R., and Trebbi, F. (2020). Tax-Exempt Lobbying: Corporate Philanthropy as a Tool for Political Influence. *American Economic Review*, 110(7):2065–2102.

Bloom, N., Garicano, L., Sadun, R., and Reenen, J. V. (2014). The Distinct Effects of Information Technology and Communication Technology on Firm Organization. *Management Science*, 60:2859–2885.

Board, S. and Meyer-ter-Vehn, M. (2013). Reputation for quality. *Econometrica*, 81:2381–2462.

Brogaard, J., Denes, M., and Duchin, R. (2021). Political influence and the renegotiation of government contracts. *The Review of Financial Studies*, 34(6):3095–3137.

Cai, Y., Xu, J., and Yang, J. (2021). Paying by donating: Corporate donations affiliated with independent directors. *Review of Financial Studies*. Forthcoming.

Campbell, K., Gordon, L. A., Loeb, M. P., and Zhou, L. (2003). The Economic Cost of Publicly Announced Information Security Breaches: Empirical Evidence from the Stock Market. *Journal of Computer Security*, 11(3):431–448.

Carmichael, H. L. (1984). Reputations in the labor market. *The American Economic Review*, 74(4):713–725.

Chakravarthy, J., DeHaan, E., and Rajgopal, S. (2014). Reputation Repair after a Serious Restatement. *The Accounting Review*, 89(4):1329–1363.

Chatterji, A. K., Luo, J., and Seamans, R. C. (2015). Competition between Organizational Forms: Banks vs. Credit Unions after the Financial Crisis. Working Paper, Duke University.

Christensen, H. B., De George, E. T., Joffre, A., and Macciocchi, D. (2023). Consumer responses to corporate social irresponsibility. Working Paper, SSRN ID 4496599.

Cohen, L. H. and Gurun, U. G. (2018). Buying the Verdict. Working paper, National Bureau of Economic Research.

Cooper, M., Gulen, H., and Ovtchinnikov, A. (2010). Corporate Political Contributions and Stock Returns. *Journal of Finance*, 65:687–724.

Corhay, A., Kung, H., and Schmid, L. (2020). Q: Risk, Rents, or Growth? Working paper, London Business School.

Corrado, C. A. and Hulten, C. R. (2010). How Do You Measure a "Technological Revolution"? *American Economic Review: Papers and Proceedings*, 100:99–104.

Crouzet, N. and Eberly, J. (2018). Intangibles, Investment, and Efficiency. *American Economic Review: Papers and Proceedings*, 108:426–431.

Dai, R., Liang, H., and Ng, L. K. (2020). Socially Responsible Corporate Customers. *Journal of Financial Economics*, forthcoming.

De Chaisemartin, C. and d'Haultfoeuille, X. (2023). Two-way fixed effects and differences-in-differences with heterogeneous treatment effects: A survey. *Econometrics Journal*. Forthcoming.

De Chaisemartin, C. and d'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9):2964–2996.

Derrien, F., Krueger, P., Landier, A., Yao, T., et al. (2021). How do ESG incidents affect firm value? Working Paper, Swiss Finance Institute.

Dube, S., Lee, H. S. G., and Wang, D. (2023). Do consumers vote with their feet in response to negative ESG news? evidence from consumer foot traffic to retail locations. Working Paper, SSRN ID 4506950.

Duchin, R., Gao, J., and Xu, Q. (2022). Sustainability or greenwashing: Evidence from the asset market for industrial pollution. Working Paper, Available at SSRN.

Electronic copy available at: https://ssrn.com/abstract=3143740

Edmans, A. (2011). Does the Stock Market Fully Value Intangibles? Employee Satisfaction and Equity Prices. *Journal of Financial Economics*, 101(3):621–640.
Elfenbein, D. W., Fisman, R., and McManus, B. (2012). Charity as a Substitute for Reputation: Evidence from an Online Marketplace. *The Review of Economic Studies*, 79(4):1441–1468.
Ferrell, A., Liang, H., and Renneboog, L. (2016). Socially Responsible Firms. *Journal of Financial Economics*, 122(3):585–606.
Flammer, C. (2013). Corporate social responsibility and shareholder reaction: The environmental awareness of investors. *Academy of Management Journal*, 56:758–781.
Flammer, C. (2015). Does Corporate Social Responsibility Lead to Superior Financial Performance? A Regression Discontinuity Approach. *Management Science*, 61(11):2549–2568.
Flammer, C. (2018). Competing for government procurement contracts: The role of corporate social responsibility. *Strategic Management Journal*, 39(5):1299–1324.
Florackis, C., Louca, C., Michaely, R., and Weber, M. (2023). Cybersecurity risk. *The Review of Financial Studies*, 36(1):351–407.
Forman, C., Goldfarb, A., and Greenstein, S. (2012). The Internet and Local Wages: A Puzzle. *American Economic Review*, 102:556–575.
Freeman, R. E. (2015). *Stakeholder Theory*, pages 1–6. John Wiley and Sons, Ltd.
Fudenberg, D. and Levine, D. K. (1989). Reputation and equilibrium selection in games with a patient player. *Econometrica*, 57(4):759–778.
Gantchev, N., Gianetti, M., and Li, R. (2020). Does Money Talk? Market Discipline Through Selloffs and Boycotts. Southern Methodosist University working paper.
Gantchev, N., Goldman, J., and Zhang, S. (2022). The role of g(overnment) in corporate ESG policies. Working Paper, Available at SSRN 4280531.
Gardner, J. (2022). Two-stage differences in differences. *arXiv preprint arXiv:2207.05943*.
Godfrey, P. C., Merrill, C. B., and Hansen, J. M. (2009). The Relationship Between Corporate Social Responsibility and Shareholder Value: An Empirical Test of the Risk Management Hypothesis. *Strategic Management Journal*, 30(4):425–445.
Gompers, P., Ishii, J., and Metrick, A. (2003). Corporate Governance and Equity Prices. *Quarterly Journal of Economics*, 118:107–155.
Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2):254–277.
Hales, J. and Williamson, M. G. (2010). Implicit employment contracts: The limits of management reputation for promoting firm productivity. *Journal of Accounting Research*, 48(1):51–80.
Hassan, T. A., Hollander, S., van Lent, L., and Tahoun, A. (2019). Firm-level political risk: Measurement and effects. *Quarterly Journal of Economics*.
Hayek, F. A. (1948). *The Meaning of Competition*, chapter 5, pages 92–106. University of Chicago Press.
Heal, G. (2005). Corporate Social Responsibility: An Economic and Financial Framework. *Geneva Papers*, 30:387–409.
Holmstrom, B. (1999). Managerial Incentive Problems: A Dynamic Perspective. *Review of Economic Studies*, 66:169–182.
Hong, H. and Kacperczyk, M. (2009). The Price of Sin: The Effects of Social Norms on Markets. *Journal of Financial Economics*, 93(1):15 – 36.
Hong, H. and Liskovich, I. (2014). Crime, Punishment and the Halo Effect of Corporate Social Responsibility. Working paper, Princeton University.
Hong, H. G., Kubik, J. D., Liskovich, I., and Scheinkman, J. A. (2019). Crime, punishment and the value of corporate social responsibility. *Available at SSRN 2492202*.
Houston, J. F., Lin, C., Shan, H., and Shen, M. (2023). How does ESG shape consumption? Working Paper, SSRN ID 4243071.
Houston, J. F. and Shan, H. (2022). Corporate ESG profiles and banking relationships. *The Review of Financial Studies*, 35(7):3373–3417.
Huang, C., Li, F., and Weng, X. (2020). Star ratings and the incentives of mutual funds. *Journal of Finance*, 75(3):1715–1765.
Jäger, S. (2016). How Substitutable Are Workers? Evidence from Worker Deaths. Working Paper, MIT.
Jarrell, G. and Peltzman, S. (1985). The Impact of Product Recalls on the Wealth of Sellers. *Journal of Political Economy*, 93(3):512–536.
Kamiya, S., Kang, J.-K., Kim, J., Milidonis, A., and Stulz, R. M. (2021). Risk management, firm reputation, and the impact of successful cyberattacks on target firms. *Journal of Financial Economics*, 139:719–749.
Karpoff, J. M. (2012). Does reputation work to discipline corporate misconduct? In Barnett, M. L. and Pollock, T. G., editors, *The Oxford Handbook of Corporate Reputation*, chapter 18, pages 361–382. Oxford University Press.
Karpoff, J. M., Lee, D. S., and Martin, G. S. (2008). The Cost to Firms of Cooking the Books. *Journal of Financial and Quantitative Analysis*, 43(3):581–611.
Karpoff, J. M., Lott Jr, J. R., and Wehrly, E. W. (2005). The Reputational Penalties for Environmental Violations: Empirical Evidence. *Journal of Law and Economics*, 48(2):653–675.
Kennedy, D. B. and Stratopoulos, T. C. (2017). Mapping IT Spending Across Industry Classifications: An Open Source Dataset. Working paper, University of Waterloo.
Kitzmueller, M. and Shimshack, J. (2012). Economic Perspectives on Corporate Social Responsibility. *Journal of Economic Literature*, 50(1):51–84.
Klein, B. and Leffler, K. B. (1981). The role of market forces in assuring contractual performance. *Journal of Political Economy*, 89:615–641.
Kreps, D. M. and Wilson, R. (1982). Reputation and imperfect information. *Journal of Economic Theory*, 27:253–279.
Larkin, Y. (2013). Brand perception, cash flow stability, and financial policy. *Journal of Financial Economics*, 110(1):232–253.
Lending, C., Minnick, K., and Schorno, P. J. (2018). Corporate Governance, Social Responsibility, and Data Breaches. *The Financial Review*, 53:413–455.
Levine, D. K. (2021). The reputation trap. *Econometrica*, 89(6):2659–2678.

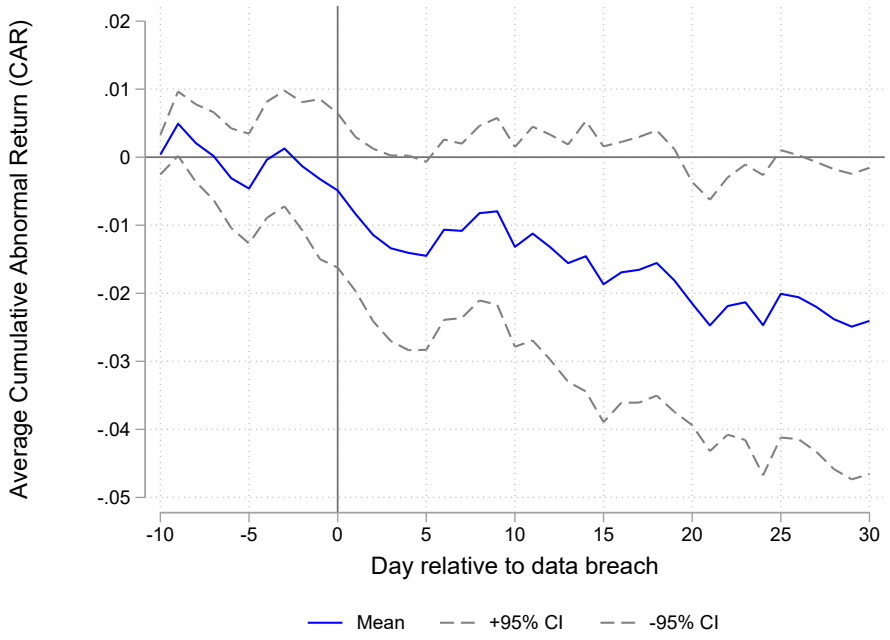
Electronic copy available at: https://ssrn.com/abstract=3143740

Li, J. and Wu, D. (2020). Do corporate social responsibility engagements lead to real environmental, social, and governance impact? *Management Science*, 66(6):2564–2588.
Liang, H. and Renneboog, L. (2017). On the Foundations of Corporate Social Responsibility. *Journal of Finance*, 72(2):853–910.
Lins, K. V., Servaes, H., and Tamayo, A. (2017). Social Capital, Trust, and Firm Performance: The Value of Corporate Social Responsibility during the Financial Crisis. *Journal of Finance*, 72(4):1785–1824.
Liu, Y. and Shankar, V. (2015). The Dynamic Impact of Product-Harm Crises on Brand Preference and Advertising Effectiveness: An Empirical Analysis of the Automobile Industry. *Management Science*, 61(10):2514–2535.
Loughran, T. and McDonald, B. (2014). Measuring Readability in Financial Disclosures. *Journal of Finance*, 69(4):1643–1671.
Mailath, G. J. and Samuelson, L. (2001). Who wants a good reputation? *Review of Economic Studies*, 68(2):415–441.
Makridis, C. A. and Dean, B. (2018). Measuring the Economic Effects of Data Breaches on Firm Outcomes: Challenges and Opportunities. *Journal of Economic and Social Measurement*, 43:59–83.
Maksimovic, V. and Titman, S. (1991). Financial policy and reputation for product quality. *Review of Financial Studies*, 4:175–200.
Margolis, J. D., Elfeinbein, H. A., and Walsh, J. P. (2009). Does it Pay to Be Good...And Does it Matter? A Meta-Analysis of the Relationship between Corporate Social and Financial Performance. Working Paper, Harvard University.
Marinovic, I., Skryzpacz, A., and Varas, F. (2018). Dynamic certification and reputation for quality. *American Economic Journal: Microeconomics*, 10(2):58–82.
Masulis, R. W. and Reza, S. W. (2015). Agency problems of corporate philanthropy. *The Review of Financial Studies*, 28(2):592–636.
Meier, J.-M., Servaes, H., Wei, J., and Xiao, S. C. (2023). Do consumers care about ESG? evidence from barcode-level sales data. Working Paper, SSRN ID 4260716.
Milgrom, P. and Roberts, J. (1982). Predation, reputation, and entry deterrence. *Journal of Economic Theory*, 27:280–312.
Mizik, N. and Jacobson, R. (2008). The financial value impact of perceptual brand attributes. *Journal of Marketing Research*, 45(1):15–32.
Murphy, D. L. and Shrieves, R. E. (2009). Determinants of the Stock Price Reaction to Allegations of Corporate Misconduct: Earnings, Risk, and Firm Size Effects. *Journal of Financial and Quantitative Analysis*, 43.3:851–612.
Naidoo, C. and Abratt, R. (2018). Brands that do good: insight into social brand equity. *Journal of Brand Management*, 25:3–13.
Noe, T. (2012). A survey of the economic theory of reputation: Its logic and limits. In Barnett, M. L. and Pollock, T. G., editors, *The Oxford Handbook of Corporate Reputation*, chapter 6, pages 114–139. Oxford University Press.
Noe, T. H., Rebello, M. J., and Rietz, T. A. (2012). Product market efficiency: The bright side of myopic, uninformed, and passive external finance. *Management Science*, 58:2019–2036.
Oster, E. (2019). Unobservable Selection and Coefficient Stability: Theory and Evidence. *Journal of Business & Economic Statistics*, 37(2):187–204.
Ovtchinnikov, A. V., Reza, S. W., and Wu, Y. (2020). Political activism and firm innovation. *Journal of Financial and Quantitative Analysis*, 55(3):989–1024.
Rehman, Z. u., Khan, A., and Rahman, A. (2020). Corporate social responsibility's influence on firm risk and firm performance: the mediating role of firm reputation. *Corporate Social Responsibility and Environmental Management*, 27(6):2991–3005.
Rhee, M. and Kim, T. (2012). After the collapse: A behavioral theory of reputatoin repair. In Barnett, M. L. and Pollock, T. G., editors, *The Oxford Handbook of Corporate Reputation*, chapter 6, pages 445–464. Oxford University Press.
Rice, A. B. and Schiller, C. (2023). When values align: Corporate philanthropy and employee turnover. *Working Paper, Available at SSRN 4172414*.
Saunders, A. and Tambe, P. (2015). Data Assets and Industry Competition: Evidence from 10-K Filings. Working paper, Wharton.
Servaes, H. and Tamayo, A. (2013). The Impact of Corporate Social Responsibility on Firm Value: The Role of Customer Awareness. *Management Science*, 59(5):1045–1061.
Spanos, G. and Angelis, L. (2016). The Impact of Information Security Events to the Stock Market: A Systematic Literature Review. *Computers & Security*, 58:216–229.
Sun, L. and Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2):175–199.
Tambe, P., Hitt, L. M., and Brynjolfsson, E. (2012). The Extroverted Firm: How External Information Practices Affect Innovation and Productivity. *Management Science*, 58:843–859.
Tirole, J. (1996). A theory of collective reputations (with applications to the persistence of corruption and to firm quality). *Review of Economic Studies*, 63(1):1–22.
Tversky, A. and Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review*, 90(4):293–315.
Vanhamme, J. and Grobben, B. (2009). 'Too Good to be True!' The Effectiveness of CSR History in Countering Negative Publicity. *Journal of Business Ethics*, 85(2):273–283.
Viceira, L. (2020). Moral Values and Corporate Social Responsibility. Harvard Kennedy School Mossavar-Rahmani Center Working paper 144.
Wartick, S. L. (1992). The relationship between intense media exposure and change in corporate reputation. *Business and Society*, 31:33–49.
Xiao, Z., Zheng, X., and Zheng, Y. (2023). The economic and financial impact of negative environmental and social practices: Evidence from consumers store visits. Working Paper, SSRN 4475050.

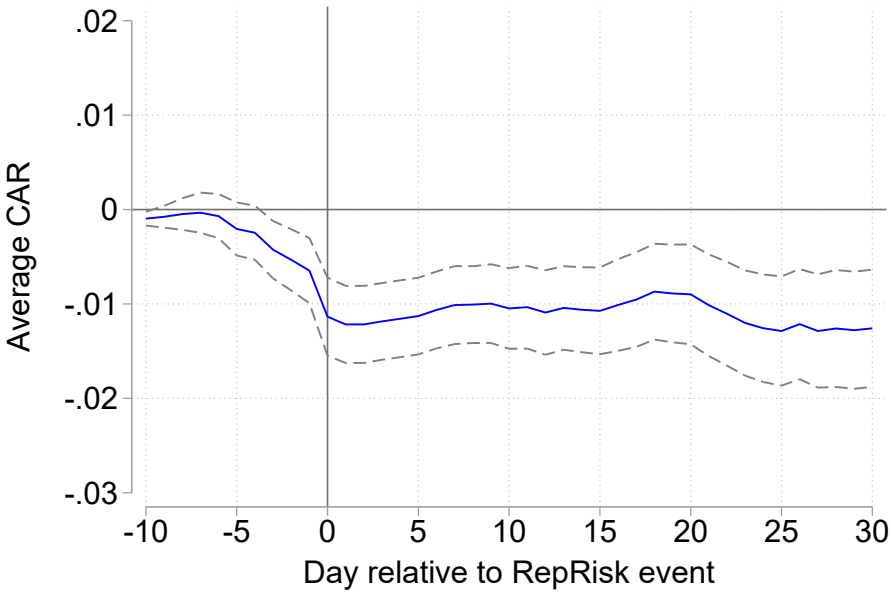
Electronic copy available at: https://ssrn.com/abstract=3143740

Figure 1: Cumulative Abnormal Returns around Reputation Shocks

(a) Data Breaches



(b) RepRisk Events

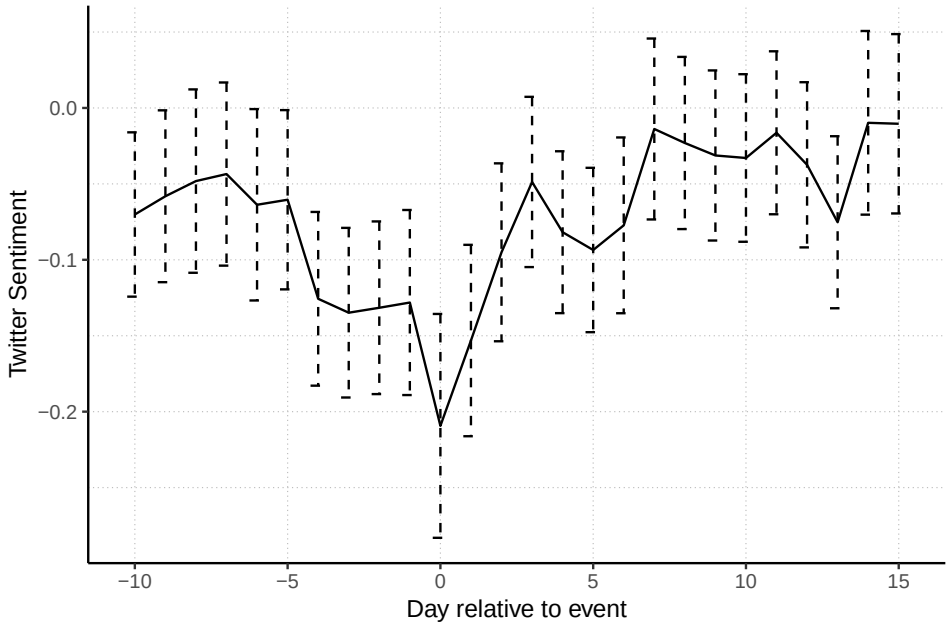


Notes: This figure plots the average cumulative abnormal returns (CARs) and 95% confidence interval around the disclosure of a data breach (Fig. 1a) and the occurrence of a RepRisk event (Fig. 1b) for the $[-10; 30]$ day event window around the reputation shock. We include only data breaches for which the number of affected records is greater than 1000 and novel RepRisk events (i.e., no duplicates) which are categorized by RepRisk as 'high reach' and 'high severity'. CARs are computed using the Fama-French three factor model.

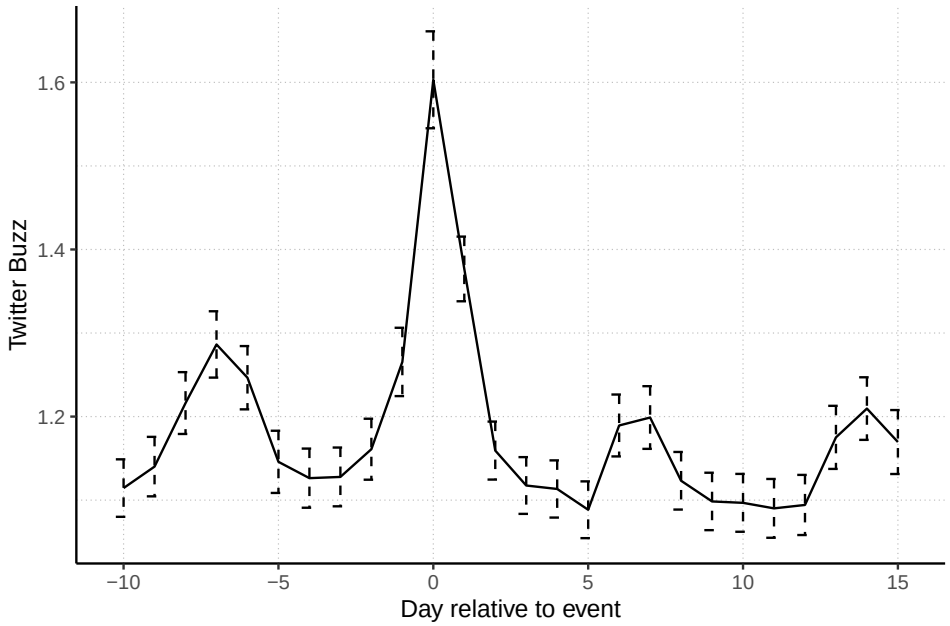
Electronic copy available at: https://ssrn.com/abstract=3143740

Figure 2: Social Media Response to RepRisk Events

(a) Twitter Sentiment



(b) Twitter Buzz

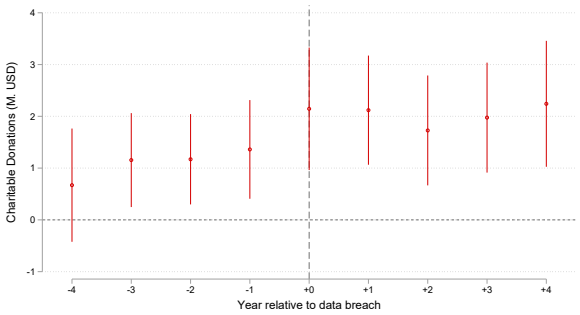


Notes: This figure displays the evolution of social media sentiment (Fig. 2a) and social media "buzz" (Fig. 2b) around the occurrence of RepRisk events at the daily frequency. Social media measures are based on Twitter posts and provided by the social media analytics firm 'Social Market Analytics' (SMA). 'Sentiment' is a text-based measure of positive vs. negative sentiment in the Tweets posted about our sample firms on a given day. "Buzz" measures the abnormal volume in Twitter posts after correcting for overall Tweet volume on a given day. Both figures show the average and 95% confidence interval on a given day relative to the RepRisk event date.

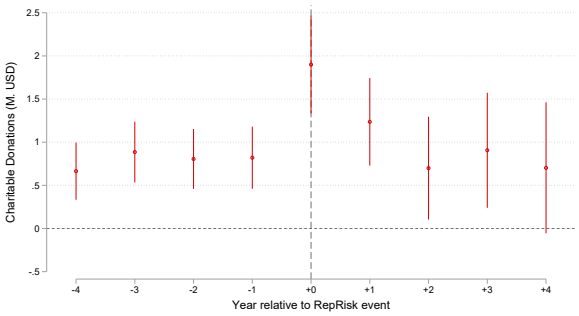
Electronic copy available at: https://ssrn.com/abstract=3143740

Figure 3: Dynamic Effects of Reputation Shocks on CSR Responses

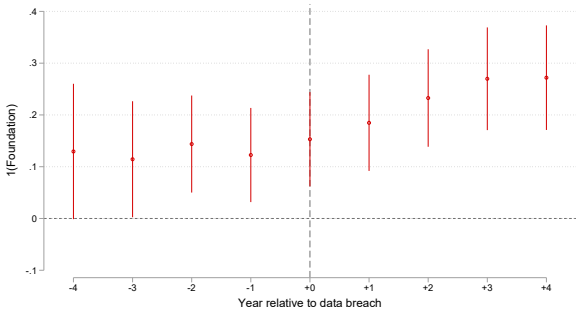
(a) Data Breaches – Donations



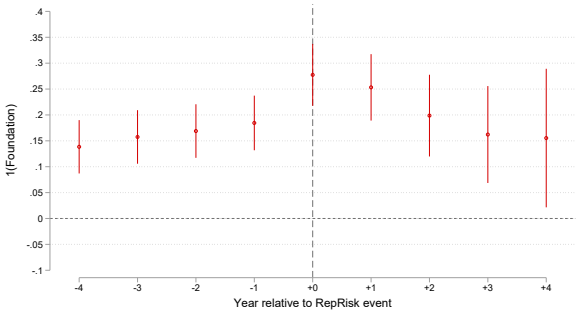
(b) RepRisk Events – Donations



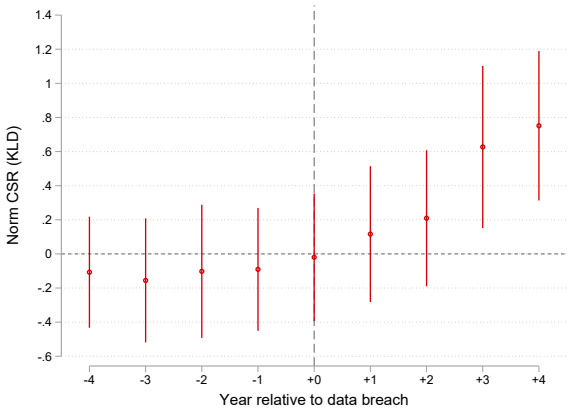
(c) Data Breaches – 1(Foundation)



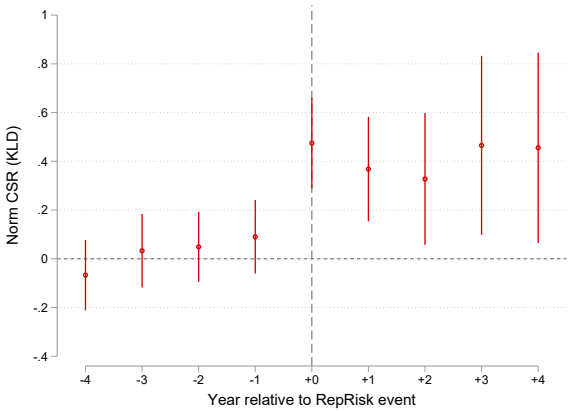
(d) RepRisk Events – 1(Foundation)



(e) Data Breaches – CSR (KLD)



(f) RepRisk Events – CSR (KLD)



Notes: These figures show the dynamic effects of data breaches and RepRisk events, respectively, on firms' CSR responses, i.e., corporate charitable donations (Fig. 3a and 3b), the presence of a corporate charitable foundation (Fig. 3c and 3d), and CSR scores from KLD (Fig. 3e and 3f). Each figure plots the coefficient estimates and corresponding confidence intervals for a regression on dummy variables indicating the distance (in years) relative to the data breach. In all figures, the estimation includes a treatment indicator, industry-by-year-quarter fixed effects and controls for $ln(Assets)$, $ln(Assets)^2$, and *market leverage*.

Electronic copy available at: https://ssrn.com/abstract=3143740

Table 1: Determinants of Reputation Shocks

Notes: This table summarizes OLS regressions of the occurrence of large data breaches (Panel 1a) and RepRisk events (Panel 1b) on CSR scores and other firm characteristics. The outcome variable in Panel 1a is an indicator variable that takes the value of one if the firm suffers a data breach with at least 1,000 compromised records, and zero otherwise. The dependent variable in Panel 1b is a dummy that indicates the occurrence of a 'novel' and 'high reach' RepRisk event (as classified as by RepRisk), in the given year. All firm are described in Table IA.1 and measured as of the year prior to the reputation shock. Firms are only included if there has ever been a data breach or RepRisk event, respectively, in their six-digit GIC industry. Compustat variables have been Winsorized at the 5th percentiles. Year-by-industry fixed effects ("Yr×GIC FE") and firm fixed effects are included as indicated. For ease of readability we suppress the intercept in column (1). Standard errors are clustered at the firm level and reported in parentheses. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

(a) Data Breaches

	Dependent Variable: Large Breach (0/1)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Norm CSR			.0021	.00266	.00274		.00187
			(.00151)	(.00181)	(.0021)		(.00195)
E-Index				.000146	.000265		.000838
				(.0022)	(.00279)		(.00302)
G-Index				-.0008	-.00104		-.00113
				(.00141)	(.0018)		(.00183)
IT Security (0/1)					-.0122**	-.00352*	-.0113**
					(.00498)	(.00203)	(.00499)
IT Investment (0/1)					.00542	.00065	.00462
					(.0129)	(.00697)	(.0133)
1(Risk Committee)					-.0143	-.0113*	-.0164
					(.0154)	(.0064)	(.0164)
1(Dual Role CEO)					-.00191	.0000381	-.00116
					(.00433)	(.00179)	(.00442)
Inst. Block Own.					.0352	.00592	.0262
					(.0449)	(.0103)	(.043)
ln(Assets)	-.00226***	-.00262***	-.0229***	-.0526***	-.0617***	-.0123***	-.0598***
	(.000677)	(.000564)	(.0088)	(.0188)	(.0228)	(.00433)	(.0226)
ln(Assets)^2	.000289***	.000361***	.00189***	.00352***	.00434***	.00124***	.00396***
	(.0000655)	(.0000732)	(.000635)	(.00119)	(.00149)	(.000396)	(.00148)
Leverage	-.00306***	-.000837	-.00641	-.000693	-.0075	-.00429	-.00582
	(.000742)	(.00101)	(.00582)	(.0106)	(.0137)	(.00372)	(.0154)
ROA	.00113***	.000172	.00238	.00852	.0139	-.00123	-.00691
	(.000308)	(.000261)	(.0024)	(.00695)	(.0103)	(.00179)	(.0145)
M/B	-3.50e-07	-.000306***	-.00105***	-.00221***	-.00309***	-.000919***	-.0035***
	(.0000558)	(.0000798)	(.000375)	(.000756)	(.001)	(.000294)	(.00125)
Log(Firm Age)						-.00139	.0101
						(.00469)	(.0158)
Tobin's Q (t-1)						-.000107	.00118
						(.000399)	(.00248)
Sales Growth						-.000107	-.000967
						(.00027)	(.00476)
BHAR 12-Mths						.000804	.00323
						(.000524)	(.00434)
1(Financial Constraint)						-.00188*	-.00761*
						(.00098)	(.00414)
Return Volatilty						-.0166	-.16
						(.0282)	(.251)
R&D/Assets						-.00723	-.0521
						(.0186)	(.0953)
CAPX/Assets						-.00113	.0197
						(.0107)	(.04)
Asset intangibility						-.000113	.00351
						(.00697)	(.0285)
1(S&P500)						.00148	.01
						(.00524)	(.00681)
Yr × GIC FE	No	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	No	Yes	Yes	Yes	Yes	Yes	Yes
Observations	74464	73010	22036	13108	9646	25590	8767
R^2	.00951	.13	.161	.192	.217	.159	.209
Within-R^2	.	.000935	.00251	.00381	.00544	.00224	.00537

Electronic copy available at: https://ssrn.com/abstract=3143740

. . . continued

(b) RepRisk Events

	Dependent Variable: RRI Event (0/1)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Norm CSR (KLD)			.00148*	.00186*	.00183		.00194
			(.000877)	(.00105)	(.00112)		(.00121)
E-Index				-.000967	-.000735		-.000607
				(.000749)	(.000765)		(.000765)
G-Index				.000753	.000912*		.00084
				(.000501)	(.000518)		(.000529)
1(Risk Committee)					-.00194	-.00091	-.00188
					(.00186)	(.000579)	(.00208)
1(Dual Role CEO)					.000736	.000136	.000701
					(.00086)	(.000302)	(.00093)
Inst. Block Own.					-.00986	-.000777	-.00986
					(.0141)	(.00283)	(.0156)
ln(Assets)	-.000574**	-.000682***	-.0103***	-.0221**	-.0188**	-.00386**	-.0177*
	(.000224)	(.000252)	(.00397)	(.0092)	(.00884)	(.00161)	(.00929)
ln(Assets)^2	.0000683***	.0000647**	.000688**	.0013**	.00109**	.000289**	.00103*
	(.0000246)	(.0000259)	(.00027)	(.000557)	(.000538)	(.000128)	(.000575)
Leverage	-.00103**	.000288	.00264	.00757*	.0056	.001	.00787*
	(.000463)	(.000382)	(.00202)	(.00406)	(.00362)	(.00107)	(.00414)
ROA	.000282**	.000393***	.00403***	.00785***	.00767***	.00197***	.00751**
	(.000121)	(.000123)	(.00125)	(.00264)	(.0027)	(.000671)	(.00346)
M/B	-.0000224	-.000029	-.000107	-.000141	-.0000203	-.0000164	-.0000349
	(.0000192)	(.0000277)	(.000131)	(.000263)	(.000273)	(.0000628)	(.000284)
Log(Firm Age)						-.00362***	-.00809
						(.00128)	(.00533)
Tobin's Q (t-1)						-.000011	.0000357
						(.0000487)	(.00028)
Sales Growth						-.0000198	.00126
						(.000061)	(.000783)
BHAR 12-Mths						-6.75e-07	-.0000935
						(.000101)	(.000829)
1(Financial Constraint)						.000117	.000987
						(.000196)	(.000997)
Return Volatilty						-.0167**	-.154**
						(.00775)	(.0734)
R&D/Assets						.000702	.0096
						(.00161)	(.01)
CAPX/Assets						.000834	.00355
						(.00205)	(.00786)
Asset Intangibility						.000988	.0105
						(.00231)	(.0107)
1(S&P500)						-.000772	-.00183
						(.00141)	(.00211)
Yr × GIC FE	No	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	No	Yes	Yes	Yes	Yes	Yes	Yes
Observations	96913	95059	28687	17639	15827	42388	14689
R^2	.00274	.158	.191	.214	.237	.19	.238
Within-R^2	.	.000164	.00227	.00357	.00318	.000732	.00411

Electronic copy available at: https://ssrn.com/abstract=3143740

Table 2: Data Breaches and Firm Reputation

Notes: This table presents OLS regression results for the effect of data breaches on firm reputation across various dimensions. The dependent variables are dummy variables that indicate the mention of "data breach", "reputation", and "CSR" terms, respectively in the presentation and Q&A section of the firm's earnings conference calls in Panels 2a and 2b, local and national newspaper sentiment (averaged at the annual frequency) from Ravenpack Edge in Panel 2c, "brand strength" from the Brand Asset Valuator (BAV) Model in Panel 2d, and the market-to-book (M/B) ratio in Panel 2e. "Years 0-1 Post" is an indicator variable that takes the value of one if a firm has disclosed a data breach in the current or previous year, and zero otherwise. Similarly, "Years 0-4 Post" indicates whether a firm has disclosed a data breach within the past five years. "Treated" takes the value of one if a firm was ever affected by a data breach, and zero otherwise. Data breaches are included if the number of affected records is known and is at least 1,000. Firms are only included if there has ever been a data breach in their respective six-digit GIC industry. Controls include ln(Assets), ln(Assets)^2, and market leverage. Compustat variables have been Winsorized at the 5th percentiles. Year fixed effects, industry fixed effects (GIC), Year-by-industry fixed effects ("Yr×GIC FE"), and firm fixed effects are included as indicated. Standard errors are clustered at the firm level and reported in parentheses. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

(a) Conference Calls – Presentation

	1(Data Breach)				1(Reputation)				1(CSR)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Years 0-1 Post	.0363**	.0301**			.0776***	.0632**			.095***	.0786**		
	(.0156)	(.0143)			(.0289)	(.0285)			(.035)	(.0344)		
Years 0-4 Post			.0301**	.0233**			.0817***	.0659**			.13***	.105***
			(.014)	(.0118)			(.0276)	(.0272)			(.0321)	(.0314)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Yr × GIC FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Observations	30341	30223	30341	30223	30341	30223	30341	30223	30341	30223	30341	30223
R^2	0.049	0.219	0.049	0.219	0.164	0.357	0.164	0.357	0.164	0.345	0.164	0.345

(b) Conference Calls – Q&A

	1(Data Breach)				1(Reputation)				1(CSR)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Years 0-1 Post	.00423	.000547			-.0381	-.0464			.0756**	.062*		
	(.00748)	(.00788)			(.0366)	(.0347)			(.0333)	(.0336)		
Years 0-4 Post			.00614	.0025			.0272	.0177			.103***	.0818***
			(.00892)	(.00934)			(.031)	(.0287)			(.0276)	(.0271)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Yr × GIC FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Observations	30341	30223	30341	30223	30341	30223	30341	30223	30341	30223	30341	30223
R^2	0.035	0.186	0.035	0.186	0.129	0.292	0.129	0.292	0.261	0.408	0.261	0.408

Electronic copy available at: https://ssrn.com/abstract=3143740

. . . continued

(c) News Sentiment from Ravenpack Edge

	Dep. Variable: Regional News Sentiment					
	(1)	(2)	(3)	(4)	(5)	(6)
Years 0-1 Post	-.0278***	-.0281***	-.0326***			
	(.01)	(.00967)	(.00968)			
Years 0-4 Post				-.0253***	-.0183**	-.0212**
				(.00967)	(.0093)	(.00965)
Treated	.0119*	-.00263		.0183**	.000043	
	(.00701)	(.00679)		(.00789)	(.00801)	
Controls	No	Yes	Yes	No	Yes	Yes
Year FE	Yes	No	No	Yes	No	No
Yr × GIC FE	No	Yes	Yes	No	Yes	Yes
Firm FE	No	No	Yes	No	No	Yes
Observations	18261	18054	17084	18261	18054	17084
R^2	0.052	0.124	0.394	0.052	0.124	0.394

	Dep. Variable: National News Sentiment					
	(1)	(2)	(3)	(4)	(5)	(6)
Years 0-1 Post	-.017	-.00966	-.0112			
	(.0109)	(.0104)	(.0107)			
Years 0-4 Post				-.0222**	-.0118	-.0105
				(.00973)	(.00881)	(.00951)
Treated	.0117**	.00722		.0177***	.0103	
	(.00585)	(.00563)		(.00666)	(.00638)	
Controls	No	Yes	Yes	No	Yes	Yes
Year FE	Yes	No	No	Yes	No	No
Yr × GIC FE	No	Yes	Yes	No	Yes	Yes
Firm FE	No	No	Yes	No	No	Yes
Observations	20279	20018	19087	20279	20018	19087
R^2	0.057	0.137	0.398	0.057	0.137	0.398

(d) BAV Brand Strength

	Dependent Variable: Brand Strength					
	(1)	(2)	(3)	(4)	(5)	(6)
Years 0-1 Post	-8.28***	-5.19***	-2.47**			
	(2.56)	(1.93)	(1.22)			
Years 0-4 Post				-8.6***	-6.77***	-4.04***
				(2.92)	(2.18)	(1.52)
Treated	2.16	5.15**		3.72	6.58***	
	(3.3)	(2.38)		(3.49)	(2.5)	
Controls	No	Yes	Yes	No	Yes	Yes
Year FE	Yes	No	No	Yes	No	No
Yr × GIC FE	No	Yes	Yes	No	Yes	Yes
Firm FE	No	No	Yes	No	No	Yes
Observations	4522	4183	4119	4522	4183	4119
R^2	0.005	0.396	0.819	0.006	0.397	0.819

(e) Valuation – M/B

	Dep. Variable: M/B					
	(1)	(2)	(3)	(4)	(5)	(6)
Years 0-1 Post	-.715***	-.855***	-.49***			
	(.15)	(.151)	(.116)			
Years 0-4 Post				-.525***	-.598***	-.267*
				(.175)	(.164)	(.145)
Treated	.643***	.75***		.696***	.804***	
	(.193)	(.161)		(.197)	(.16)	
Controls	No	Yes	Yes	No	Yes	Yes
Year FE	Yes	No	No	Yes	No	No
Yr × GIC FE	No	Yes	Yes	No	Yes	Yes
Firm FE	No	No	Yes	No	No	Yes
Observations	74898	74592	73152	74898	74592	73152
R^2	0.023	0.274	0.666	0.023	0.274	0.666

Electronic copy available at: https://ssrn.com/abstract=3143740

Table 3: RepRisk Events and Firm Reputation

Notes: This table presents OLS regression results for the effect of RepRisk events on firm reputation across various dimensions. The dependent variables are dummy variables that indicate the mention of "reputation" and "CSR" terms, respectively, in the presentation and Q&A section of the firm's earnings conference calls in Panels 3a and 3b, local and national newspaper sentiment (averaged at the annual frequency) from Ravenpack Edge in Panel 3c, "customer churn", i.e., the number of existing consumers who did not buy from this firm again, from the customer churn data of Baker et al. (2023) in Panel 3d, and the market-to-book (M/B) ratio in Panel 3e. "Years 0-1 Post" is an indicator variable that takes the value of one if a firm has disclosed a data breach in the current or previous year, and zero otherwise. Independent variables, controls, and fixed effects are similar to Table 2. Panel 3d is organized at the quarterly frequency and therefore includes dummies that indicate the occurrence of a RepRisk event in the current and previous 1, 4, and 16 quarters, respectively, and time-fixed effects defined at the quarterly frequency. Standard errors are clustered at the firm level and reported in parentheses. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

(a) Conference Calls – Presentation

	1(Reputation)				1(CSR)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Years 0-1 Post	.0406**	.0116			.0665***	.0538***		
	(.0172)	(.016)			(.0194)	(.0181)		
Years 0-4 Post			.0393**	.0135			.0788***	.0637***
			(.0168)	(.016)			(.019)	(.0181)
Treated	.0302***		.0285***		.00122		-.00508	
	(.0102)		(.0102)		(.0107)		(.0106)	
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Yr × GIC FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	No	Yes	No	Yes	No	Yes	No	Yes
Observations	40277	40140	40277	40140	40277	40140	40277	40140
R^2	0.155	0.341	0.155	0.341	0.169	0.347	0.169	0.347

(b) Conference Calls – Q&A

	1(Reputation)				1(CSR)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Years 0-1 Post	.038**	.0257			.124***	.115***		
	(.0166)	(.0158)			(.0162)	(.0152)		
Years 0-4 Post			.046***	.0365**			.12***	.116***
			(.0158)	(.0151)			(.016)	(.0153)
Treated	.0295***		.0256***		.00982		.00459	
	(.0088)		(.00877)		(.0102)		(.0104)	
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Yr × GIC FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	No	Yes	No	Yes	No	Yes	No	Yes
Observations	40277	40140	40277	40140	40277	40140	40277	40140
R^2	0.126	0.282	0.126	0.282	0.273	0.413	0.273	0.413

Electronic copy available at: https://ssrn.com/abstract=3143740

. . . continued

(c) Ravenpack News Sentiment

	Dep. Variable: Regional News Sentiment					
	(1)	(2)	(3)	(4)	(5)	(6)
Years 0-1 Post	-.0127**	-.021***	-.00169			
	(.00532)	(.0052)	(.00563)			
Years 0-4 Post				-.011**	-.0164***	.00587
				(.00544)	(.00528)	(.00622)
Treated	.00427	-.00847*		.00452	-.00851*	
	(.00422)	(.00446)		(.00459)	(.00481)	
Controls	No	Yes	Yes	No	Yes	Yes
Year FE	Yes	No	No	Yes	No	No
Year × GIC FE	No	Yes	Yes	No	Yes	Yes
Firm FE	No	No	Yes	No	No	Yes
Observations	23634	23362	22198	23634	23362	22198
R^2	0.051	0.126	0.393	0.051	0.125	0.393

	Dep. Variable: National News Sentiment					
	(1)	(2)	(3)	(4)	(5)	(6)
Years 0-1 Post	-.00951*	-.0164***	.000298			
	(.00495)	(.00469)	(.00506)			
Years 0-4 Post				-.0126**	-.0174***	.0021
				(.00519)	(.00488)	(.00542)
Treated	-.00316	-.0111***		-.0012	-.00948**	
	(.00388)	(.00394)		(.00411)	(.00416)	
Controls	No	Yes	Yes	No	Yes	Yes
Year FE	Yes	No	No	Yes	No	No
Year × GIC FE	No	Yes	Yes	No	Yes	Yes
Firm FE	No	No	Yes	No	No	Yes
Observations	26133	25787	24665	26133	25787	24665
R^2	0.057	0.140	0.395	0.057	0.140	0.395

(d) Customer Churn

	Customer Churn (Existing Customers)					
	(1)	(2)	(3)	(4)	(5)	(6)
Qtrs 0-1 Post	-0.0274	0.0122*				
	(0.0218)	(0.0065)				
Qtrs 0-4 Post			-0.0172	0.0090*		
			(0.0197)	(0.0054)		
Qtrs 0-16 Post					0.0233	0.0140**
					(0.0266)	(0.0063)
Treated	0.0250		0.0256		0.0059	
	(0.0229)		(0.0238)		(0.0279)	
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Yr×Qtr×GIC FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	No	Yes	No	Yes	No	Yes
Observations	4,014	4,014	4,014	4,014	4,014	4,014
R^2	0.5808	0.9582	0.5804	0.9581	0.5807	0.9582

(e) Valuation – M/B

	Dep. Variable: M/B					
	(1)	(2)	(3)	(4)	(5)	(6)
Years 0-1 Post	-.108	-.503***	-.205**			
	(.111)	(.108)	(.0902)			
Years 0-4 Post				-.198*	-.428***	-.179*
				(.114)	(.111)	(.0942)
Treated	.501***	.707***		.528***	.721***	
	(.0948)	(.0874)		(.0959)	(.0875)	
Controls	No	Yes	Yes	No	Yes	Yes
Yr × GIC FE	No	Yes	Yes	No	No	Yes
Firm FE	No	No	Yes	No	No	Yes
Observations	97476	97099	95258	97476	97105	95258
R^2	0.023	0.266	0.666	0.023	0.237	0.666

Electronic copy available at: https://ssrn.com/abstract=3143740

Table 4: Reputation Shocks, IT, and Advertising

Notes: This table presents linear probability and OLS regression results for the effect of reputation shocks, i.e., data breaches in Panel 4a and RepRisk events in Panel 4b, on measures of IT investment and advertising. The dependent variables in both panels are an indicator variable that takes the value of one if the firm mentions terms related to 'IT security' in their 10K (columns 1–6) and advertising expenses scaled by assets (columns 7–12), respectively. Columns 1 through 6 additionally include controls for the length of the firm's 10K (i.e., number of words) and the complexity of the vocabulary (i.e., the number of unique words). "Years 0-1 Post" is an indicator variable that takes the value of one if a firm has disclosed a data breach in the current or previous year, and zero otherwise. Similarly, "Years 0-4 Post" indicates whether a firm has disclosed a data breach within the past five years. "Treated" takes the value of one if a firm was ever affected by a data breach, and zero otherwise. Data breaches are included if the number of affected records is known and is at least 1,000. Firms are only included if there has ever been a data breach in their respective six-digit GIC industry. Controls include ln(Assets), ln(Assets)^2, and market leverage. Compustat variables have been Winsorized at the 5th percentiles. Year fixed effects, industry fixed effects (GIC), Year-by-industry fixed effects ("Yr×GIC FE"), and firm fixed effects are included as indicated. Standard errors are clustered at the firm level and reported in parentheses. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

(a) Data Breaches

	1(IT Security)						Adv/Assets					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Years 0-1 Post	.0649**	.0563*	.0384				-.00108	.0000615	-.000183			
	(.0312)	(.0304)	(.0285)				(.00125)	(.00133)	(.000931)			
Years 0-4 Post				.126***	.121***	.0916***				-.00159	.000297	-.000133
				(.0323)	(.031)	(.0293)				(.00167)	(.00162)	(.00112)
Treated	.183***	.0676***		.151***	.0358		-.00107	-.00169		-.000704	-.00178	
	(.0268)	(.0242)		(.0266)	(.0235)		(.00265)	(.00219)		(.00273)	(.00229)	
Length 10K	1.51***	1.69***	1.59***	1.5***	1.68***	1.57***						
	(.2)	(.234)	(.254)	(.201)	(.234)	(.254)						
10K Vocab. Complexity	-.196*	.247**	.243**	-.197*	.246**	.241**						
	(.114)	(.113)	(.111)	(.114)	(.114)	(.111)						
Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Yr × GIC FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Firm FE	No	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes
Observations	53491	50170	49092	53491	50170	49092	37728	34329	33566	37728	34329	33566
R^2	0.142	0.297	0.631	0.143	0.298	0.631	0.010	0.351	0.834	0.010	0.351	0.834

(b) RepRisk Events

	1(IT Security)						Adv/Assets					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Years 0-1 Post	.0238	.0252	.00599				.000319	.00304*	-.000537			
	(.0278)	(.0254)	(.0219)				(.00204)	(.00172)	(.000954)			
Years 0-4 Post				.135***	.0281	.0129				-5.68e-06	.00283	-.000512
				(.0213)	(.0293)	(.0286)				(.002)	(.00175)	(.00108)
Treated	.0392***	-.000873		.0749***	-.00344		.00908***	.00323		.00915***	.00313	
	(.013)	(.0125)		(.0135)	(.0125)		(.0022)	(.00197)		(.00225)	(.00201)	
Length 10K	1.41***	1.56***	1.31***	1.34***	1.56***	1.31***						
	(.123)	(.188)	(.203)	(.126)	(.188)	(.203)						
10K Vocab. Complexity	-.241***	.206**	.219***	-.285***	.206**	.219***						
	(.0769)	(.0852)	(.0829)	(.0807)	(.0853)	(.0828)						
Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Yr × GIC FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Firm FE	No	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes
Observations	68030	62464	61079	70081	62464	61079	45184	41185	40270	45184	41185	40270
R^2	0.128	0.270	0.617	0.148	0.270	0.617	0.013	0.351	0.837	0.013	0.351	0.837

Electronic copy available at: https://ssrn.com/abstract=3143740

Table 5: CSR Reaction to Reputation Shocks

Notes: This table presents OLS and linear probability regression results for the effect of reputation shocks, i.e., data breaches in Panels 5a, 5c, and 5e and RepRisk events in Panels 5b, 5d, and 5f, on various CSR outcomes. The dependent variables are the amount of charitable donations ($M) made by the firm in the given year (Panels 5a and 5b), an dummy, "1(Has Foundation)" that indicates whether the firm had a corporate charitable foundation in the given year (Panels 5c and 5d), and "Norm CSR (KLD)" (Panels 5e and 5f), which is the CSR score from KLD normalized such that within the full Compustat sample the mean is 0 and the standard deviation is 1, as described in Section 2. "Years 0-1 Post" is an indicator variable that takes the value of one if a firm has disclosed a data breach in the current or previous year, and zero otherwise. Similarly, "Years 0-4 Post" indicates whether a firm has disclosed a data breach within the past five years. "Treated" takes the value of one if a firm was ever affected by a data breach, and zero otherwise. Data filters, controls, and fixed effects are similar as in Tables 2 and 3. Standard errors are clustered at the firm level and reported in parentheses. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

(a) Data Breaches – Donations

	Charitable Donations (M. USD)					
	(1)	(2)	(3)	(4)	(5)	(6)
Years 0-1 Post	1.49***	1.13***	.703***			
	(.441)	(.329)	(.225)			
Years 0-4 Post				1.45***	1.24***	.808***
				(.368)	(.286)	(.217)
Treated	1.64***	.668***		1.41***	.456**	
	(.317)	(.248)		(.279)	(.226)	
Controls	No	Yes	Yes	No	Yes	Yes
Year FE	Yes	No	No	Yes	No	No
Year × GIC FE	No	Yes	Yes	No	Yes	Yes
Firm FE	No	No	Yes	No	No	Yes
Observations	30688	30435	30385	30688	30435	30385
R^2	0.064	0.259	0.716	0.066	0.261	0.717

(b) RepRisk Events – Donations

	Charitable Donations (M. USD)					
	(1)	(2)	(3)	(4)	(5)	(6)
Years 0-1 Post	1.85***	1.41***	1.02***			
	(.251)	(.216)	(.129)			
Years 0-4 Post				1.57***	1.26***	.987***
				(.226)	(.199)	(.125)
Treated	.959***	.393***		.939***	.357***	
	(.12)	(.11)		(.121)	(.109)	
Controls	No	Yes	Yes	No	Yes	Yes
Year FE	Yes	No	No	Yes	No	No
Year × GIC FE	No	Yes	Yes	No	Yes	Yes
Firm FE	No	No	Yes	No	No	Yes
Observations	39759	39438	39372	39759	39438	39372
R^2	0.106	0.262	0.714	0.102	0.261	0.714

(c) Data Breaches – 1(Foundation)

	1(Has Foundation)					
	(1)	(2)	(3)	(4)	(5)	(6)
Years 0-1 Post	.145***	.108***	.0936***			
	(.0307)	(.0273)	(.024)			
Years 0-4 Post				.21***	.192***	.156***
				(.0316)	(.0304)	(.0263)
Treated	.265***	.107***		.222***	.0638**	
	(.0304)	(.0301)		(.0287)	(.029)	
Controls	No	Yes	Yes	No	Yes	Yes
Year FE	Yes	No	No	Yes	No	No
Year × GIC FE	No	Yes	Yes	No	Yes	Yes
Firm FE	No	No	Yes	No	No	Yes
Observations	30688	30435	30385	30688	30435	30385
R^2	0.083	0.245	0.773	0.086	0.247	0.775

(d) RepRisk Events – 1(Foundation)

	1(Has Foundation)					
	(1)	(2)	(3)	(4)	(5)	(6)
Years 0-1 Post	.268***	.208***	.168***			
	(.0194)	(.0188)	(.0131)			
Years 0-4 Post				.246***	.2***	.167***
				(.0208)	(.0204)	(.0144)
Treated	.209***	.0866***		.202***	.078***	
	(.0177)	(.0188)		(.0175)	(.0186)	
Controls	No	Yes	Yes	No	Yes	Yes
Year FE	Yes	No	No	Yes	No	No
Year × GIC FE	No	Yes	Yes	No	Yes	Yes
Firm FE	No	No	Yes	No	No	Yes
Observations	39759	39438	39372	39759	39438	39372
R^2	0.134	0.257	0.788	0.133	0.257	0.788

(e) Data Breaches – CSR

	Norm CSR (KLD)					
	(1)	(2)	(3)	(4)	(5)	(6)
Years 0-1 Post	.318**	.18	.143			
	(.15)	(.137)	(.118)			
Years 0-4 Post				.515***	.425***	.387***
				(.149)	(.145)	(.129)
Treated	.488***	.141		.364***	.0258	
	(.114)	(.112)		(.117)	(.116)	
Controls	No	Yes	Yes	No	Yes	Yes
Year FE	Yes	No	No	Yes	No	No
Yr × GIC FE	No	Yes	Yes	No	Yes	Yes
Firm FE	No	No	Yes	No	No	Yes
Observations	23275	23137	22738	23275	23137	22738
R^2	0.041	0.168	0.605	0.044	0.170	0.607

(f) RepRisk Events – CSR

	Norm CSR (KLD)					
	(1)	(2)	(3)	(4)	(5)	(6)
Years 0-1 Post	.655***	.481***	.409***			
	(.0829)	(.0751)	(.0579)			
Years 0-4 Post				.651***	.508***	.476***
				(.0799)	(.0744)	(.059)
Treated	.189***	-.0261		.162***	-.0544	
	(.0528)	(.0528)		(.0529)	(.0531)	
Controls	No	Yes	Yes	No	Yes	Yes
Year FE	Yes	No	No	Yes	No	No
Yr × GIC FE	No	Yes	Yes	No	Yes	Yes
Firm FE	No	No	Yes	No	No	Yes
Observations	30238	30071	29580	30238	30071	29580
R^2	0.055	0.186	0.613	0.056	0.187	0.614

Electronic copy available at: https://ssrn.com/abstract=3143740

Table 6: Reputation Shocks and Political Contributions

Notes: This table presents OLS regression results for the effect of reputation shocks, i.e., data breaches in Panel 6a and RepRisk events in Panel 6b, on firms' political contributions. The dependent variable is the amount of political contributions ($M) by the firm and the firm's political action committees (PAC) in a given election cycle (which has a length of two years). The data in both panels is organized at the firm-by-election-cycle frequency. Data on political contributions are obtained from the Federal Election Commission, and represent the total dollar amount of contributions to candidates for the U.S. House of Representatives and U.S. Senate. "Years 0-1 Post" and "Years 0-4 Post", which indicate the occurrence of a data breach (Panel 6a) or RepRisk Event (Panel 6b) in the current and previous years, is collapsed at the election-cycle frequency. All other data filters, controls, and fixed effects are similar as in Tables 2 and 3. Standard errors are clustered at the firm level and reported in parentheses. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

(a) Data Breaches

	Political Contributions (M.USD)					
	(1)	(2)	(3)	(4)	(5)	(6)
Years 0-1 Post	0.2582**	0.1730	0.1064**			
	(0.1206)	(0.1096)	(0.0524)			
Years 0-4 Post				0.1976**	0.1753**	0.1332***
				(0.0839)	(0.0739)	(0.0476)
Treated	0.2880***	0.1405***		0.2654***	0.1097***	
	(0.0658)	(0.0450)		(0.0620)	(0.0405)	
Controls	No	Yes	Yes	No	Yes	Yes
Cycle FE	Yes	No	No	Yes	No	No
Cycle × GIC FE	No	Yes	Yes	No	Yes	Yes
Firm FE	No	No	Yes	No	No	Yes
Observations	4,147	4,136	4,136	4,147	4,136	4,136
R^2	0.1090	0.4641	0.8599	0.1096	0.4664	0.8613

(b) RepRisk Events

	Political Contributions (M.USD)					
	(1)	(2)	(3)	(4)	(5)	(6)
Years 0-1 Post	0.2300***	0.1762***	0.1116***			
	(0.0423)	(0.0360)	(0.0241)			
Years 0-4 Post				0.1914***	0.1555***	0.1088***
				(0.0421)	(0.0367)	(0.0258)
Treated	0.1676***	0.0669***		0.1684***	0.0638***	
	(0.0226)	(0.0159)		(0.0232)	(0.0162)	
Controls	No	Yes	Yes	No	Yes	Yes
Cycle FE	Yes	No	No	Yes	No	No
Cycle × GIC FE	No	Yes	Yes	No	Yes	Yes
Firm FE	No	No	Yes	No	No	Yes
Observations	5,531	5,505	5,505	5,531	5,505	5,505
R^2	0.1457	0.4774	0.8626	0.1405	0.4759	0.8626

Electronic copy available at: https://ssrn.com/abstract=3143740

Table 7: Reputation Shocks and Employee Salaries

Notes: This table presents OLS regression results for the effect of reputation shocks, i.e., data breaches in Panel 7a and RepRisk events in Panel 7b, on employee salaries. The dependent variable is the employee's annual salary (in thousands of $). In both Panels, the data is organized at the individual employee-by-year level. "Years 0-1 Post" ("Years 0-4 Post") is an indicator variable that takes the value of one if the employer of the given employee has disclosed a data breach (Panel 7a) or RepRisk event (Panel 7b), respectively, in the current or previous one (four) years, and zero otherwise. Each regression includes controls for the job experience (self-reported in number of years) of a given employee in addition to firm-level controls for ln(Assets), ln(Assets)^2, and market leverage. In addition to year-by-industry (GIC) and firm fixed effects, we include employee-level fixed effects for the metro area, occupation group, highest degree (professional, post-graduate, undergraduate, etc.), and gender of the given employee. All employee-level data are self-reported and obtained from Glassdoor.com. Standard errors are clustered at the firm level and reported in parentheses. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

(a) Data Breaches

	Annual Salary (Thsd. USD)			
	(1)	(2)	(3)	(4)
Years 0-1 Post	0.6225**	0.6716**		
	(0.2666)	(0.3255)		
Years 0-4 Post			1.194***	1.093***
			(0.2395)	(0.2576)
Experience (Yrs)	1.874***	1.926***	1.874***	1.926***
	(0.0374)	(0.0375)	(0.0374)	(0.0375)
Controls	Yes	Yes	Yes	Yes
Yr × GIC FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Metro Area FE	Yes	Yes	Yes	Yes
Occupation FE	Yes	Yes	Yes	Yes
Highest Degree FE	No	Yes	No	Yes
Gender FE	No	Yes	No	Yes
Observations	449,716	185,878	449,716	185,878
R^2	0.7172	0.7448	0.7173	0.7449

(b) RepRisk Events

	Annual Salary (Thsd. USD)			
	(1)	(2)	(3)	(4)
Years 0-1 Post	0.2105	0.0891		
	(0.1991)	(0.2445)		
Years 0-4 Post			0.1880	-0.1544
			(0.2513)	(0.3041)
Experience (Yrs)	1.861***	1.914***	1.861***	1.914***
	(0.0338)	(0.0340)	(0.0338)	(0.0340)
Controls	Yes	Yes	Yes	Yes
Yr × GIC FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Metro Area FE	Yes	Yes	Yes	Yes
Occupation FE	Yes	Yes	Yes	Yes
Highest Degree FE	No	Yes	No	Yes
Gender FE	No	Yes	No	Yes
Observations	504,725	207,617	504,725	207,617
R^2	0.7134	0.7403	0.7134	0.7403

Electronic copy available at: https://ssrn.com/abstract=3143740

Table 8: CSR Response in Consumer- and Business-facing Industries

Notes: This table presents OLS and linear probability model estimates for the heterogeneous effect of reputation shocks, i.e., data breaches in Panel 8a and RepRisk events in Panel 8b, on CSR responses across consumer- and business facing industries. The dependent variables are corporate charitable donations ($M) (columns 1–4), an indicator taking the value of one if the firm has a private charitable foundation (columns 5–8), and the CSR score from KLD (columns 9–12), respectively. "Years 0-1 Post" ("Years 0-4 Post") is an indicator variable that takes the value of one if the firm has disclosed a data breach (Panel 8a) or RepRisk event (Panel 8b), respectively, in the current or previous one (four) years, and zero otherwise. We interact "Years 0-1 Post" and "Years 0-4 Post" with dummy variables that indicate consumer-facing and business-facing industries. "Has CCus-Ind" is defined at the industry-level and measures the proportion of firms in an industry that have major corporate customers as reported in the Compustat Segment Files database. We define industries below (above) the median as consumer-facing (business-facing) industries. All other data filters, control variables, and fixed effects are similar as in Table 5. 'Chi-Sq(Diff. Low-High)' and 'p(Chi-Sq)' report the results of a χ^2-test comparing coefficient estimates interacted with "Has CCus-Ind = Low" and "Has CCus-Ind = High". *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

(a) Data Breaches

	Donations				1(Foundation)				CSR (KLD)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Years 0-1 × Has CCus-Ind. = Low	1.57***	.891***			.136***	.11***			.186	.118		
	(.449)	(.281)			(.0356)	(.0278)			(.181)	(.148)		
Years 0-1 × Has CCus-Ind. = High	-.155	.177			.0298	.0474			.168	.216		
	(.509)	(.305)			(.0735)	(.0474)			(.212)	(.167)		
Years 0-4 × Has CCus-Ind. = Low			1.64***	1.06***			.212***	.161***			.459***	.403***
			(.401)	(.279)			(.0386)	(.0318)			(.176)	(.154)
Years 0-4 × Has CCus-Ind. = High			.0894	.0867			.135*	.14***			.325	.335
			(.461)	(.219)			(.0696)	(.0458)			(.261)	(.219)
Treated	.671***		.459**		.107***		.0638**		.14		.0263	
	(.249)		(.227)		(.0301)		(.029)		(.113)		(.116)	
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Yr × GIC FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
p(Chi-Sq)	.0268	.0876	.029	.00633	.244	.257	.374	.698	.951	.663	.675	.8
Chi-Sq(Diff. Low-High)	4.91	2.92	4.78	7.46	1.36	1.29	.79	.15	.00376	.19	.175	.0644
Observations	30438	30388	30438	30388	30438	30388	30438	30388	23139	22741	23139	22741
R2	.26	.717	.263	.718	.245	.773	.248	.775	.168	.605	.17	.607

(b) RepRisk Events

	Donations				1(Foundation)				CSR (KLD)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Years 0-1 × Has CCus-Ind. = Low	1.77***	1.33***			.245***	.187***			.416***	.425***		
	(.315)	(.193)			(.0258)	(.0174)			(.0964)	(.0746)		
Years 0-1 × Has CCus-Ind. = High	.973***	.622***			.161***	.142***			.564***	.391***		
	(.351)	(.166)			(.0317)	(.0199)			(.13)	(.0894)		
Years 0-4 × Has CCus-Ind. = Low			1.57***	1.26***			.233***	.182***			.41***	.472***
			(.295)	(.184)			(.0286)	(.0193)			(.094)	(.0742)
Years 0-4 × Has CCus-Ind. = High			.86***	.635***			.156***	.146***			.629***	.481***
			(.31)	(.163)			(.0324)	(.0213)			(.127)	(.093)
Treated	.396***		.362***		.0873***		.0789***		-.0278		-.0566	
	(.11)		(.109)		(.0188)		(.0186)		(.0528)		(.0531)	
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Yr × GIC FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
p(Chi-Sq)	.118	.00649	.119	.0126	.0525	.0904	.0907	.217	.381	.768	.177	.935
Chi-Sq (Diff. Low-High)	2.44	7.42	2.43	6.23	3.76	2.87	2.86	1.52	.767	.0868	1.83	.00669
Observations	39434	39368	39434	39368	39434	39368	39434	39368	30062	29573	30062	29573
R2	.263	.715	.262	.715	.257	.788	.257	.788	.186	.613	.187	.614

Electronic copy available at: https://ssrn.com/abstract=3143740

Electronic copy available at: https://ssrn.com/abstract=3143740

Table 9: Political Contributions in Government-facing and other Firms

Notes: This table presents OLS estimates for the heterogeneous effect of reputation shocks, i.e., data breaches in Panel 9a and RepRisk events in Panel 9b, on political contributions across firms with- and without government customers. The dependent variable is the amount of political contributions ($M) by the firm and the firm's political action committees (PAC) in a given election cycle. Data on political contributions are obtained from the Federal Election Commission, and represent the total dollar amount of contributions to candidates for the U.S. House of Representatives and U.S. Senate. "Years 0-1 Post" and "Years 0-4 Post" and all other control variables, fixed effects, and data filters are defined similarly as in Table 6. We additionally interact "Years 0-1 Post" and "Years 0-4 Post" with a dummy variable that indicates whether the firm has at least one major government customer ("Has GovCus = Yes") or not ("Has GovCus = No"), as reported in the Compustat Segment Files database. 'Chi-Sq(Diff. Y-N)' and 'p(Chi-Sq)' report the results of a χ^2-test comparing coefficient estimates interacted with "Has GovCus = Yes" and "Has GovCus = No". Standard errors are clustered at the firm level and reported in parentheses. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

(a) Data Breaches

	Dep. Var.: Political Contributions			
	(1)	(2)	(3)	(4)
Years 0-1 Post × Has GovCus = Yes	0.7003***	0.3620**		
	(0.1993)	(0.1454)		
Years 0-1 Post × Has GovCus = No	0.1034	0.0720		
	(0.1281)	(0.0534)		
Years 0-4 Post × Has GovCus = Yes			0.6139**	0.4708***
			(0.2419)	(0.1725)
Years 0-4 Post × Has GovCus = No			0.1010	0.0730*
			(0.0831)	(0.0382)
Treated	0.1359***		0.1076***	
	(0.0447)		(0.0404)	
Controls	Yes	Yes	Yes	Yes
Cycle × GIC FE	Yes	Yes	Yes	Yes
Firm FE	No	Yes	No	Yes
Observations	4,135	4,135	4,135	4,135
R^2	0.4684	0.8615	0.4747	0.8650
Chi-Sq(Diff. Y-N)	6.226**	1.777	5.683**	5.627**
p(Chi-Sq)	(0.013)	(0.183)	(0.017)	(0.018)

(b) RepRisk Events

	Dep. Var.: Political Contributions			
	(1)	(2)	(3)	(4)
Years 0-1 Post × Has GovCus = Yes	0.6547***	0.3523***		
	(0.1846)	(0.1141)		
Years 0-1 Post × Has GovCus = No	0.1068***	0.0731***		
	(0.0275)	(0.0179)		
Years 0-4 Post × Has GovCus = Yes			0.5746***	0.3390***
			(0.1748)	(0.1168)
Years 0-4 Post × Has GovCus = No			0.0861***	0.0688***
			(0.0252)	(0.0180)
Treated	0.0646***		0.0623***	
	(0.0159)		(0.0162)	
Controls	Yes	Yes	Yes	Yes
Cycle × GIC FE	Yes	Yes	Yes	Yes
Firm FE	No	Yes	No	Yes
Observations	5,504	5,504	5,504	5,504
R^2	0.4970	0.8676	0.4960	0.8679
Chi-Sq(Diff. Y-N)	7.906***	7.394***	6.773***	8.147***
p(Chi-Sq)	(0.005)	(0.007)	(0.009)	(0.004)

Table 10: Political Contributions after RepRisk 'Violation' and other Events

Notes: This table presents OLS estimates for the heterogeneous effect of RepRisk reputation events on political contributions across different types of RepRisk events.The dependent variable is the amount of political contributions ($M) by the firm and the firm's political action committees (PAC) in a given election cycle. Data on political contributions are obtained from the Federal Election Commission, and represent the total dollar amount of contributions to candidates for the U.S. House of Representatives and U.S. Senate. "Years 0-1 Post" and "Years 0-4 Post" and all other control variables, fixed effects, and data filters are defined similarly as in Table 6. We additionally interact "Years 0-1 Post" and "Years 0-4 Post" with dummy variables that indicates whether the reputation event is classified as a "violation" of a domestic or foreign law, legal statute or agreement (i.e., "RRI-V=Yes") in the RepRisk data or not (i.e., "RRI-V=No"). 'Chi-Sq(Diff. Y-N)' and 'p(Chi-Sq)' report the results of a χ^2-test comparing coefficient estimates interacted with "Yes" and "No". Standard errors are clustered at the firm level and reported in parentheses. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

	Dep. Var.: Political Contributions			
	(1)	(2)	(3)	(4)
Years 0-1 Post × RRI-V = Yes	0.2046***	0.1326***		
	(0.0415)	(0.0266)		
Years 0-1 Post × RRI-V = No	0.1289***	0.0788***		
	(0.0490)	(0.0303)		
Years 0-4 Post × RRI-V = Yes			0.1815***	0.1295***
			(0.0408)	(0.0273)
Years 0-4 Post × RRI-V = No			0.1164**	0.0780**
			(0.0495)	(0.0347)
Treated	0.0667***		0.0635***	
	(0.0158)		(0.0162)	
Controls	Yes	Yes	Yes	Yes
Cycle × GIC FE	Yes	Yes	Yes	Yes
Firm FE	No	Yes	No	Yes
Observations	5,505	5,505	5,505	5,505
R^2	0.4781	0.8629	0.4765	0.8629
Chi-Sq(Diff. Y-N)	1.695	3.663*	1.720	3.257*
p(Chi-Sq)	(0.193)	(0.056)	(0.190)	(0.071)

Electronic copy available at: https://ssrn.com/abstract=3143740

Table 11: Employee Salaries after Employee- and Customer Data Breaches

Notes: This table presents OLS estimates for the heterogeneous effect of data breaches on employee salaries across different types of data breaches. The dependent variable is the annual employee salary (in $K). The data is organized at the individual employee level. "Years 0-1 Post" and "Years 0-4 Post" and all other control variables, fixed effects, and data filters are defined similarly as in Table 7. All employee level variables are self-reported and obtained from Glassdoor.com. Compared to Table 7, we additionally interact "Years 0-1 Post" and "Years 0-4 Post" with dummy variables that indicates whether the data breach affected employee records (i.e., "Empl. Hack=Yes") or not (i.e., "Empl. Hack=No"). 'Chi-Sq(Diff. Y-N)' and 'p(Chi-Sq)' report the results of a χ^2-test comparing coefficient estimates interacted with "Yes" and "No". Standard errors are clustered at the firm level and reported in parentheses. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

	Annual Base Salary (Thsd. USD)			
	(1)	(2)	(3)	(4)
Yr 0-1 Post×1(Empl. Hack=Y)	1.214***	1.116***		
	(0.4037)	(0.4258)		
Yr 0-1 Post×1(Empl. Hack=N)	0.2226	0.2758		
	(0.2944)	(0.3742)		
Yr 0-4 Post×1(Empl. Hack=Y)			1.535***	1.342***
			(0.2702)	(0.2943)
Yr 0-4 Post×1(Empl. Hack=N)			0.6389*	0.5746
			(0.3264)	(0.4212)
Experience (Yrs)	1.874***	1.897***	1.874***	1.896***
	(0.0374)	(0.0370)	(0.0374)	(0.0370)
Controls	Yes	Yes	Yes	Yes
Yr × GIC FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Metro Area FE	Yes	Yes	Yes	Yes
Occupation FE	Yes	Yes	Yes	Yes
Highest Degree FE	No	Yes	No	Yes
Gender FE	No	Yes	No	Yes
Observations	450,004	228,133	450,004	228,133
R^2	0.7171	0.7407	0.7172	0.7407
Chi-Sq.(Diff. Y-N)	5.567**	2.875*	5.797**	2.976*
p(Chi-Sq.)	(0.018)	(0.090)	(0.016)	(0.084)

Electronic copy available at: https://ssrn.com/abstract=3143740

Internet Appendix

Intended for online publication only

Electronic copy available at: https://ssrn.com/abstract=3143740

IA.I Appendix: Model Proofs

Proof of Proposition 1

Taking FOCs of the firm's optimization problem yields the following solutions:

$$\begin{aligned} K^* &= \frac{\rho^2 + p\theta(2 - \rho p)}{2\rho + 4p\theta}, \\ I^* &= \frac{p\left[\rho^2 + 2\theta - \rho\theta(2 - p)\right]}{2\rho + 4p\theta}, \\ R^* &= \frac{2(1 + p\theta) - \rho(1 + p)}{2\rho + 4p\theta}. \end{aligned}$$

The budget constraint is binding, so investments sum to one (after discounting R^*). Under standard assumptions about the parameters ($0 < \rho < 1, 0 < p \leq 1, \theta \geq 0$), all three investments are positive and non-zero by inspection.■

Examining the solutions, we see that if $p = 0$ and there is no potential for a negative event, the firm invests exclusively in physical capital K.[1] In contrast, when $p = 1$ and a negative event is certain, the firm would still invest in all three types of capital: K is efficient because it is not affected by p, I is efficient because it directly offsets the effects of p occurring, and R is efficient due to the time value of money and the efficiency parameter θ. In fact, even when $\theta = 0$, in which case an investment in reputation repair has neither costs nor benefits, there is still a small optimal investment in R due to the time value of money in the budget constraint, though naturally, nearly all investment comes in the forms of physical capital and reputation insurance.

Proof of Proposition 2

To keep the solutions reasonable, we assume p is sufficiently large such that $\rho(2 + p) > 2$. Given this assumption, the proof for proposition 2 is straightforward:

$$\begin{aligned} \frac{\partial K^*}{\partial p} &= \frac{-\rho\theta\left[\theta p^2 + \rho(1 + p) - 1\right]}{(\rho + 2p\theta)^2} < 0, \\ \frac{\partial I^*}{\partial p} &= \frac{1}{4}\left[\rho + \frac{\rho^3 + 4\theta\rho(1 - \rho)}{(\rho + 2p\theta)^2}\right] > 0, \\ \frac{\partial R^*}{\partial p} &= \frac{4\theta(\rho - 1) - \rho^2}{2(\rho + 2p\theta)^2} < 0. \end{aligned}$$

This makes sense: since I is directly attached to p (because it directly reduces the harm caused by an attack), firms will respond to an increase in p by increasing I relative to the other two forms of investment.

[1] The solution is non-zero for reputation repair as well, but the firm would only invest in repair if a negative event occurs.

Electronic copy available at: https://ssrn.com/abstract=3143740

Statics for ρ:

$$\begin{aligned}
\frac{\partial K^*}{\partial \rho} &= \frac{\rho^2 + 4p\rho\theta - 2p\theta(\theta p^2 + 1)}{2(\rho + 2p\theta)^2} <> 0, \\
\frac{\partial I^*}{\partial \rho} &= \frac{p[\rho^2 + 2\theta(2\rho p - 1 - p\theta(2-p))]}{2(\rho + 2p\theta)^2} <> 0, \\
\frac{\partial R^*}{\partial \rho} &= -\frac{p\theta(2+p)+1}{(\rho + 2p\theta)^2} < 0.
\end{aligned}$$

This also makes sense: while the statics for K^* and I^* depend on the parameter values, a higher discount factor (i.e. lower discount rate) reduces the time value of money and makes investing in R less appealing.

Finally, statics for θ:

$$\begin{aligned}
\frac{\partial K^*}{\partial \theta} &= \frac{-p\rho(\rho(2+p)-2)}{2(\rho + 2p\theta)^2} < 0, \\
\frac{\partial I^*}{\partial \theta} &= \frac{-p\rho(\rho(2+p)-2)}{2(\rho + 2p\theta)^2} < 0, \\
\frac{\partial R^*}{\partial \theta} &= \frac{p(\rho(2+p)-2)}{(\rho + 2p\theta)^2} > 0.
\end{aligned}$$

These statics make sense as well: when investing in R is more efficient, the firm will relatively increase its investment in R.

Electronic copy available at: https://ssrn.com/abstract=3143740

IA.II Expanded data description

This section provides additional details on the data sources we use in the paper, as described in Section 2.

IA.II.1 Corporate data breaches

We obtain data on corporate data breaches from the Privacy Rights Clearinghouse (PRC) website.[2] The PRC is a non-profit foundation that advocates to educate consumers about privacy protection. In addition to providing educational services, it has compiled a database of publicly disclosed data breaches starting in 2005. We download the list of breaches that affected private organizations (as opposed to government agencies or universities) and match these organizations to publicly traded firms. The PRC data includes information about which firms were affected by the breach, a short description of the breach and, when available, the number of records that were affected. These data breaches can take several forms, including external hacks, lost or stolen portable devices, insider employees improperly accessing data, physical theft of documents containing information, and inadvertent disclosure of sensitive information. We classify hacks as affecting customer records (such as account information or personal details), employee records, or internal company documents when such information is available (these categories are not mutually exclusive). Panel IA.1a of Table IA.1 reports summary statistics for these variables. Overall, there are 287 data breaches, of which the vast majority represent breaches involving customer records (66%) or employee records (33%). Figures IA.1 and IA.2 in the Internet Appendix provide descriptive data on the frequency of data breaches during our sample, the industries that are affected, the form of the data breaches and the number of records.

IA.II.2 Reputation data

We also obtain time-varying measures of corporate reputation from RepRisk. RepRisk is a commercial company that provides data on ESG and business conduct risk. The firm compiles daily instances of company-specific news events related to 28 distinct CSR and reputation risk issues, such as air pollution, product controversies, discrimination, and labor practices, as well as broader scandals including violations of national legislation or international standards. RepRisk screens over 80,000 public sources in 20 languages, including print, online and social media, government bodies, regulators, think tanks, and newsletters to construct this data. RepRisk's analysts further classify each news item according to its novelty, severity and reach. Novelty records the first occurrence of a company-specific issue, severity captures the consequences, extent, and cause of the risk incident, and reach describes the influence (e.g. readership and circulation) of the underlying information source. RepRisk coverage starts in January 2007 and includes approximately 4,000 publicly listed companies in the Compustat North America universe (see, e.g., Dai et al., 2020 and Gantchev et al., 2020, among others). Our RepRisk sample ends in 2019. RepRisk data has been widely used in finance research, including Li and Wu (2020), Bansal et al. (2022) and Houston and Shan (2022), among others. We also use RepRisk's daily news data to construct a sample of negative corporate reputation events. Our main analysis retains only novel events (i.e., events when they were first reported) with a high level of reach, credibility and influence – for example, issues that were reported

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

Electronic copy available at: https://ssrn.com/abstract=3143740

by international news organizations. In total, we identify 2,700 negative reputation events for the firms in our sample.

IA.II.3 Corporate charitable contributions data

We obtain data on corporate charitable contributions from Foundation Directory Online (FDO) and the Urban Institute's National Center for Charitable Statistics (NCCS) database.[3] The sample period for our donations data is 2003 to 2014. We begin by searching the enterprise version of FDO to identify all foundations associated with the firms in our sample. We then collect data on each donation made by a foundation during our sample period from FDO. If a firm does not have charitable contributions data in FDO, we search for the firm in the NCCS database and collect charitable donations data from NCCS as needed. The process of supplementing FDO data with NCCS data ensures that we have both the broadest and deepest possible coverage for our charitable contributions data (though we later confirm that our results are similar using only FDO data or only NCCS data, respectively).

IA.II.4 Corporate social responsibility data

We measure corporate social responsibility using the widely-used MSCI ESG KLD Stats ("KLD") measure of CSR. These scores are developed to provide an independent assessment of firms' social responsibility, similar to the manner in which credit rating agencies assign credit ratings. To calculate the score, MSCI first determines the presence or absence of a series of social responsibility "strengths" or "concerns" within a firm. The score itself is an index that equals the number of strengths minus the number of concerns. Therefore, a firm can increase its index score by one point if it eliminates one concern or adds one strength. The score can be further broken down into several dimensions of CSR: community relations, product characteristics, environmental impact, employee relations, diversity, and governance.

We make several modifications to the KLD methodology to account for the fact that the calculation of the index has changed over the years. The individual strengths and concerns making up the index have variously been added, deleted, and redefined. Therefore, the index itself has not referred to a consistent set of actions over time. This is especially problematic around 2009, which saw a large redefinition in index components. To ensure that we study a consistent measure of CSR, we create a time-consistent index. We take the following three steps: (1) we match indicators that changed names but covered the same concepts over time, (2) we only use the indicators that are covered from 1991 through 2015, and (3) we limit our index to those indicators that were non-missing for the full sample in 2010, following the major redefinition of the index. This process yields a time-consistent CSR score that we use in our analysis. Our final CSR score is made up of eighteen strengths and six concerns.[4] For ease of interpretation, we normalize our final measure to have a mean of 0 and a standard deviation of 1 throughout the sample. We refer to this measure in our tables as "Norm CSR." In some tests, we also examine the robustness of our CSR findings using time-varying, firm-level environmental, social, and governance (ESG) scores from Asset4.

[3]Among other papers, Chatterji et al. (2015), Masulis and Reza (2015), Viceira (2020), and Cai et al. (2021) use FDO data, while Bertrand et al. (2018), Ahn et al. (2020), and Bertrand et al. (2020) use data from NCCS.

[4]We do not include any risk management or governance measures in our CSR score, as none of the KLD governance metrics (which include risk management) meet the criteria for time consistency.

Electronic copy available at: https://ssrn.com/abstract=3143740

IA.II.5 IT security and investment data

We also use textual analysis of corporate disclosure documents to measure firms' investments in IT security and infrastructure. We begin by scanning each company's 10-K filing for keywords related to 'IT security' and count the number of occurrences we observe.[5] To construct a proxy for *investment* in IT security, we also scan each 10-K filing containing at least one IT Security reference for keywords related to 'investment' that are located within 100 characters before and after each IT Security reference.[6] We remove instances where firms appear to be discussing the data breach itself by excluding the cases where the firm mentions 'data breach', 'hack', and 'hacking'. We also construct control variables capturing the length and vocabulary complexity (i.e., the ratio of unique words to total number of words) of each 10-K filing, following Loughran and McDonald (2014). Our 10-K based IT security and IT investment measures are new to the literature and complement the 10-K based textual analysis measures developed by Saunders and Tambe (2015) to capture firms' "data assets."[7]

Internet Appendix Figure IA.3 shows that firms have become much more likely to mention IT security and IT investment in their 10-K filings over time. Panel A shows that references to IT security and IT investment in firms' 10-K filings rose more than ten-fold between 2000 and 2015. Panel B shows that these increases are largely concentrated within a subset of industries. For example, disclosure is far more prevalent in the information technology (28%) and the telecommunication services sectors (21%) than in the energy and materials sectors (less than 10%). These findings suggest that corporate reputation concerns related to IT security may matter more for firms with the heaviest reliance on information technology and data.

IA.II.6 Media sentiment data

In addition, we construct measures of national and regional news media sentiment using data from Ravenpack (RP) Edge for the period from 2000 to 2016. Ravenpack Edge provides proprietary sentiment scores for each individual news item, including media from TV segments, radio features, blog posts, or newspaper articles, across thousands of individual sources. The RP sentiment measure is distributed between -1 and 1, indicating highly negative to highly positive news media sentiment. RP Edge includes separate sentiment scores for the 'document' in question (i.e., the news paper article or TV segment), the underlying story, e.g., an ongoing M&A process, and the involved entities, such as the target and acquiring firm. We focus on the document sentiment score as it most directly allows us to aggregate sentiment from individual news items at the yearly frequency to match to our sample.

[5]For our main analysis, the list of IT Security search terms we use includes 'IT infrastructure', 'cyber infrastructure', 'digital infrastructure', 'data infrastructure', 'IT security', 'cyber security', 'digital security', 'data security', 'data protection', 'IT protection', 'Information Technology security', 'enterprise security', and 'cybersecurity'. Additional robustness tests add the terms 'hack', 'hacking', and 'data breach' and find similar results.

[6]The list of IT investment terms includes 'spend money', 'invest in', 'investment in', 'investments in', 'invested in', 'improv', 'updat', 'moderniz'.

[7]While firm-level IT data exists until around 2010 from the marketing firm Harte Hanks (Forman et al., 2012; Bloom et al., 2014), the reliability of the Harte Hanks / Computer Intelligence Technology Database after that date has been questioned (Tambe et al., 2012; Saunders and Tambe, 2015). There are also papers that use industry-level measures from sources such as *InformationWeek* (see e.g., Aldasoro et al., 2020; Kennedy and Stratopoulos, 2017).

Electronic copy available at: https://ssrn.com/abstract=3143740

We process Ravenpack Edge news media data in the following way. First, we download all 'full-article' news media items related to 'business' topics, that are relevant (at least 90/100 RP relevance score) to our sample firms and published in the US in English-speaking outlets using the RP Edge API. We retain only news reports from outlets with a source rating of 3 or better (1 through 8 from best to worst source reputation). This procedure yields a total of 10,481 unique news sources. Next, we classify each news source as either national or regional by querying the OpenAI GPT API, using GPT-3.5 with the 'davinci' engine.[8] We repeat this procedure three times to reduce the effect of noise, retain the majority answer from GPT, and verify the accuracy of the assignment for the 200 news sources with the most news reports in the sample. In the same way we classify news sources into different categories, i.e., TV, newspaper, radio, etc.[9] Our classification yields 446 blogs, 1141 magazines, 1767 newspapers, 297 radio stations, 47 scientific publications, 536 TV stations, 5146 websites, 79 wire services, and 1021 other outlets; approximately 61% of outlets are classified as 'national'.

We then map the entities covered by RP to the firms in our sample using the RP-to-WRDS crosswalk provided by Ravenpack, and aggregate document-level news sentiment across all published items at the firm-by-year level. We further do this separately for national and regional news media, retaining only publications from TV stations, newspapers, and radio stations, which can be more unambiguously assigned as either national or regional compared to blog posts or websites.

IA.II.7 Twitter data

In addition to traditional news media, we obtain social media data based on Twitter content from the data provider 'Social Market Analytics (SMA)' (now known as 'Context Analytics').[10] SMA is a social media monitoring firm and data provider that markets high-frequency sentiment measures and other textual metrics based on Twitter posts to investors, firms, and academics, using a proprietary sentiment algorithm. From SMA, we obtain data on Twitter sentiment, Tweet volume (i.e., the number of Tweets posted), and Twitter 'buzz', i.e., a measure of unusual volume activity compared to a universe of stocks in the cross-section, from SMA at the daily frequency. SMA's algorithm considers only relevant Tweets from 'credible' accounts when calculating sentiment and buzz, to address the presence of bots.

SMA data is available to us at the firm-by-day frequency and includes ticker symbols, which allows us to match daily Twitter sentiment, Tweet volume, and Twitter 'buzz' to our sample firms. SMA's coverage begins in 2011 (Twitter was launched only shortly before) and grows exponentially over time along with the growth of overall social media usage, from 39,614 unique firm-days with Twitter sentiment from 264,028 total tweets in 2011, to 1,028,724 firm-days from 6,319,524 total tweets in 2020, after merging with CRSP/Compustat firm identifiers. Given the late start date of SMA coverage, merging SMA social media data with the firms in our data breach sample yields too few firms for analysis. (Only 47 unique events

[8] Specifically, we submit the name of each news source to GPT with the following prompt "For the following news source, tell me if it is a regional or national news source. Start your answer with 'regional' or 'national' ", and record the response.

[9] Specifically, we submit the query "For the following news source, tell me the type of news source and your confidence on a scale from 1 to 10. Start with either 'TV', 'newspaper', 'website', 'blog', 'wire service', 'radio', 'scientific publication', 'magazine' or 'other', followed by a ';' and the confidence number:".

[10] Twitter is now called X. We use the name Twitter in our paper since the site was called Twitter during our sample period.

Electronic copy available at: https://ssrn.com/abstract=3143740

survive the merge, the wide majority of which occurred during 2012 and 2013 when Twitter usage was much lower than in the later part of the sample). We therefore focus on RepRisk events in our analysis of social media reactions to reputation shocks.

Twitter sentiment from SMA is constructed as an aggregation of sentiment of unique tweets received from credible accounts in the past 24h, and has an average value of 0 and a standard deviation of 1 by construction. On average, firms receive about 6 Tweets that enter the sample on a given day with a standard deviation of 23.35. This distribution is highly skewed; the 95th percentile is close to 20 Tweets with a maximum of 7,160 Tweets in one day (about ticker 'GME' on Jan 27, 2021). 'Buzz', which is constructed as abnormal Tweet volume, i.e., it is normalized in the daily cross section, has an average value and standard deviation of 1.34 and 0.79, with a minimum of 0 and a maximum of 45.49.

IA.II.8 Conference call data

We further construct measures to capture the content of earnings conference calls with equity analysts, using conference call transcripts from the data provider Streetevents. For this purpose, we distinguish between the management presentation section and the Q&A section of each earnings call, and construct indicator and count variables that capture if the section of the earnings call discusses issues related to reputation, CSR, data security, and data breaches.[11]

Conference call transcripts are available to us at the quarterly frequency from 2004 through 2014. For analyses conducted at the firm-year level, we collapse the quarterly data at the annual frequency by summing over the occurrence of individual words in our content dictionaries. Our conference call data covers a total of 26,447 firm-years.

IA.II.9 Brand value data

Further, we obtain data on brand values from Brand Asset Valuator (BAV). BAV is a proprietary brand metrics model provided by 'Brand Asset Consulting', a subsidiary of Young & Rubicam Brands. The BAV model relies on a consumer survey-based approach from a wide base of respondents, representative of the US population, and does not require accounting or market valuation variables. BAV brand value data is widely cited in marketing research (e.g., Mizik and Jacobson, 2008; Naidoo and Abratt, 2018) and is used by Young & Rubicam's clients to analyze different aspects of brand image.[12]

BAV data is available to us starting in 2001 and ending in 2011. Since our RepRisk data starts in late 2009, the sample periods in two data sources do not overlap sufficiently to conduct meaningful analysis on the relationship of RepRisk events and BAV brand values. We therefore focus on data breaches in our analysis of brand value effects.

BAV data is organized at the individual brand level, e.g., 'Fanta', 'Diet Coke', and 'Sprite' for the Coca Cola company, and is available at the quarterly frequency. We manually match

[11] Specifically, the 'reputation' dictionary includes the words 'reputation' and 'trust'. The 'data security' dictionary includes the words 'data* protec*', 'data* securit*', 'it* protec*', 'it* securit*', and 'cyber* sec*'. The 'data breaches' dictionary includes the words 'hack*', 'data* breach', and 'cyber* attack*'. The 'CSR and Philanthropy' dictionary include the words 'human right*', 'sustainab*', 'soc* resp', 'corp* soc*', 'ethical', 'philanthrop*', 'char* contr*', 'char* grant*', and 'charity'.

[12] Further, according to the BAV documentation, the list of brands included in the dataset is not biased towards clients of Brand Asset Consulting, as the company tries to maintain a fair representation of all major industry competitors.

Electronic copy available at: https://ssrn.com/abstract=3143740

brands to publicly listed firms and collapse the data at the quarterly frequency by averaging across brands for a given firm. For our analyses at the annual frequency, we retain the last value for a given brand in a given year and average across brands for each firm-year. In total, our BAV data covers 5,921 firm-years (21,676 firm-quarters) from 2001 to 2011 across 826 unique firms. Our main metric of brand value is 'brand strength', which captures how much regard and loyalty consumers have towards the brand and is distributed between 0 and 100.

IA.II.10 Customer churn data

We additionally obtain data on customer churn at the firm level from (Baker et al., 2023).[13] Baker et al. (2023) construct measures of customer churn using credit card and checking account transaction-level data from a large online financial account aggregation service, which offers individual users the ability to track their finances across multiple financial institutions and accounts on a single website or mobile app. Based on the transaction dates, amounts, and identity of the seller of each individual transaction, Baker et al. (2023) construct the 'spending-share-adjusted change' as their measure of customer churn, i.e., the difference between two given periods t and $t-k$ in the share of firm f's revenue coming from individual i averaged across all individuals $i = 1, ..., I$.

Baker et al. (2023) differentiate between overall churn, churn from new customers (i.e., calculated from shoppers who buy at firm f only in t), and churn of old customers (i.e., calculated from shoppers who purchased from firm f only in $t-k$). Because our analysis focuses on the effect of reputation shocks on the relationships with important stakeholders, we use the loss of existing customers as our main measure of customer churn. By construction, churn is between 0 and 1. The data covers 6,277 firm-year-quarters across 420 unique firms between 2011 and 2015 (5,112 firm-quarters after merging with our RepRisk events sample).[14]

IA.II.11 Political contributions data

We obtain political connections data from the Federal Election Commission. We manually match firms' political action committees (PACs) to their Compustat identifiers and use detailed PAC contribution data to calculate the total dollar amount of contributions to each candidate for the U.S. House of Representatives and U.S. Senate at the firm-candidate-election cycle frequency.

IA.II.12 Employee wage data

To examine the effect of firm-level reputation events on relationships with employees, we further obtain data on person-level salaries from Glassdoor.com. Glassdoor.com allows individuals to self-report their occupation, job title, work status, and salary along with a number of personal characteristics. We hand-match employer names from Glassdoor.com to firm identifiers from Compustat, retain all employees with full-time status, and convert hourly and monthly salaries to annual salaries to allow for a comparison across individuals. After matching employer names to public firm identifiers, our data covers 976,293 individual salary observations across 1,900 unique firms between 2006 to 2016. The number of individual data points grows exponentially during this sample period with the growth of the Glassdoor.com user base, from 2,959 self-reported salaries in 2006 to 210,139 in 2015.

[13] We thank Scott Baker for graciously sharing customer churn data at the quarterly frequency with us.
[14] There are too few matches with firm in the hacking sample to use the churn data for analysis.

Electronic copy available at: https://ssrn.com/abstract=3143740

In addition to salary information, we collect data on work experience, occupation group, employee location at the metropolitan area level, gender, and education from Glassdoor.com. The average annual salary in the sample of public firms $73,000, the average employee is 34 years old. A large fraction of observations do not include information on education or gender (62% and 45%); among the remaining observations, 60% hold a Bachelor's degree, 23% hold a Masters or MBA degree, and 1.2% have a PhD. About 61% of users are male. The most commonly represented occupation groups (in order) are branch manager, software engineer, sales representative, customer service, analyst, store manager, engineer, finance specialist, project manager, and marketing manager. The top 5 most widely represented metropolitan regions in the data are New York City, Chicago, San Jose, Seattle, and San Francisco.

In addition, we also obtain ratings with respect to corporate culture and firm values posted by employees about their employers from Glassdoor.com

IA.II.13 Other firm outcome and control variables

We obtain data on firm returns from CRSP and firm fundamentals from Compustat. Following Kamiya et al. (2021), we also construct measures of risk committee membership, CEO dual role, and institutional block ownership. Using data from BoardEx, we construct an indicator variable that takes the value of one if the name of a firm's board committee includes the word "risk", and zero otherwise, for 'risk committee membership', as in Kamiya et al. (2021). Similarly, the indicator variable 'Dual Role CEO' takes the value of one if the CEO is also the chair of the board, and zero otherwise. We further collect indicators of S&P 500 membership for our sample firms from CRSP. Some variables in our analysis, such as firms' market-to-book (M/B) ratios, are computed by merging data from both CRSP and Compustat.

In addition, we use data from the Compustat Segment Files on supplier firms and their corporate and government customers to construct measures of final consumer proximity. SEC regulation S-K requires public firms in the US to report the identity of customers who represent more than 10% of their total revenue. Many firms additionally disclose the dollar value of sales to a specific customer. We use this data to construct variables capturing whether or not a given firm has corporate and government customers, respectively, and the percentage these customer groups represent relative to total revenues. By design, a firm with a high share of corporate or government customers is further removed from the final consumer than a firm with few corporate or government customers. In total, the Segment Files from Compustat cover 12,416 unique firms between 1978 and 2020. Among our sample firms and period, approximately 32.5% (4.5%) report at least one corporate (government) customer in a given year on average.

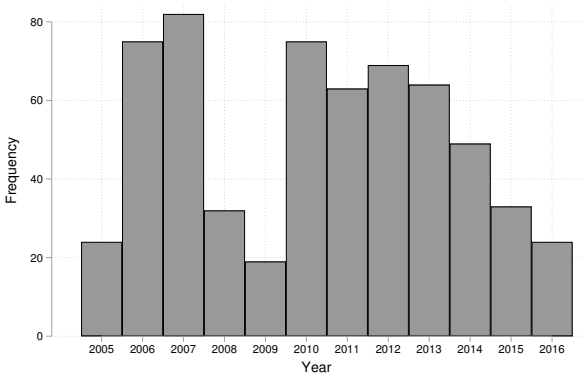
Finally, we use Thomson Reuters' Institutional Holdings (13f) database to obtain the proportion of shares held by institutional investors who own at least 5% of a firm's outstanding equity.

Electronic copy available at: https://ssrn.com/abstract=3143740

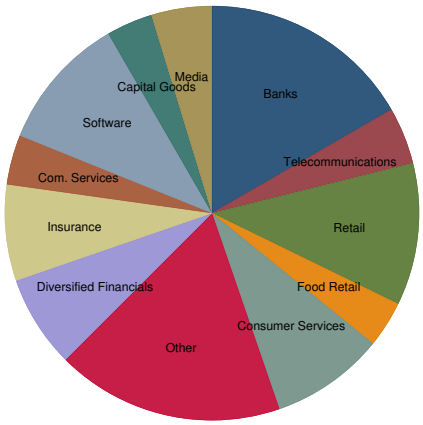
IA.III Additional Tables and Figures

Figure IA.1: Frequency an Types of Breaches

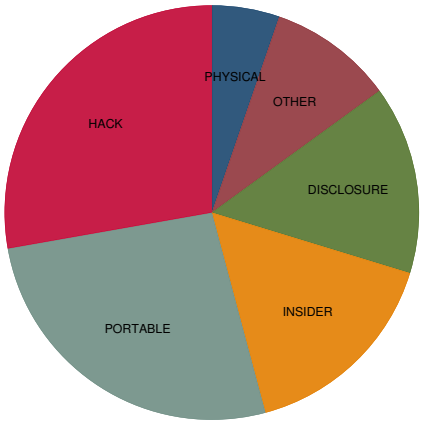
(a) Frequency of Breaches over Time



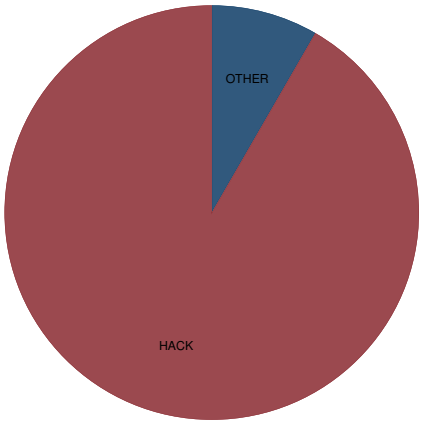
(b) Industries Affected by Breaches



(c) Breaches by Type



(d) Records by Breach Type

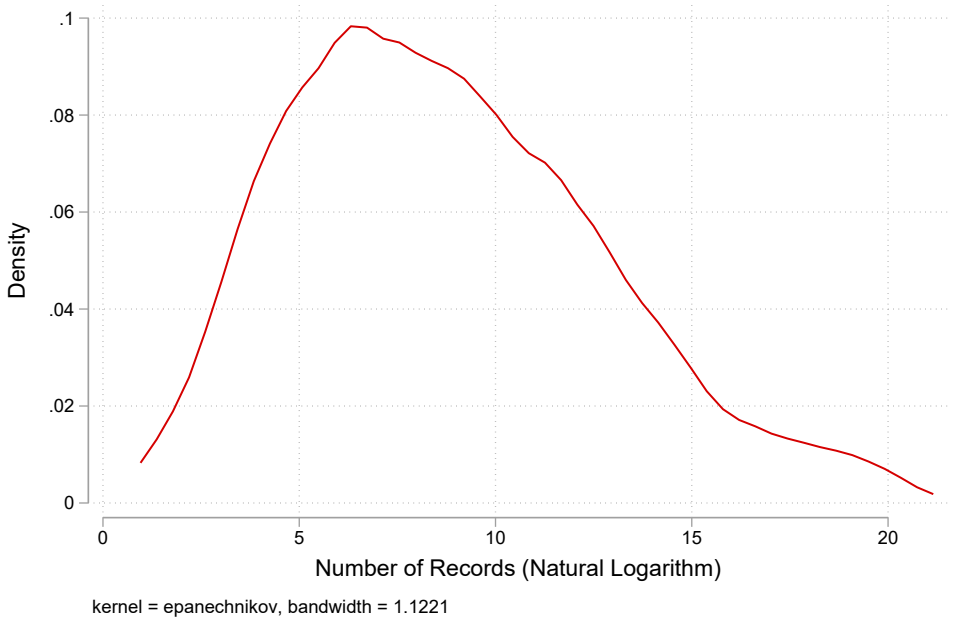


Notes: These figures show the distribution of data breaches across time and industry, and the distribution of the types of security breaches. Panel IA.1a presents the number of breaches per year and Panel IA.1b presents the proportion of breaches across four-digit Global Industry Classification Standard (GICS) industries. Panel IA.1c presents the frequency of different types of breaches. Panel IA.1d presents the proportion of records compromised by different types of breaches.

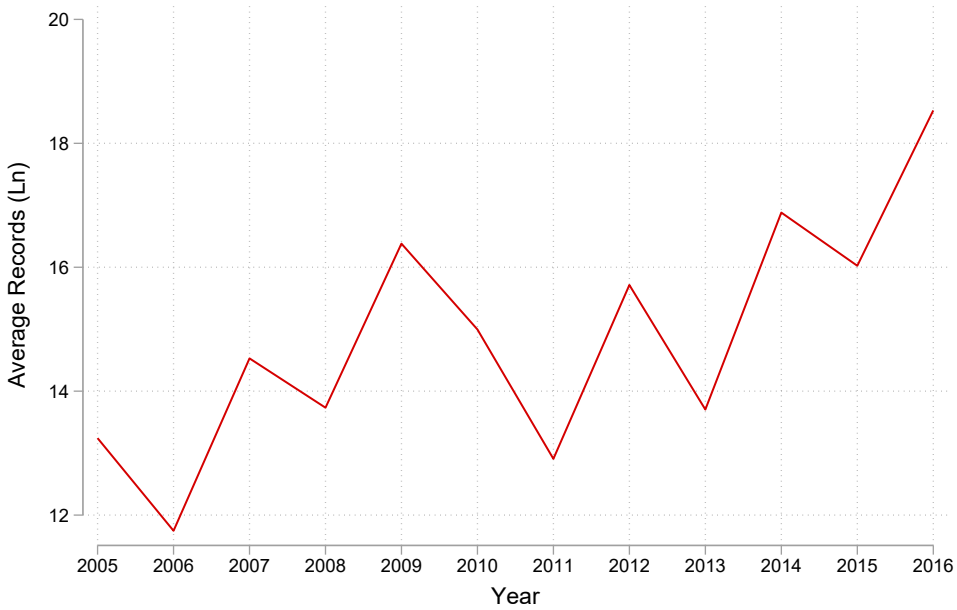
Electronic copy available at: https://ssrn.com/abstract=3143740

Figure IA.2: Number of Records Affected by Breaches

(a) Number of Records



(b) Number of Records over Time

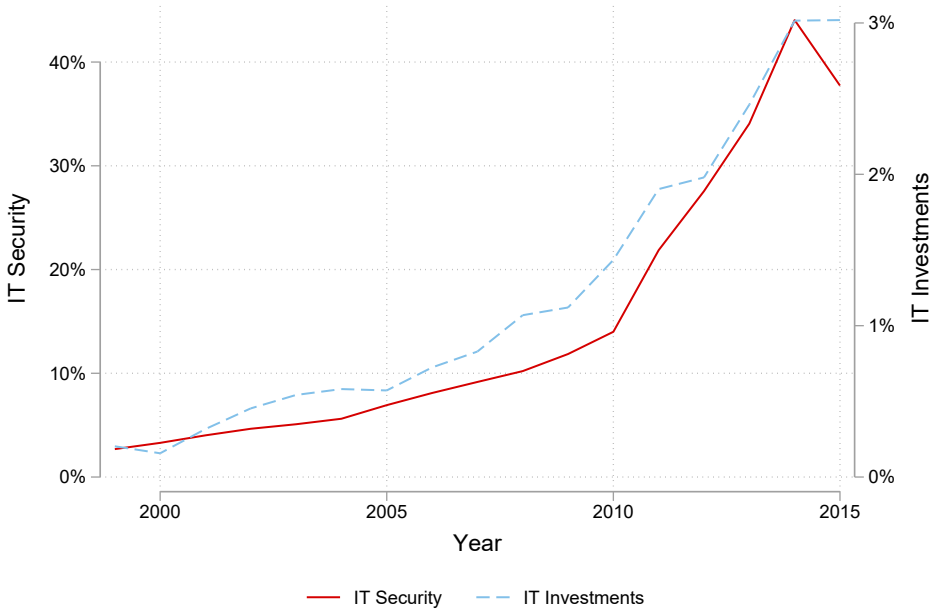


Notes: These figures show the distribution of the number of records affected by security breaches. Panel IA.2a presents a kernel density plot of the natural logarithm of the number of records breached. Panel IA.2b presents the natural logarithm of the average number of records breached throughout the sample period.

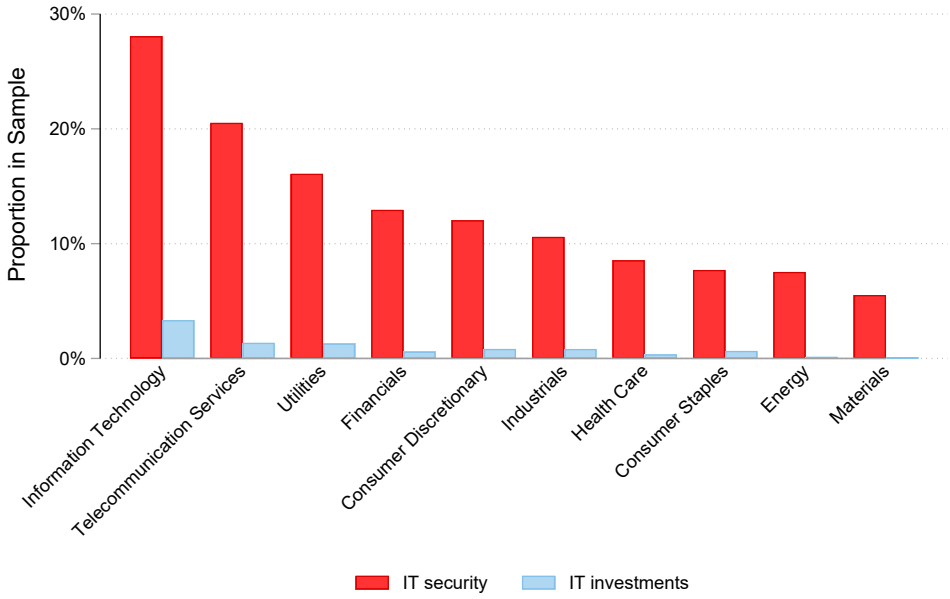
Electronic copy available at: https://ssrn.com/abstract=3143740

Figure IA.3: Frequency of IT Security Investments

(a) Frequency over time



(b) Frequency by industry



Notes: These figures document the proportion of firms that mention key phrases related to IT Security and IT Investments in their 10K filings over time (Panel IA.3a) and across sectors (Panel IA.3b), respectively. Details of how these measures are constructed are provided in Section 2.6.

Electronic copy available at: https://ssrn.com/abstract=3143740

Table IA.1: Summary Statistics

Notes: Panel IA.1a of this table reports summary statistics for all data breaches that have been matched to public firms and have a non-missing value for total number of records compromised. The remaining variables are non mutually exclusive indicators for whether the affected records included employee data, customer data, or internal documents. The last two variables are indicators for whether the compromised entity was a subsidiary and whether the data breach affected multiple firms. Panel IA.1b reports similar summary statistics for the RepRisk events in our sample. We only include novel (i.e., not duplicated) events, categorized by RepRisk as 'high reach' and 'high severity'. Similar to data breaches, RepRisk categories (environmental, social, governance, product, and violations) are not mutually exclusive. Panel IA.1c reports summary statistics for all firms that appear either in our data breach or in our RepRisk events sample. "Norm CSR" is the standardized measure of CSR from KLD. "Charitable Donations (M. USD)" and "Foundation (0/1)" are the annual amount of charitable donations and an indicator variable taking the value of one if a firm has made a donation through its foundation, respectively. "(Regional/National) News Sentiment" is the average sentiment of the firm's news coverage (between -1 and 1) from Ravenpack Edge obtained from regional and national news outlets, respectively. "Brand Strength" is obtained from the BAV model, "Political Contributions" is the annual amount of contributions to house- and state races by firms' PACs. "1(CSR)", "1(Data Breach)", and "1(Reputation)" (CC PPT and CC QnA) are dummy variables indicating the occurrence of terms related to CSR, data breach, and reputation, respectively, in the firm's presentation (PPT) and Q&A (QnA) section of their earnings conference calls. "Book Assets", "Leverage", "ROA", "ROE", "P/E", "M/B", and "Advertising Expenses/Assets" are Winsorized at the 5% level within the full Compustat universe. "Inst. Block Own (%)", "Dual Role CEO (0/1)", and "Risk Committee (0/1)" are constructed as in Kamiya et al. (2021). "Reputation Rating" is a measure of firm reputation from RepRisk. "IT Security (0/1)" and "IT Investment (0/1)" indicate if firms mention IT security and investment in their 10K filings. "E-Index" and "G-Index" are corporate governance measures from Gompers et al. (2003) and Bebchuk et al. (2009), respectively. "ESG Score (A4)" is the standarized ESG score from Asset4. Details on data sources and variable construction for both Panels are summarized in Section 2.

(a) Data Breaches

	Mean	Median	StDev
Total Records	6,173,208	3,482	40,340,917
Employee Records	.331	0	.471
Customer Records	.655	1	.476
Internal Documents	.0139	0	.117
Subsidiary	.0906	0	.288
Multiple Firms	.0523	0	.223
Observations	287		

(b) RepRisk Events

	Mean	Median	StDev
Environmental	.104	0	.305
Social	.307	0	.461
Governance	.539	1	.499
Product	.217	0	.412
Violation	.541	1	.498
Observations	3121		

Electronic copy available at: https://ssrn.com/abstract=3143740

. . . continued

(c) Full Sample

	N	Mean	StDev	p25	Median	p75
Norm CSR	30380	.00546	1.01	-.289	-.289	.615
Charitable Donations (M. USD)	39928	.378	1.88	0	0	0
Foundation (0/1)	39928	.151	.358	0	0	0
News Sentiment (All)	51979	.103	.533	-.294	.0154	.5
Regional News Sentiment	23795	.187	.158	.091	.198	.29
National News Sentiment	26297	.162	.158	.07	.18	.266
Brand Strength	5407	46.3	26.4	24.5	45.8	66.9
Political Contributions (USD)	10947	.0871	.165	.00825	.0295	.0908
1(CSR CC PPT)	44915	.192	.394	0	0	0
1(CSR CC QnA)	44915	.266	.442	0	0	1
1(Data Breach CC PPT)	44915	.00401	.0632	0	0	0
1(Data Breach CC QnA)	44915	.00443	.0664	0	0	0
1(Reputation CC PPT)	44915	.159	.366	0	0	0
1(Reputation CC QnA)	44915	.13	.337	0	0	0
Book Assets (M. USD)	126636	2535	5370	36.2	291	1683
Leverage	111157	.255	.262	.0119	.171	.436
ROA	125549	-.169	.48	-.114	.00946	.0532
ROE	126034	.0265	.566	-.0687	.0803	.177
PE	111255	8.22	23.6	-3.46	8.83	19.1
M/B	98048	2.93	2.92	1.09	1.84	3.42
Q	110828	2.57	3.01	1.03	1.37	2.45
Adv. Expenses/Assets	45523	.0227	.0354	.000999	.00636	.0262
Sales/Turnover	126156	1325	2796	18.6	133	899
Nonrecurring	131767	.0937	.291	0	0	0
Inst. Block Own. (%)	49885	.122	.0894	.0754	.0974	.133
Risk Committee (0/1)	53880	.0357	.186	0	0	0
Dual Role CEO (0/1)	54885	.505	.5	0	1	1
Reputation Rating	16657	8.47	1.46	8	9	9.17
IT Security (0/1)	70468	.139	.346	0	0	0
IT Investment (0/1)	70468	.0114	.106	0	0	0
E-Index	29706	2.93	1.07	3	3	3
G-Index	29706	6.28	1.53	5	6	7
WW Index	103144	-.101	.446	-.0208	-.000573	.00217
ESG Score (A4)	10444	.162	.921	-.503	-.0977	.74
E Score (A4)	10434	-.189	.949	-.884	-.633	.402
S Score (A4)	10444	-.0201	.922	-.753	-.155	.599
G Score (A4)	10444	.63	.636	.277	.667	1.03
Observations	131767					

Electronic copy available at: https://ssrn.com/abstract=3143740

Table IA.2: Summary Statistics by Sample

Notes: This table presents summary statistics at the firm-year level for the firms in our sample, presented separately for the data breach sample in Panel IA.2a and the RepRisk event sample in Panel IA.2b. We retain only firms with at least one data breach or RepRisk event, respectively, in their 6-digit GIC industry. The sample period is from 1999 to 2015. Compustat variables have been Winsorized at the 5th percentiles. Variables are defined similarly as in Table IA.1. Details on data sources and variable construction for both Panels are summarized in Section 2.

(a) Data Breaches – Full Sample of Treated and Control Firms

	N	Mean	StDev	p25	Median	p75
Norm CSR	23276	.0184	.996	-.289	-.289	.615
Charitable Donations (M. USD)	30692	.401	1.97	0	0	0
Foundation (0/1)	30692	.146	.353	0	0	0
News Sentiment (All)	39740	.115	.534	-.28	.0455	.5
Regional News Sentiment	18261	.186	.157	.0909	.198	.289
National News Sentiment	20279	.162	.157	.0711	.18	.265
Brand Strength	4523	46.7	26.7	24.8	46	67.7
Political Contributions (USD)	8175	.0976	.183	.00875	.0326	.0983
1(CSR CC PPT)	33849	.186	.389	0	0	0
1(CSR CC QnA)	33849	.254	.435	0	0	1
1(Hack CC PPT)	33849	.00464	.0679	0	0	0
1(Hack CC QnA)	33849	.00455	.0673	0	0	0
1(Reputation CC PPT)	33849	.166	.372	0	0	0
1(Reputation CC QnA)	33849	.134	.34	0	0	0
Assets	95646	2579	5501	38.7	293	1617
Leverage	84675	.259	.265	.0106	.174	.446
ROE	95185	.0234	.572	-.0711	.0799	.177
PE	84779	8.46	23.6	-3.35	9.15	19.3
M/B	74901	2.97	2.95	1.09	1.86	3.5
Adv. Expenses/Assets	37726	.0225	.035	.000901	.00603	.0269
Inst. Block Own. (%)	37575	.121	.0887	.0754	.0972	.133
Risk Committee (0/1)	41288	.0408	.198	0	0	0
Dual Role CEO (0/1)	42021	.505	.5	0	1	1
Reputation Rating	12379	8.5	1.46	8	9	9.25
IT Security (0/1)	53491	.152	.359	0	0	0
IT Investment (0/1)	53491	.0122	.11	0	0	0
E-Index	22262	2.89	1.06	3	3	3
G-Index	22262	6.24	1.52	5	6	7
ESG Score (A4)	7557	.104	.907	-.534	-.148	.631
Observations	97448					

Electronic copy available at: https://ssrn.com/abstract=3143740

Continued...

(b) RepRisk Events – Full Sample of Treated and Control

	N	Mean	StDev	p25	Median	p75
Reputation Rating	16600	8.47	1.47	8	9	9.17
Norm CSR (KLD)	30230	.00573	1.01	-.289	-.289	.615
Charitable Donations (M. USD)	39749	.379	1.88	0	0	0
Foundation (0/1)	39749	.151	.358	0	0	0
News Sentiment (All)	51635	.104	.533	-.292	.0174	.5
Regional News Sentiment	23634	.187	.158	.0909	.198	.29
National News Sentiment	26133	.162	.158	.07	.18	.266
Brand Strength	5402	46.3	26.4	24.5	45.8	66.8
Political Contributions (USD)	10913	.0872	.165	.00825	.0295	.0909
1(CSR CC PPT)	44614	.193	.394	0	0	0
1(CSR CC QnA)	44614	.266	.442	0	0	1
1(Hack CC PPT)	44614	.00401	.0632	0	0	0
1(Hack CC QnA)	44614	.00446	.0666	0	0	0
1(Reputation CC PPT)	44614	.159	.366	0	0	0
1(Reputation CC QnA)	44614	.13	.336	0	0	0
Assets	125849	2548	5382	36.8	295	1702
Leverage	110445	.256	.262	.0123	.173	.437
ROA	124765	-.168	.479	-.112	.00951	.0531
ROE	125248	.0267	.565	-.0679	.0804	.177
PE	110544	8.23	23.5	-3.43	8.86	19.1
M/B	97467	2.92	2.91	1.09	1.83	3.41
Q	110117	2.56	3	1.02	1.37	2.43
Adv. Expenses/Assets	45190	.0228	.0355	.000995	.00636	.0263
Sales/Turnover	125370	1331	2802	18.9	135	907
Nonrecurring	130924	.0939	.292	0	0	0
Inst. Block Own. (%)	49605	.122	.0895	.0754	.0974	.133
Risk Committee (0/1)	53558	.0359	.186	0	0	0
Dual Role CEO (0/1)	54551	.505	.5	0	1	1
Reputation Rating	16600	8.47	1.47	8	9	9.17
IT Security (0/1)	70066	.138	.345	0	0	0
IT Investment (0/1)	70066	.0114	.106	0	0	0
E-Index	29545	2.93	1.07	3	3	3
G-Index	29545	6.28	1.53	5	6	7
WW Index	102468	-.0013	.00488	-.000511	-.0000441	.0000534
ESG Score (A4)	10418	.163	.921	-.502	-.0971	.742
E Score (A4)	10408	-.188	.949	-.884	-.633	.402
S Score (A4)	10418	-.0195	.922	-.753	-.154	.6
G Score (A4)	10418	.631	.636	.277	.667	1.03
Observations	130924					

Electronic copy available at: https://ssrn.com/abstract=3143740

Table IA.3: Determinants of Data Breaches and Reputation Shocks — Relation to the Related Literature

Notes: This table summarizes Logit and linear probability models estimating the predictability of data breaches and RepRisk events using various firm- and industry characteristics, following the specifications in Florackis et al. (2023) and Kamiya et al. (2021). The dependent variable in Panels IA.3a through IA.3d is an indicator that takes the value of one if the firm experienced a large data breach that exposed Social Security Numbers (SSN), similar to the variable definition in Florackis et al. (2023) and Kamiya et al. (2021). The dependent variable in Panel IA.3e is a dummy indicating the occurrence of a RepRisk event. Panels IA.3a and IA.3b report estimates from Logit and linear probability model regressions, respectively. "Cyber Risk (FLMW)" is the text-based measure of cybersecurity risk of Florackis et al. (2023) and "1(S&P500)" is a dummy indicating that the firm was a member of the S&P500 index in the given year. Panels IA.3c through IA.3e implement the specification of Table 3 in Kamiya et al. (2021). All independent variables are constructed following Kamiya et al. (2021) and described in detail in the Appendix. All explanatory variables are measured one year before the reputation shock except for Tobin's Q, which is measured with a lag of two years. Similar to Kamiya et al. (2021), the sample in Panel IA.3c starts in 2005, while Panels IA.3d and IA.3e include our full sample starting in 1999. Standard errors are clustered at the firm level and reported in parentheses. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

(a) Cybersecurity Risk (Florackis et al., 2023) and Large Data Breaches – Logit

	Logit: Breach with SSN (0/1)		
	(1)	(2)	(3)
Cyber Risk (FLMW)	2.5880***	1.1224**	1.1065**
	(0.4315)	(0.5284)	(0.5200)
Firm Size		0.7410***	0.7551***
		(0.0751)	(0.0765)
Tobin's Q			0.1590**
			(0.0701)
ROA			0.3425
			(1.5857)
Leverage			0.1898
			(0.6058)
1(S&P500)		0.3496	0.2914
		(0.2712)	(0.2749)
Year FE	Yes	Yes	Yes
SIC2 FE	Yes	Yes	Yes
Observations	27948	21344	21205
Pseudo R^2	.114	.301	.304

Electronic copy available at: https://ssrn.com/abstract=3143740

Continued...

(b) Cybersecurity Risk (Florackis et al., 2023) and Large Data Breaches – OLS

	OLS: Breach with SSN (0/1)				
	(1)	(2)	(3)	(4)	(5)
Cyber Risk (FLMW)	0.0181***	0.0183***	0.0047	0.0038	0.0041
	(0.0032)	(0.0032)	(0.0030)	(0.0030)	(0.0044)
Firm Size			0.0050***	0.0035***	0.0006
			(0.0006)	(0.0005)	(0.0010)
Tobin's Q				0.0006**	-0.0002
				(0.0003)	(0.0004)
ROA				-0.0079***	-0.0003
				(0.0016)	(0.0019)
Leverage				-0.0017	0.0020
				(0.0025)	(0.0035)
1(S&P500)				0.0214***	-0.0135
				(0.0033)	(0.0084)
Year FE	Yes	Yes	Yes	Yes	Yes
SIC2 FE	Yes	No	Yes	Yes	No
SIC2-by-Year FE	No	Yes	No	No	No
Firm FE	No	No	No	No	Yes
Observations	32289	32242	32287	32077	31614
R^2	.0129	.0347	.0258	.032	.226
Within-R^2	.00132	.00135	.0144	.0207	.000333

Electronic copy available at: https://ssrn.com/abstract=3143740

Continued...

(c) Determinants of Large Data Breaches (Kamiya et al., 2021) – After 2005

	OLS: Breach with SSN (0/1)					
	(1)	(2)	(3)	(4)	(5)	(6)
Firm Size	0.0050***	0.0051***	0.0024*	0.0018	0.0031***	0.0032***
	(0.0007)	(0.0007)	(0.0014)	(0.0014)	(0.0005)	(0.0004)
Tobin's Q (t-1)	0.0005	0.0004	-0.0004	-0.0004	0.0005**	0.0005**
	(0.0003)	(0.0003)	(0.0004)	(0.0005)	(0.0002)	(0.0002)
ROA	-0.0082***	-0.0085***	-0.0032	-0.0034	-0.0055***	-0.0054***
	(0.0019)	(0.0020)	(0.0023)	(0.0023)	(0.0014)	(0.0013)
Leverage	-0.0039	-0.0053	-0.0046	-0.0094*	-0.0005	-0.0028
	(0.0035)	(0.0036)	(0.0051)	(0.0055)	(0.0025)	(0.0025)
Log(Firm Age)	0.0018	0.0018	0.0218***	0.0219***	0.0009	0.0017**
	(0.0011)	(0.0012)	(0.0071)	(0.0082)	(0.0008)	(0.0009)
Sales Growth	0.0001	-0.0001	0.0002	-0.0001	-0.0001	0.0001
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)
BHAR 12-Mths	0.0004	0.0005	-0.0001	0.0002	0.0003	0.0003
	(0.0008)	(0.0008)	(0.0007)	(0.0008)	(0.0006)	(0.0006)
1(Financial Constraint)	-0.0007	-0.0010	-0.0023*	-0.0028**	-0.0007	-0.0002
	(0.0010)	(0.0011)	(0.0013)	(0.0014)	(0.0008)	(0.0008)
Return Volatilty	0.0872***	0.0909***	0.0205	0.0236	0.0450**	0.0576***
	(0.0299)	(0.0308)	(0.0238)	(0.0245)	(0.0186)	(0.0201)
Inst. Block Own.	-0.0122	-0.0096	-0.0140	-0.0109	-0.0118*	-0.0097
	(0.0093)	(0.0092)	(0.0126)	(0.0128)	(0.0064)	(0.0064)
R&D/Assets	0.0293***	0.0289***	0.0324	0.0158	0.0020	0.0143**
	(0.0094)	(0.0094)	(0.0232)	(0.0222)	(0.0066)	(0.0069)
CAPX/Assets	-0.0066	-0.0007	-0.0139	-0.0044	-0.0007	-0.0041
	(0.0110)	(0.0108)	(0.0101)	(0.0105)	(0.0063)	(0.0065)
Asset intangibility	0.0062	0.0067	0.0043	0.0050	0.0081***	0.0016
	(0.0040)	(0.0041)	(0.0074)	(0.0083)	(0.0026)	(0.0027)
1(S&P500)	0.0224***	0.0227***	-0.0003	0.0016	0.0288***	0.0288***
	(0.0034)	(0.0034)	(0.0068)	(0.0068)	(0.0035)	(0.0035)
1(Risk Committee)	-0.0161***	-0.0158***	-0.0067	-0.0052		
	(0.0040)	(0.0043)	(0.0099)	(0.0072)		
N Board Committees	0.0011	0.0010	-0.0013	-0.0012		
	(0.0008)	(0.0008)	(0.0013)	(0.0013)		
Industry HHI					0.0315**	
					(0.0126)	
1(Unique Industry)					0.0016	
					(0.0017)	
Industry Tobin's Q					0.0000	
					(0.0012)	
1(Wholesale and retail trade)						0.0151***
						(0.0032)
1(Finance)						0.0073***
						(0.0018)
1(Service Industries)						0.0069***
						(0.0015)
1(Transportation and Communication)						-0.0003
						(0.0012)
Year FE	Yes	No	Yes	No	Yes	Yes
SIC2 FE	Yes	No	No	No	No	No
SIC2-by-Year FE	No	Yes	No	Yes	No	No
Firm FE	No	No	Yes	Yes	No	No
Observations	34060	34005	33519	33305	43658	43658
R^2	.0381	.0617	.215	.241	.0271	.0293
Within-R^2	.0262	.0269	.000821	.00079	.0259	.028

Electronic copy available at: https://ssrn.com/abstract=3143740

Continued...

(d) Determinants of Large Data Breaches (Kamiya et al., 2021) – Full Sample

	OLS: Breach with SSN (0/1)					
	(1)	(2)	(3)	(4)	(5)	(6)
Firm Size	0.0046***	0.0047***	0.0030***	0.0027**	0.0021***	0.0021***
	(0.0006)	(0.0006)	(0.0009)	(0.0011)	(0.0003)	(0.0003)
Tobin's Q (t-1)	0.0003	0.0002	-0.0005	-0.0007**	0.0002	0.0002
	(0.0002)	(0.0003)	(0.0003)	(0.0004)	(0.0001)	(0.0001)
ROA	-0.0066***	-0.0072***	-0.0034**	-0.0035*	-0.0040***	-0.0037***
	(0.0015)	(0.0016)	(0.0017)	(0.0018)	(0.0008)	(0.0007)
Leverage	-0.0040	-0.0054*	-0.0012	-0.0048	-0.0006	-0.0017
	(0.0028)	(0.0030)	(0.0034)	(0.0038)	(0.0016)	(0.0016)
Log(Firm Age)	0.0012	0.0012	-0.0022	-0.0037	0.0003	0.0009
	(0.0009)	(0.0009)	(0.0037)	(0.0046)	(0.0006)	(0.0006)
Sales Growth	0.0000	-0.0001	0.0002	0.0001	-0.0001	-0.0000
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0001)	(0.0002)
BHAR 12-Mths	0.0009**	0.0009*	0.0004	0.0004	0.0003	0.0004
	(0.0004)	(0.0004)	(0.0005)	(0.0005)	(0.0002)	(0.0002)
1(Financial Constraint)	-0.0010	-0.0013	-0.0019**	-0.0023**	-0.0015***	-0.0012**
	(0.0008)	(0.0009)	(0.0009)	(0.0010)	(0.0005)	(0.0005)
Return Volatilty	0.0751***	0.0771***	0.0167	0.0103	0.0446***	0.0484***
	(0.0232)	(0.0240)	(0.0187)	(0.0195)	(0.0111)	(0.0121)
Inst. Block Own.	-0.0104	-0.0088	-0.0116	-0.0104	-0.0084**	-0.0067
	(0.0073)	(0.0073)	(0.0086)	(0.0089)	(0.0041)	(0.0041)
R&D/Assets	0.0252***	0.0253***	0.0276	0.0214	0.0036	0.0100**
	(0.0075)	(0.0075)	(0.0169)	(0.0164)	(0.0040)	(0.0044)
CAPX/Assets	-0.0124**	-0.0002	-0.0407***	-0.0020	-0.0057**	-0.0090***
	(0.0059)	(0.0084)	(0.0130)	(0.0085)	(0.0028)	(0.0030)
Asset intangibility	0.0044	0.0057*	-0.0036	0.0027	0.0046***	0.0001
	(0.0030)	(0.0033)	(0.0058)	(0.0062)	(0.0016)	(0.0015)
1(S&P500)	0.0153***	0.0154***	0.0005	0.0012	0.0208***	0.0208***
	(0.0025)	(0.0025)	(0.0042)	(0.0043)	(0.0025)	(0.0025)
1(Risk Committee)	-0.0158***	-0.0145***	-0.0146	-0.0144		
	(0.0040)	(0.0038)	(0.0120)	(0.0094)		
N Board Committees	0.0005	0.0005	-0.0009	-0.0011		
	(0.0005)	(0.0006)	(0.0009)	(0.0008)		
Industry HHI					0.0243***	
					(0.0094)	
1(Unique Industry)					0.0013	
					(0.0012)	
Industry Tobin's Q					-0.0004	
					(0.0006)	
1(Wholesale and retail trade)						0.0097***
						(0.0020)
1(Finance)						0.0052***
						(0.0013)
1(Service Industries)						0.0043***
						(0.0009)
1(Transportation and Communication)						-0.0001
						(0.0009)
Year FE	Yes	No	Yes	No	Yes	Yes
SIC2 FE	Yes	No	No	No	No	No
SIC2-by-Year FE	No	Yes	No	Yes	No	No
Firm FE	No	No	Yes	Yes	No	No
Observations	44544	44473	44257	44026	67880	67880
R^2	.0323	.0579	.167	.195	.0217	.023
Within-R^2	.0214	.0219	.000892	.000471	.0184	.0197

Electronic copy available at: https://ssrn.com/abstract=3143740

Continued...

(e) Determinants of RepRisk Events – Full Sample

	OLS: RRI Event (0/1)					
	(1)	(2)	(3)	(4)	(5)	(6)
Firm Size	0.0007***	0.0007***	-0.0002	-0.0000	0.0005***	0.0005***
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Log(Firm Age)	0.0005	0.0004	-0.0043**	-0.0023	0.0006**	0.0006**
	(0.0004)	(0.0004)	(0.0021)	(0.0022)	(0.0003)	(0.0003)
Tobin's Q (t-1)	0.0001*	0.0000	0.0001	-0.0001	0.0000	0.0000
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0000)	(0.0000)
ROA	-0.0006*	-0.0008**	0.0011***	0.0009**	-0.0006**	-0.0006**
	(0.0003)	(0.0003)	(0.0004)	(0.0004)	(0.0003)	(0.0003)
Sales Growth	0.0000	-0.0001	0.0001	-0.0001	0.0001	0.0001
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0000)	(0.0000)
BHAR 12-Mths	-0.0002	-0.0002	-0.0001	-0.0000	-0.0002	-0.0002
	(0.0002)	(0.0002)	(0.0001)	(0.0002)	(0.0001)	(0.0001)
Leverage	-0.0019	-0.0023	0.0001	-0.0010	-0.0017	-0.0017
	(0.0013)	(0.0015)	(0.0010)	(0.0014)	(0.0010)	(0.0012)
1(Financial Constraint)	0.0004	0.0004	0.0004	0.0003	0.0002	0.0002
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0002)	(0.0002)
Return Volatilty	0.0066	0.0066	-0.0044	-0.0031	0.0083*	0.0090*
	(0.0058)	(0.0056)	(0.0076)	(0.0070)	(0.0048)	(0.0053)
Inst. Block Own.	-0.0038	-0.0042	-0.0030	-0.0031	-0.0030	-0.0028
	(0.0028)	(0.0029)	(0.0036)	(0.0038)	(0.0020)	(0.0018)
R&D/Assets	0.0017	0.0018	0.0009	0.0029*	0.0027*	0.0019
	(0.0019)	(0.0019)	(0.0013)	(0.0017)	(0.0015)	(0.0018)
CAPX/Assets	-0.0012	-0.0011	0.0002	0.0009	-0.0003	-0.0010
	(0.0025)	(0.0025)	(0.0017)	(0.0021)	(0.0014)	(0.0016)
Asset intangibility	-0.0015	-0.0014	0.0014	0.0037	-0.0010	-0.0015
	(0.0014)	(0.0015)	(0.0026)	(0.0037)	(0.0007)	(0.0011)
1(S&P500)	0.0014**	0.0014**	0.0014	0.0015	0.0031***	0.0031***
	(0.0006)	(0.0007)	(0.0011)	(0.0016)	(0.0009)	(0.0010)
1(Risk Committee)	-0.0009	-0.0007	-0.0010	-0.0004		
	(0.0007)	(0.0007)	(0.0006)	(0.0006)		
N Board Committees	0.0001	0.0001	0.0004	0.0003		
	(0.0003)	(0.0003)	(0.0004)	(0.0004)		
Industry HHI					0.0137*	
					(0.0071)	
1(Unique Industry)					-0.0003	
					(0.0006)	
Industry Tobin's Q					0.0001	
					(0.0003)	
1(Wholesale and retail trade)						0.0023
						(0.0014)
1(Finance)						0.0001
						(0.0004)
1(Service Industries)						0.0002
						(0.0003)
1(Transportation and Communication)						0.0001
						(0.0006)
Year FE	Yes	No	Yes	No	Yes	Yes
SIC2 FE	Yes	No	No	No	No	No
SIC2-by-Year FE	No	Yes	No	Yes	No	No
Firm FE	No	No	Yes	Yes	No	No
Observations	34043	33988	33497	33283	43635	43635
R^2	.0287	.0806	.213	.266	.00687	.00627
Within-R^2	.00349	.00368	.000381	.000247	.00636	.00576

Electronic copy available at: https://ssrn.com/abstract=3143740

Table IA.4: Data Breaches and Brand Strength by Number of Brands per Firm

Notes: This table presents OLS regression results on the heterogeneous effects of data breaches on brand strength from BAV across firms with a high- and low number of individual brands. The dependent variable is the brand strength measure from BAV, and "Yrs 0-1" ("Yrs 0-4") indicates the occurrence of a data breach in the previous two (five) years. For each firm-year, we count the number of individual brands (such as Sprite, Coca Cola, and Fanta for Coca Cola Inc.) covered in the BAV and split the sample at the median to define '1(High N Brands)' and '1(Low N Brands)'. Year-by-industry (GIC) and firm-fixed effects are included as indicated. All control variables and data filters are similar to the corresponding panel in Table 2. 'Chi-Sq(Diff. High-Low)' and 'p(Chi-Sq)' report the results of a χ^2-test comparing coefficient estimates interacted with "1(High N Brands)" and "1(Low N Brands)". *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

	Brand Strength			
	(1)	(2)	(3)	(4)
Yrs 0-1 Post × 1(High N Brands)	-4.08*	-3.28**		
	(2.42)	(1.34)		
Yrs 0-1 Post × 1(Low N Brands)	-7.84*	-.555		
	(4.25)	(2.51)		
Yrs 0-4 Post × 1(High N Brands)			-6.27**	-6.16***
			(2.46)	(1.79)
Yrs 0-4 Post × 1(Low N Brands)			-7.91*	1.48
			(4.4)	(1.88)
Treated	5.16**		6.59***	
	(2.38)		(2.5)	
Controls	Yes	Yes	Yes	Yes
Yr × GIC FE	Yes	Yes	Yes	Yes
Firm FE	No	Yes	No	Yes
p(Chi-Sq)	.472	.337	.744	.00121
Chi-Sq (Diff. High-Low)	.518	.923	.107	10.6
Observations	4185	4121	4185	4121
R2	.396	.819	.397	.82

Electronic copy available at: https://ssrn.com/abstract=3143740

Table IA.5: Robustness – Reputation Shocks and Accounting Performance

Notes: This table presents OLS estimates of the effect of reputation shocks, i.e., data breaches (Panels IA.5a and IA.5b) and RepRisk events (Panels IA.5c and IA.5d), on measures of accounting performance. The dependent variables are the return-on-equity (ROE), price-earnings ratio (P/E), sales scaled by assets, and market-to-book ratio (M/B). "Years 0-1 Post" ("Years 0-4 Post") is an indicator variable that takes the value of one if a data breach or RepRisk event occurred in the current or previous one (four) years, and zero otherwise. Treated takes the value of one if a firm was ever affected by a reputation shock, and zero otherwise. Data breaches are included if the number of affected records is known and is at least 1,000. Firms are only included if there has ever been a data breach or RepRisk event, respectively, in their six-digit GIC industry. Controls in all panels include ln(Assets), ln(Assets)^2, and market leverage. Panels IA.5b and IA.5d additionally include controls indicating the presence of a Board "Risk Committee (0/1)", firms in which the CEO also holds the position of board chairman ("Dual Role CEO"), and the percentage of institutional block owners ("Inst. Block Own"). Compustat variables have been Winsorized at the 5th percentiles. Year-by-industry fixed effects ("Yr×GIC FE"), and firm fixed effects are included as indicated. Standard errors are clustered at the firm level and reported in parentheses. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

(a) Data Breaches

	ROE				P/E				Sales / Assets			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Years 0-1 Post	-.0637**	-.032			-3.38**	-3.13**			.0159	.0192*		
	(.0279)	(.0251)			(1.34)	(1.32)			(.022)	(.0115)		
Years 0-4 Post			-.0658***	-.0349*			-3.5***	-2.61**			.0162	-.0014
			(.0203)	(.0184)			(1.17)	(1.19)			(.0245)	(.0145)
Treated	.00272		.0135		2.19**		2.76***		.0271		.0245	
	(.0173)		(.0169)		(1.02)		(1.07)		(.0375)		(.0378)	
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year × GIC FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Observations	84125	82897	84125	82897	84117	82888	84117	82888	84130	82901	84130	82901
R^2	0.075	0.330	0.075	0.330	0.134	0.353	0.134	0.353	0.437	0.845	0.437	0.845

(b) Data Breaches – with Risk Management Controls

	M/B				ROE				P/E			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Years 0-1 Post	-.58***	-.359***			-.0533**	-.0385*			-2.17	-2.39*		
	(.145)	(.115)			(.0231)	(.0219)			(1.33)	(1.31)		
Years 0-4 Post			-.311**	-.146			-.0351*	-.0196			-2.33**	-1.93
			(.154)	(.135)			(.0188)	(.0182)			(1.17)	(1.19)
Treated	.503***		.51***		-.0126		-.00964		1.94*		2.37**	
	(.165)		(.163)		(.015)		(.0143)		(1.1)		(1.16)	
Risk Committee (0/1)	.656	-.148	.649	-.151	.0761	.0187	.0752	.0181	1.74	2.4	1.67	2.31
	(1.27)	(.678)	(1.27)	(.679)	(.0508)	(.0454)	(.0511)	(.0455)	(2.64)	(3.79)	(2.65)	(3.81)
Dual Role CEO (0/1)	-.0136	.0521	-.0138	.0516	.0247***	.00919	.0247***	.00916	.942***	.525	.944***	.525
	(.0471)	(.0444)	(.0471)	(.0444)	(.00609)	(.00808)	(.00609)	(.00808)	(.33)	(.426)	(.33)	(.426)
Inst. Block Own. (%)	.573*	-.384	.573*	-.383	.0857**	-.0384	.0859**	-.0382	1.73	-1.29	1.74	-1.27
	(.314)	(.306)	(.314)	(.306)	(.039)	(.0596)	(.039)	(.0596)	(2.06)	(2.79)	(2.06)	(2.79)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Yr × GIC FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Observations	37712	37274	37712	37274	39588	39183	39588	39183	39585	39179	39585	39179
R^2	0.257	0.706	0.257	0.706	0.117	0.387	0.117	0.387	0.122	0.363	0.122	0.363

Electronic copy available at: https://ssrn.com/abstract=3143740

Continued...

(c) RepRisk Events

	ROE				P/E				Sales / Assets			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Years 0-1 Post	-.0475***	-.0243*			-1.19*	-1.44**			.0187	-.0103		
	(.0137)	(.0128)			(.639)	(.677)			(.0181)	(.0114)		
Years 0-4 Post			-.0346***	-.00892			-1.38**	-1.58**			.00874	-.0142
			(.0127)	(.012)			(.616)	(.66)			(.0198)	(.0124)
Treated	.0359***		.0352***		1.38***		1.49***		.107***		.109***	
	(.00879)		(.00896)		(.465)		(.477)		(.0236)		(.0242)	
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Yr × GIC FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Observations	109691	108097	109691	108097	109683	108088	109683	108088	109701	108106	109701	108106
R^2	0.071	0.324	0.071	0.324	0.132	0.343	0.132	0.343	0.395	0.836	0.395	0.836

(d) RepRisk Events – with Risk Management Controls

	MtB				ROE				P/E			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Years 0-1 Post	-.311***	-.17**			-.023*	-.0174			-.579	-1.02		
	(.108)	(.0838)			(.0134)	(.0127)			(.683)	(.715)		
Years 0-4 Post			-.302***	-.145			-.0111	.00162			-.695	-.973
			(.114)	(.0898)			(.0127)	(.0124)			(.65)	(.693)
Treated	.497***		.51***		.00469		.00299		.497		.555	
	(.0953)		(.0964)		(.00886)		(.00912)		(.528)		(.539)	
Risk Committee (0/1)	.561	-.132	.568	-.128	.0747	.0166	.0752	.0164	1.65	2.43	1.67	2.46
	(1.36)	(.684)	(1.36)	(.684)	(.0478)	(.0474)	(.0475)	(.0475)	(2.68)	(3.78)	(2.68)	(3.77)
Dual Role CEO (0/1)	-.00894	.0533	-.00956	.0529	.0274***	.00546	.0274***	.00554	1.06***	.641*	1.06***	.636*
	(.0407)	(.0371)	(.0407)	(.0371)	(.00533)	(.00703)	(.00533)	(.00703)	(.284)	(.374)	(.284)	(.374)
Inst. Block Own. (%)	.252	-.401	.253	-.4	.112***	-.0477	.112***	-.0472	4.03**	.415	4.04**	.417
	(.267)	(.255)	(.267)	(.255)	(.0342)	(.0498)	(.0342)	(.0498)	(1.84)	(2.46)	(1.84)	(2.46)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Yr × GIC FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Observations	49026	48514	49026	48514	51425	50955	51425	50955	51424	50953	51424	50953
R^2	0.241	0.705	0.241	0.705	0.114	0.388	0.114	0.388	0.118	0.351	0.118	0.351

Electronic copy available at: https://ssrn.com/abstract=3143740

Table IA.6: Robustness — Alternative Measures of Charitable Donations

Notes: This table presents robustness tests for the results summarized in Table 5 with respect to corporate charitable donations. Panels IA.6a and IA.6c use only donations data obtained from Foundation Directory Online (FDO), Panels IA.6b and IA.6d use only donations data obtained from the National Center for Charitable Statistics (NCCS). Analogous to Table 5, "Years 0-1 Post" ("Years 0-4 Post") is a dummy that takes the value of one if a firm has disclosed a data breach or RepRisk event, respectively, in the current or previous year (previous four years), and zero otherwise. All other variables, controls, data filters, and fixed effects are similar as in Table 5. Standard errors are clustered at the firm level and reported in parentheses. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

(a) Data Breaches – Only FDO Data

	Charitable Donations (M. USD)									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Years 0-1 Post	6.97**	4.53***	4.01***	4.46***	2.74**					
	(2.8)	(1.71)	(1.46)	(1.5)	(1.08)					
Years 0-4 Post						5.54**	3.79***	3.48***	3.81***	2.39**
						(2.17)	(1.34)	(1.19)	(1.24)	(1.03)
Treated	2.74**	.784	.503	.466		2.09**	.304	.0495	-.0169	
	(1.19)	(.877)	(.806)	(.852)		(.991)	(.812)	(.768)	(.797)	
Controls	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	No	No	Yes	Yes	Yes	No	No
GIC FE	No	No	Yes	No	No	No	No	Yes	No	No
Year × GIC FE	No	No	No	Yes	Yes	No	No	No	Yes	Yes
Firm FE	No	No	No	No	Yes	No	No	No	No	Yes
Observations	5120	5083	5083	5018	5016	5120	5083	5083	5018	5016
R^2	0.073	0.325	0.406	0.468	0.726	0.072	0.325	0.406	0.469	0.727

(b) Data Breaches – Only NCCS Data

	Charitable Donations (M. USD)									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Years 0-1 Post	1.39***	.892***	.876***	.885***	.262*					
	(.433)	(.296)	(.283)	(.293)	(.154)					
Years 0-4 Post						1.46***	1.13***	1.14***	1.15***	.545***
						(.359)	(.283)	(.271)	(.276)	(.193)
Treated	2.83***	1.42***	1.41***	1.41***		2.59***	1.2***	1.19***	1.19***	
	(.501)	(.412)	(.384)	(.388)		(.482)	(.406)	(.376)	(.379)	
Controls	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	No	No	Yes	Yes	Yes	No	No
GIC FE	No	No	Yes	No	No	No	No	Yes	No	No
Year × GIC FE	No	No	No	Yes	Yes	No	No	No	Yes	Yes
Firm FE	No	No	No	No	Yes	No	No	No	No	Yes
Observations	30688	30439	30439	30435	30385	30688	30439	30439	30435	30385
R^2	0.079	0.261	0.301	0.310	0.878	0.080	0.262	0.303	0.312	0.879

Electronic copy available at: https://ssrn.com/abstract=3143740

Continued...

(c) RepRisk Events – Only FDO Data

	Charitable Donations (M. USD)									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Years 0-1 Post	5.99***	3.43***	3.23***	3.85***	3.08***					
	(1.24)	(.919)	(.83)	(.959)	(.608)					
Years 0-4 Post						5.41***	3.11***	2.98***	3.69***	3.26***
						(1.17)	(.872)	(.792)	(.911)	(.596)
Treated	2.18***	.701*	1.42**	1.3**		2.13***	.648	1.35**	1.2*	
	(.47)	(.41)	(.573)	(.616)		(.487)	(.417)	(.583)	(.625)	
Controls	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	No	No	Yes	Yes	Yes	No	No
GIC FE	No	No	Yes	No	No	No	No	Yes	No	No
Year × GIC FE	No	No	No	Yes	Yes	No	No	No	Yes	Yes
Firm FE	No	No	No	No	Yes	No	No	No	No	Yes
Observations	6724	6672	6672	6594	6591	6724	6672	6672	6594	6591
R^2	0.100	0.306	0.392	0.462	0.727	0.095	0.304	0.391	0.461	0.727

(d) RepRisk Events – Only NCCS Data

	Charitable Donations (M. USD)									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Years 0-1 Post	1.87***	1.18***	1.17***	1.2***	.631***					
	(.281)	(.248)	(.234)	(.24)	(.113)					
Years 0-4 Post						1.53***	.994***	1***	1.03***	.631***
						(.251)	(.223)	(.212)	(.22)	(.109)
Treated	1.81***	.919***	.94***	.922***		1.8***	.902***	.92***	.901***	
	(.194)	(.154)	(.161)	(.163)		(.198)	(.155)	(.161)	(.163)	
Controls	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	No	No	Yes	Yes	Yes	No	No
GIC FE	No	No	Yes	No	No	No	No	Yes	No	No
Year × GIC FE	No	No	No	Yes	Yes	No	No	No	Yes	Yes
Firm FE	No	No	No	No	Yes	No	No	No	No	Yes
Observations	39747	39435	39435	39426	39360	39747	39435	39435	39426	39360
R^2	0.125	0.262	0.299	0.308	0.871	0.121	0.261	0.298	0.307	0.871

Electronic copy available at: https://ssrn.com/abstract=3143740

Table IA.7: Robustness — Scaled Charitable Donations

Notes: This table presents robustness tests for the results summarized in Table 5 with respect to charitable donations. The dependent variable is the amount of charitable donations from FDO and NCCS, scaled by the firm's one-year lagged sales in Panels IA.7a and IA.7c, and the log-transformed charitable donations in Panels IA.7b and IA.7d. Analogous to Table 5, "Years 0-1 Post" ("Years 0-4 Post") is a dummy that takes the value of one if a firm has disclosed a data breach (Panels IA.7a and IA.7b) or RepRisk event (Panels IA.7c and IA.7d), respectively, in the current or previous year (previous four years), and zero otherwise. All other variables, controls, data filters, and fixed effects are similar as in Table 5. Standard errors are clustered at the firm level and reported in parentheses. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

(a) Data Breaches – Scaled by Revenues

	Donations (t) / Revenues (t-1)									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Years 0-1 Post	.0101**	.0074**	.00758**	.00799**	.00311					
	(.00415)	(.00341)	(.00331)	(.00343)	(.00215)					
Years 0-4 Post						.0105***	.00837**	.00901***	.00923***	.00381*
						(.00384)	(.00339)	(.00327)	(.00338)	(.00222)
Treated	.0156***	.00652	.00832*	.00797*		.0137***	.00486	.00647	.00611	
	(.00451)	(.00437)	(.00458)	(.00467)		(.00421)	(.00417)	(.0044)	(.00442)	
Controls	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	No	No	Yes	Yes	Yes	No	No
GIC FE	No	No	Yes	No	No	No	No	Yes	No	No
Year × GIC FE	No	No	No	Yes	Yes	No	No	No	Yes	Yes
Firm FE	No	No	No	No	Yes	No	No	No	No	Yes
Observations	26712	26524	26524	26513	26482	26712	26524	26524	26513	26482
R^2	0.006	0.018	0.057	0.076	0.716	0.006	0.019	0.057	0.076	0.716

(b) Data Breaches – Log-Transformed

	Log(1+Donations (M. USD))									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Years 0-1 Post	.289***	.212***	.211***	.213***	.137***					
	(.0859)	(.065)	(.0635)	(.0649)	(.048)					
Years 0-4 Post						.288***	.237***	.242***	.245***	.159***
						(.0725)	(.058)	(.0569)	(.0571)	(.0459)
Treated	.385***	.163***	.16***	.158***		.34***	.122**	.117**	.115**	
	(.0657)	(.055)	(.0537)	(.0542)		(.059)	(.0509)	(.0497)	(.0504)	
Controls	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	No	No	Yes	Yes	Yes	No	No
GIC FE	No	No	Yes	No	No	No	No	Yes	No	No
Year × GIC FE	No	No	No	Yes	Yes	No	No	No	Yes	Yes
Firm FE	No	No	No	No	Yes	No	No	No	No	Yes
Observations	30688	30439	30439	30435	30385	30688	30439	30439	30435	30385
R^2	0.069	0.214	0.238	0.256	0.707	0.071	0.216	0.240	0.257	0.708

Electronic copy available at: https://ssrn.com/abstract=3143740

Continued...

(c) RepRisk Events – Scaled by Revenues

	Donations (t) / Revenues (t-1)									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Years 0-1 Post	.012***	.00858***	.00931***	.0113***	.00731***					
	(.0026)	(.00236)	(.0023)	(.00231)	(.00139)					
Years 0-4 Post						.0128***	.0101***	.0111***	.0136***	.00845***
						(.00302)	(.00307)	(.0031)	(.00322)	(.00264)
Treated	.00894***	.0024	.00543**	.00483*		.00808***	.0015	.00442*	.00359	
	(.0025)	(.00263)	(.0027)	(.00273)		(.00221)	(.00226)	(.00238)	(.00242)	
Controls	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	No	No	Yes	Yes	Yes	No	No
GIC FE	No	No	Yes	No	No	No	No	Yes	No	No
Year × GIC FE	No	No	No	Yes	Yes	No	No	No	Yes	Yes
Firm FE	No	No	No	No	Yes	No	No	No	No	Yes
Observations	34607	34365	34365	34346	34301	34607	34365	34365	34346	34301
R^2	0.008	0.017	0.054	0.075	0.685	0.008	0.017	0.055	0.075	0.685

(d) RepRisk Events – Log-Transformed

	ln(1+Donations (M. USD))									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Years 0-1 Post	.401***	.295***	.295***	.304***	.22***					
	(.0496)	(.0443)	(.0426)	(.0432)	(.0263)					
Years 0-4 Post						.347***	.264***	.266***	.277***	.212***
						(.0458)	(.0413)	(.04)	(.0411)	(.0262)
Treated	.244***	.103***	.104***	.101***		.238***	.0952***	.0959***	.0917***	
	(.0263)	(.0234)	(.0245)	(.025)		(.0264)	(.0231)	(.0242)	(.0246)	
Controls	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	No	No	Yes	Yes	Yes	No	No
GIC FE	No	No	Yes	No	No	No	No	Yes	No	No
Year × GIC FE	No	No	No	Yes	Yes	No	No	No	Yes	Yes
Firm FE	No	No	No	No	Yes	No	No	No	No	Yes
Observations	39759	39447	39447	39438	39372	39759	39447	39447	39438	39372
R^2	0.115	0.219	0.243	0.260	0.703	0.112	0.218	0.243	0.260	0.703

Electronic copy available at: https://ssrn.com/abstract=3143740

Table IA.8: Robustness — Political Contributions

Notes: This table presents robustness tests for the results shown in Table 6, on the effect of data breaches (Panels IA.8a, IA.8c, and IA.8e) and RepRisk events (Panels IA.8b, IA.8d, and IA.8f) on political contributions. The first two panels use panel data at the annual frequency (rather than the election-cycle frequency) by evenly splitting political contributions across the two years in an election cycle. The next two panels use log-transformed political contributions. The last two panels additionally include controls for corporate governance ('E-Index' and 'G-Index') and risk management ('Risk Committee (0/1)', 'Dual Role CEO (0/1)', and 'Inst. Block Own (%)'), following Kamiya et al. (2021). All other variables, controls, data filters, and fixed effects are similar as in Table 6. Standard errors are clustered at the firm level and reported in parentheses. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

(a) Data Breaches – Annual

	Political Contributions (M.USD)					
	(1)	(2)	(3)	(4)	(5)	(6)
Years 0-1 Post	0.0907**	0.0604	0.0375**			
	(0.0420)	(0.0379)	(0.0173)			
Years 0-4 Post				0.0852**	0.0755**	0.0569***
				(0.0361)	(0.0317)	(0.0200)
Treated	0.1500***	0.0738***		0.1367***	0.0581***	
	(0.0335)	(0.0229)		(0.0313)	(0.0204)	
Controls	No	Yes	Yes	No	Yes	Yes
Year FE	Yes	No	No	Yes	No	No
Yr × GIC FE	No	Yes	Yes	No	Yes	Yes
Firm FE	No	No	Yes	No	No	Yes
Observations	8,175	8,131	8,131	8,175	8,131	8,131
R^2	0.1059	0.4643	0.8614	0.1075	0.4667	0.8628

(b) RepRisk Events – Annual

	Political Contributions (M.USD)					
	(1)	(2)	(3)	(4)	(5)	(6)
Years 0-1 Post	0.0964***	0.0731***	0.0435***			
	(0.0176)	(0.0148)	(0.0092)			
Years 0-4 Post				0.0867***	0.0694***	0.0467***
				(0.0189)	(0.0162)	(0.0108)
Treated	0.0888***	0.0364***		0.0872***	0.0338***	
	(0.0115)	(0.0080)		(0.0116)	(0.0081)	
Controls	No	Yes	Yes	No	Yes	Yes
Year FE	Yes	No	No	Yes	No	No
Yr × GIC FE	No	Yes	Yes	No	Yes	Yes
Firm FE	No	No	Yes	No	No	Yes
Observations	10,913	10,831	10,831	10,913	10,831	10,831
R^2	0.1408	0.4767	0.8639	0.1377	0.4761	0.8642

(c) Data Breaches – Log-transform

	Ln(1+Pol. Contr. M.USD)					
	(1)	(2)	(3)	(4)	(5)	(6)
Years 0-1 Post	0.0572**	0.0348	0.0213**			
	(0.0248)	(0.0214)	(0.0101)			
Years 0-4 Post				0.0535**	0.0457**	0.0338***
				(0.0223)	(0.0184)	(0.0115)
Treated	0.1118***	0.0511***		0.1035***	0.0413***	
	(0.0233)	(0.0153)		(0.0225)	(0.0144)	
Controls	No	Yes	Yes	No	Yes	Yes
Year FE	Yes	No	No	Yes	No	No
Yr × GIC FE	No	Yes	Yes	No	Yes	Yes
Firm FE	No	No	Yes	No	No	Yes
Observations	8,175	8,131	8,131	8,175	8,131	8,131
R^2	0.1091	0.5073	0.8862	0.1102	0.5091	0.8871

(d) RepRisk Events – Log-transform

	Log(1+Political Contributions)					
	(1)	(2)	(3)	(4)	(5)	(6)
Years 0-1 Post	0.0698***	0.0500***	0.0285***			
	(0.0114)	(0.0092)	(0.0057)			
Years 0-4 Post				0.0613***	0.0463***	0.0300***
				(0.0121)	(0.0100)	(0.0065)
Treated	0.0722***	0.0290***		0.0715***	0.0275***	
	(0.0086)	(0.0061)		(0.0087)	(0.0062)	
Controls	No	Yes	Yes	No	Yes	Yes
Year FE	Yes	No	No	Yes	No	No
Yr × GIC FE	No	Yes	Yes	No	Yes	Yes
Firm FE	No	No	Yes	No	No	Yes
Observations	10,913	10,831	10,831	10,913	10,831	10,831
R^2	0.1606	0.5210	0.8872	0.1569	0.5201	0.8873

(e) Data Breaches – Governance

	Political Contributions (M.USD)					
	(1)	(2)	(3)	(4)	(5)	(6)
Years 0-1 Post	0.1064**	0.1015**	0.0884*			
	(0.0524)	(0.0500)	(0.0486)			
Years 0-4 Post				0.1332***	0.1263**	0.1170**
				(0.0476)	(0.0494)	(0.0502)
Risk Committee (0/1)		0.0964*	0.1191*		0.3470***	0.3530***
		(0.0534)	(0.0649)		(0.1177)	(0.1255)
Dual Role CEO (0/1)		0.0267**	0.0317*		0.0265**	0.0317*
		(0.0131)	(0.0166)		(0.0130)	(0.0166)
Inst. Block Own (%)		0.0641	0.0376		0.0552	0.0196
		(0.0933)	(0.1533)		(0.0948)	(0.1567)
E-Index			0.00001			0.0007
			(0.0105)			(0.0103)
G-Index			-0.0027			-0.0033
			(0.0065)			(0.0065)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Cycle × GIC FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4,136	3,436	2,700	4,136	3,436	2,700
R^2	0.8599	0.8697	0.8794	0.8613	0.8709	0.8805

(f) RepRisk Events – Governance

	Political Contributions (M.USD)					
	(1)	(2)	(3)	(4)	(5)	(6)
Years 0-1 Post	0.1116***	0.1050***	0.1036***			
	(0.0241)	(0.0258)	(0.0264)			
Years 0-4 Post				0.1088***	0.1118***	0.1125***
				(0.0258)	(0.0284)	(0.0291)
Risk Committee (0/1)		0.0676	0.0803		0.1104**	0.1239**
		(0.0510)	(0.0587)		(0.0465)	(0.0555)
Dual Role CEO (0/1)		0.0232**	0.0243*		0.0242**	0.0249*
		(0.0103)	(0.0129)		(0.0105)	(0.0130)
Inst. Block Own (%)		0.0745	0.0491		0.0727	0.0543
		(0.0757)	(0.1215)		(0.0753)	(0.1210)
E-Index			0.0020			0.0024
			(0.0081)			(0.0081)
G-Index			-0.0027			-0.0032
			(0.0051)			(0.0051)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Cycle × GIC FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,505	4,581	3,578	5,505	4,581	3,578
R^2	0.8626	0.8718	0.8836	0.8626	0.8723	0.8840

Electronic copy available at: https://ssrn.com/abstract=3143740

Table IA.9: Robustness — CSR Response to Reputation Shocks: Alternative Measure of Consumer- vs. Business-facing Industries

Notes: This table presents robustness tests for the results summarized in Table 8 on the differential effect of reputation shocks (i.e., data breaches in Panel IA.9a and RepRisk events in Panel IA.9b) on CSR outcomes across consumer-facing and business-facing industries, using an alternative proxy for proximity to the consumer. "Adv. Ind" is constructed as the average ratio of advertising expenses to sales at the industry-level. We split this industry-level average at the median to delineate between consumer-facing (i.e, "Adv. Ind. = High") and business-facing (i.e., "Adv. Ind. = Low") industries. All variables, controls, data filters, and fixed effects are similar to the corresponding Table 8. 'Chi-Sq(Diff. Low-High)' and 'p(Chi-Sq)' report the results of a χ^2-test comparing coefficient estimates interacted with "Adv. Ind. = Low" and "Has Adv.-Ind. = High". Standard errors are clustered at the firm level and reported in parentheses. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

(a) Data Breaches

	Donations				1(Foundation)				CSR (KLD)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Years 0-1 × Adv. Ind. = Low	.691	.076			.0669	.0744			.0367	.127		
	(.742)	(.384)			(.0712)	(.0475)			(.285)	(.243)		
Years 0-1 × Adv. Ind. = High	1.26***	.894***			.119***	.0986***			.211	.136		
	(.439)	(.265)			(.0355)	(.0266)			(.171)	(.133)		
Years 0-4 × Adv. Ind. = Low			.792	.362			.158**	.165***			.325	.299
			(.696)	(.406)			(.0689)	(.0452)			(.241)	(.198)
Years 0-4 × Adv. Ind. = High			1.38***	.953***			.203***	.154***			.445**	.404***
			(.376)	(.25)			(.0379)	(.0296)			(.18)	(.153)
Treated	.668***		.453**		.108***		.0641**		.14		.0265	
	(.249)		(.226)		(.03)		(.029)		(.112)		(.116)	
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Yr × GIC FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
p(Chi-Sq)	.559	.0789	.506	.21	.559	.644	.595	.821	.625	.974	.701	.663
Chi-Sq(Diff. Low-High)	.341	3.09	.442	1.57	.341	.214	.282	.0511	.239	.00107	.147	.19
Observations	30366	30317	30366	30317	30366	30317	30366	30317	23089	22691	23089	22691
R2	.259	.717	.261	.718	.244	.773	.247	.774	.165	.605	.168	.606

(b) RepRisk Events

	Donations				1(Foundation)				CSR (KLD)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Years 0-1 × Adv. Ind. = Low	.539*	.632***			.133***	.142***			.145	.316***		
	(.319)	(.173)			(.0329)	(.02)			(.108)	(.0809)		
Years 0-1 × Adv. Ind. = High	1.97***	1.27***			.254***	.184***			.695***	.473***		
	(.292)	(.168)			(.0227)	(.0151)			(.101)	(.0722)		
Years 0-4 × Adv. Ind. = Low			.419	.56***			.114***	.129***			.218**	.373***
			(.284)	(.161)			(.0336)	(.0211)			(.109)	(.0857)
Years 0-4 × Adv. Ind. = High			1.83***	1.27***			.255***	.191***			.699***	.55***
			(.271)	(.16)			(.0248)	(.0165)			(.099)	(.0708)
Treated	.398***		.364***		.0875***		.0793***		-.0221		-.0499	
	(.11)		(.109)		(.0188)		(.0186)		(.0528)		(.0531)	
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Yr × GIC FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
p(Chi-Sq)	.00117	.00528	.000388	.000688	.00281	.0647	.000707	.00907	.000288	.117	.00114	.0813
Chi-Sq (Diff. Low-High)	10.6	7.79	12.6	11.5	8.94	3.42	11.5	6.82	13.2	2.45	10.6	3.04
Observations	39330	39265	39330	39265	39330	39265	39330	39265	29983	29494	29983	29494
R2	.266	.715	.265	.715	.257	.788	.257	.788	.186	.613	.187	.614

Electronic copy available at: https://ssrn.com/abstract=3143740

Table IA.10: Data Breaches and Employee Salaries: Differences in Labor Mobility

Notes: This table presents estimation results on the heterogeneous effects of data breaches on employee salaries across firms with high- and low labor mobility. The dependent variable in all panels is the annual salary (in $K) at the individual employee level, and "Yr 0-1" ("Yr 0-4") indicates the occurrence of a data breach in the previous two (five) years. All firm- and employee-level control variables and fixed effects are similar to Table 11. Panel IA.10a uses the percentage of employees at the firm with 'common' job titles (i.e., an occupation in the top 20 of all jobs in the sample) as a proxy for labor mobility. Panel IA.10b uses the proportion of employees with a post-graduate (i.e., MBA, MD, JD, PhD) degree, and Panel IA.10c uses the average employee-salary. We split each variable at the median to delineate between high- and low-labor mobility firms. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

(a) Labor mobility: percentage of employees with common job titles

	Annual Salary (Thsd. USD)			
	(1)	(2)	(3)	(4)
Yr 0-1 Post×1(High Mobility)	0.8261**	1.309***		
	(0.4138)	(0.4908)		
Yr 0-1 Post×1(Low Mobility)	0.4937	0.2802		
	(0.3451)	(0.3909)		
Yr 0-4 Post×1(High Mobility)			1.616***	1.506***
			(0.4000)	(0.4475)
Yr 0-4 Post×1(Low Mobility)			0.8129***	0.7248**
			(0.2623)	(0.3192)
Experience (Yrs)	1.874***	1.926***	1.874***	1.925***
	(0.0374)	(0.0375)	(0.0374)	(0.0375)
Firm Controls	Yes	Yes	Yes	Yes
Yr × GIC FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Metro Area FE	Yes	Yes	Yes	Yes
Occupation FE	Yes	Yes	Yes	Yes
Highest Degree FE	No	Yes	No	Yes
Gender FE	No	Yes	No	Yes
Observations	449,716	185,878	449,716	185,878
R^2	0.7172	0.7448	0.7173	0.7449

Electronic copy available at: https://ssrn.com/abstract=3143740

Continued...

(b) % of employees with post-graduate degree

	Annual Salary (Thsd. USD)			
	(1)	(2)	(3)	(4)
Yr 0-1 Post×1(High % Grad-Degree)	1.296**	1.474***		
	(0.5052)	(0.5609)		
Yr 0-1 Post×1(Low % Grad-Degree)	0.4282	0.3236		
	(0.2950)	(0.3293)		
Yr 0-4 Post×1(High % Grad-Degree)			1.794***	1.218**
			(0.5009)	(0.5285)
Yr 0-4 Post×1(Low % Grad-Degree)			0.9307***	1.221***
			(0.2617)	(0.2957)
Experience (Yrs)	1.878***	1.929***	1.878***	1.929***
	(0.0366)	(0.0365)	(0.0366)	(0.0365)
Firm Controls	No	No	No	No
Yr × GIC FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Metro Area FE	Yes	Yes	Yes	Yes
Occupation FE	Yes	Yes	Yes	Yes
Highest Degree FE	No	Yes	No	Yes
Gender FE	No	Yes	No	Yes
Observations	467,350	193,714	467,350	193,714
R^2	0.7166	0.7440	0.7166	0.7440

(c) Average employee salary

	Annual Salary (Thsd. USD)			
	(1)	(2)	(3)	(4)
Yr 0-1 Post×1(High Avg Salary)	1.301**	1.706***		
	(0.5364)	(0.5934)		
Yr 0-1 Post×1(Low Avg Salary)	0.4743*	0.2693		
	(0.2871)	(0.3121)		
Yr 0-4 Post×1(High Avg Salary)			1.886***	1.451**
			(0.5439)	(0.5693)
Yr 0-4 Post×1(Low Avg Salary)			0.9155***	1.065***
			(0.2501)	(0.2940)
Experience (Yrs)	1.878***	1.928***	1.878***	1.928***
	(0.0366)	(0.0366)	(0.0366)	(0.0365)
Firm Controls	No	No	No	No
Yr × GIC FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Metro Area FE	Yes	Yes	Yes	Yes
Occupation FE	Yes	Yes	Yes	Yes
Highest Degree FE	No	Yes	No	Yes
Gender FE	No	Yes	No	Yes
Observations	467,350	193,714	467,350	193,714
R^2	0.7166	0.7440	0.7166	0.7440

Electronic copy available at: https://ssrn.com/abstract=3143740

Table IA.11: Robustness — CSR Reaction and Governance Controls

Notes: This table presents robustness tests for the results summarized in Table 5 on the effect of reputation shocks, i.e., data breaches (Panel IA.11a) and RepRisk events (Panel IA.11b) on firms' CSR scores, charitable donations, and presence of a corporate charitable foundation. Relative to Table 5, we additionally include the "E-Index" and "G-Index" from Gompers et al. (2003) and Bebchuk et al. (2009) as corporate governance measures controls, and "Dual Role CEO (0/1)" which indicates whether the CEO is also the chair of the board, "Risk Committee (0/1)" which takes the value of one if the name of a firm's board committee includes *risk* and zero otherwise, and "Inst. Block Own", i.e., the percentage of shares owned by institutional block owners, as risk management controls, following Kamiya et al. (2021). "Years 0-1 Post" ("Years 0-4 Post") is an indicator variable that takes the value of one if a firm experienced a reputation shock in the current or previous year (previous four years), and zero otherwise. All other variables, controls, data filters, and fixed effects are similar as in Table 5. Standard errors are clustered at the firm level and reported in parentheses. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

(a) Data Breaches

	CSR (KLD)		Donations		1(Foundation)	
	(1)	(2)	(3)	(4)	(5)	(6)
Years 0-1 Post	.172		.215		.0124	
	(.129)		(.204)		(.0245)	
Years 0-4 Post		.403***		.291		.0589**
		(.148)		(.204)		(.0275)
E-Index	-.015	-.0138	-.0983*	-.0978*	.00301	.0032
	(.0291)	(.0289)	(.0558)	(.0557)	(.00848)	(.00843)
G-Index	-.00166	-.0013	.0284	.0283	-.00601	-.00595
	(.0189)	(.0188)	(.0355)	(.0355)	(.00541)	(.00536)
Risk Committee (0/1)	-.982***	-.973***	.0224	.0156	-.244**	-.247**
	(.291)	(.3)	(.423)	(.421)	(.102)	(.101)
Dual Role CEO (0/1)	.0463	.0465	-.0762	-.0755	.0198*	.0198*
	(.0373)	(.0366)	(.0748)	(.0747)	(.0108)	(.0108)
Inst. Block Own. (%)	-.22	-.259	.356	.33	.0498	.0438
	(.359)	(.354)	(.491)	(.487)	(.102)	(.102)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year × GIC FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	11749	11749	10952	10952	10952	10952
R^2	0.630	0.632	0.784	0.784	0.827	0.827

Electronic copy available at: https://ssrn.com/abstract=3143740

Continued...

(b) RepRisk Events

	CSR (KLD)		Donations		1(Foundation)	
	(1)	(2)	(3)	(4)	(5)	(6)
Years 0-1 Post	.323***		.694***		.0875***	
	(.0684)		(.136)		(.0143)	
Years 0-4 Post		.418***		.691***		.0874***
		(.0725)		(.138)		(.0161)
E-Index	-.0143	-.0106	-.041	-.0386	.01	.0103
	(.0245)	(.0243)	(.0483)	(.0483)	(.00707)	(.00706)
G-Index	-.000414	-.00173	.00261	.00182	-.00895**	-.00905**
	(.0156)	(.0155)	(.0288)	(.0289)	(.00452)	(.00452)
Risk Committee (0/1)	-.985***	-1.01***	-.029	-.0345	-.242**	-.242**
	(.279)	(.268)	(.43)	(.43)	(.102)	(.102)
Dual Role CEO (0/1)	.0345	.0363	-.0161	-.0123	.0182**	.0186**
	(.0303)	(.0303)	(.0598)	(.0599)	(.00888)	(.0089)
Inst. Block Own. (%)	-.209	-.204	.488	.498	.0899	.0912
	(.295)	(.294)	(.418)	(.42)	(.0858)	(.0862)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year × GIC FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	15809	15809	14591	14591	14591	14591
R^2	0.638	0.639	0.780	0.780	0.833	0.833

Electronic copy available at: https://ssrn.com/abstract=3143740

Table IA.12: Robustness — Employee Salaries and Governance Controls

Notes: This table presents robustness tests for the results shown in Table 7, on the effect of data breaches (Panel IA.12a) and RepRisk events (Panel IA.12b) on employee salaries. The data is organized at the individual employee level and each regression includes employee-level fixed effects for the employee's metro area and occupation. Relative to Table 7, we additionally include controls for corporate governance ('E-Index' and 'G-Index') and risk management ('Risk Committee (0/1)', 'Dual Role CEO (0/1)', and 'Inst. Block Own (%)'), following Kamiya et al. (2021). All other variables, controls, data filters, and fixed effects are similar as in Table 7. Standard errors are clustered at the firm level and reported in parentheses. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

(a) Data Breaches

	Annual Salary (Thsd. USD)					
	(1)	(2)	(3)	(4)	(5)	(6)
Years 0-1 Post	0.6247**	0.5077**	0.3615			
	(0.2665)	(0.2519)	(0.2589)			
Years 0-4 Post				1.194***	1.163***	1.068***
				(0.2402)	(0.2539)	(0.2642)
Experience (Yrs)	1.875***	1.849***	1.844***	1.875***	1.849***	1.844***
	(0.0374)	(0.0396)	(0.0421)	(0.0374)	(0.0396)	(0.0421)
Risk Committee (0/1)		0.3444	0.0404		0.1789	-0.1833
		(0.4916)	(0.6056)		(0.5191)	(0.6349)
Dual Role CEO (0/1)		-0.0189	0.2166		0.0377	0.2842
		(0.2782)	(0.3272)		(0.2803)	(0.3297)
Inst. Block Own (%)		7.236***	11.48***		6.947***	11.06***
		(2.284)	(3.667)		(2.330)	(3.741)
E-Index			-0.4912**			-0.4475**
			(0.2135)			(0.2102)
G-Index			0.4103**			0.4203**
			(0.1749)			(0.1712)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Yr × GIC FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Metro Area FE	Yes	Yes	Yes	Yes	Yes	Yes
Occupation FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	449,702	404,325	316,543	449,702	404,325	316,543
R^2	0.7172	0.7223	0.7252	0.7173	0.7224	0.7252

(b) RepRisk Events

	Annual Salary (Thsd. USD)					
	(1)	(2)	(3)	(4)	(5)	(6)
Years 0-1 Post	0.2105	0.3291*	0.1433			
	(0.1991)	(0.1948)	(0.2169)			
Years 0-4 Post				0.1880	0.2394	0.0201
				(0.2513)	(0.2511)	(0.2926)
Experience (Yrs)	1.861***	1.837***	1.843***	1.861***	1.837***	1.842***
	(0.0338)	(0.0360)	(0.0382)	(0.0338)	(0.0360)	(0.0382)
Risk Committee (0/1)		0.1785	-0.1234		0.1779	-0.1213
		(0.4942)	(0.6047)		(0.4939)	(0.6051)
Dual Role CEO (0/1)		-0.0147	0.2408		-0.0026	0.2399
		(0.2545)	(0.3004)		(0.2537)	(0.2999)
Inst. Block Own (%)		6.375***	9.088***		6.296***	9.034***
		(2.160)	(3.411)		(2.151)	(3.399)
E-Index			-0.5323***			-0.5414***
			(0.2033)			(0.2043)
G-Index			0.4343**			0.4387**
			(0.1696)			(0.1712)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Yr × GIC FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Metro Area FE	Yes	Yes	Yes	Yes	Yes	Yes
Occupation FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	504,725	452,496	355,100	504,725	452,496	355,100
R^2	0.7134	0.7188	0.7222	0.7134	0.7188	0.7222

Electronic copy available at: https://ssrn.com/abstract=3143740

Table IA.13: Robustness — Risk Management Interactions

Notes: This table presents robustness tests for the results presented in Table 5. We additionally include interaction terms of data breach indicators with the risk management measure "Risk Committee (0/1)". The dependent variables are the amount of donations ($M), existence of a corporate charitable foundation, and scaled CSR score from KLD, respectively. "Years 0-1 Post" ("Years 0-4 Post") is an indicator variable that takes the value of one if a firm experienced a reputation shock in the current or previous year (previous four years), and zero otherwise. All other variables, controls, data filters, and fixed effects are similar as in Table 5. Standard errors are clustered at the firm level and reported in parentheses. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

(a) Data Breaches

	Donations (M. USD)		1(Foundation)		Norm CSR (KLD)	
	(1)	(2)	(3)	(4)	(5)	(6)
Years 0-1 Post	.388**		.0557**		.0999	
	(.192)		(.0238)		(.122)	
Years 0-4 Post		.385**		.11***		.434***
		(.17)		(.0276)		(.139)
Years 0-1 Post × 1(Risk Committee)	.75		-.00706		.491	
	(.7)		(.0571)		(.401)	
Years 0-4 Post × 1(Risk Committee)		1.38**		-.02		.00145
		(.671)		(.0589)		(.36)
Risk Committee (0/1)	.27**	.21	.0493**	.0499**	.0678	.0779
	(.135)	(.129)	(.0211)	(.0213)	(.0695)	(.0652)
Dual Role CEO (0/1)	1.91e-06	-.000815	.0202***	.0199***	.0246	.0237
	(.0405)	(.04)	(.00767)	(.00761)	(.0271)	(.0267)
Inst. Block Own. (%)	.0444	.0322	.00805	.00469	-.117	-.138
	(.213)	(.21)	(.0521)	(.0525)	(.202)	(.201)
Other Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year × GIC FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	21947	21947	21947	21947	19589	19589
R^2	0.746	0.747	0.805	0.806	0.599	0.601

(b) RepRisk Events

	Donations (M. USD)		1(Foundation)		Norm CSR (KLD)	
	(1)	(2)	(3)	(4)	(5)	(6)
Years 0-1 Post	.743***		.129***		.362***	
	(.123)		(.0141)		(.0637)	
Years 0-4 Post		.747***		.131***		.444***
		(.121)		(.0152)		(.0649)
Years 0-1 Post × 1(Risk Committee)	1.04**		-.00205		.249	
	(.457)		(.0347)		(.183)	
Years 0-4 Post × 1(Risk Committee)		.809*		-.00523		.175
		(.418)		(.0354)		(.181)
Risk Committee (0/1)	.115	.108	.0382*	.0371*	.0358	.0381
	(.0965)	(.0973)	(.0201)	(.0202)	(.0625)	(.0605)
Dual Role CEO (0/1)	.0283	.0335	.0176***	.0184***	.0131	.0164
	(.0339)	(.0341)	(.00632)	(.00632)	(.0224)	(.0224)
Inst. Block Own. (%)	.153	.15	.0173	.017	-.0688	-.0723
	(.177)	(.177)	(.0432)	(.0433)	(.173)	(.173)
Other Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year × GIC FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	28721	28721	28721	28721	25615	25615
R^2	0.743	0.743	0.815	0.815	0.606	0.608

Electronic copy available at: https://ssrn.com/abstract=3143740

Table IA.14: Robustness — Difference-in-Difference Estimation with Staggered Treatment

Notes: This table presents robustness tests with respect to our findings in Tables 2, 3, 4, 5, and 6 on the effect of reputation shocks, i.e., data breaches (Panels IA.14a and IA.14b) and RepRisk events (Panels IA.14c and IA.14d), on CSR outcomes, political contributions, and employee salaries. The estimates presented in this table explicitly account for the staggered treatment of firms with reputation shocks in our empirical setting by implementing the staggered- and heterogeneous treatment-robust difference-in-difference estimators of Sun and Abraham (2021) (Panels IA.14a and IA.14c) and Gardner (2022) (Panels IA.14b and IA.14d). The Sun and Abraham (2021) estimator requires the explicit declaration of a reference period, which we set to 'all', 't=0', and 't=1', respectively, in Panels IA.14a and IA.14c. "ATT" represents the 'average-treatment effect on the treated' for the effect of reputation shocks on the respective outcome. Panels IA.14b and IA.14d do not include 'employee salaries' as a dependent variable due to computational limitations. Controls, data filters, fixed effects, and standard errors are similar to the corresponding estimations summarized in Tables 2, 3, 4, 5, and 6. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

(a) Data Breaches – Sun & Abraham (2021)

	CSR (KLD)			1(Foundation)			Donations (USD)			Pol. Contrib.			Salary		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
ATT	0.3196**	0.3468***	0.3324***	0.2146***	0.0418**	0.0618***	2.394***	0.2099	0.0552	0.1019***	0.0445	0.0455**	0.8079	0.8468*	0.9124**
	(0.1623)	(0.1237)	(0.1283)	(0.0417)	(0.0196)	(0.0213)	(0.5102)	(0.1737)	(0.1595)	(0.0336)	(0.0304)	(0.0232)	(0.9361)	(0.4583)	(0.4125)
Ref. Period	All	t=0	t=-1	All	t=0	t=-1	All	t=0	t=-1	All	t=0	t=-1	All	t=0	t=-1
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year×GIC FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Metro Area FE	No	No	No	No	No	No	No	No	No	No	No	No	Yes	Yes	Yes
Occupation FE	No	No	No	No	No	No	No	No	No	No	No	No	Yes	Yes	Yes
Highest Degree FE	No	No	No	No	No	No	No	No	No	No	No	No	Yes	Yes	Yes
Gender FE	No	No	No	No	No	No	No	No	No	No	No	No	Yes	Yes	Yes
Observations	22,257	22,257	22,257	20,425	20,425	20,425	22,497	22,497	22,497	7,722	7,722	7,722	208,117	208,117	208,117
R^2	0.1659	0.6188	0.6188	0.3131	0.9110	0.9110	0.4433	0.9159	0.9159	0.4969	0.8806	0.8813	0.6570	0.6810	0.6810

Electronic copy available at: https://ssrn.com/abstract=3143740

Continued...

(b) Data Breaches – Gardner (2022)

	CSR (KLD)		1(Foundation)		Donations (USD)		Pol. Contrib.	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ATT	0.2676	0.3709**	0.1933***	0.0615**	3.264***	0.0933	0.1910**	0.0638**
	(0.1708)	(0.1518)	(0.0363)	(0.0268)	(0.7960)	(0.2217)	(0.0760)	(0.0290)
Year×GIC FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	No	Yes	No	Yes	No	Yes	No	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	22,257	22,252	20,425	20,421	20,425	20,421	8,037	7,989
R^2	0.0012	0.0051	0.0118	0.0083	0.0600	0.0003	0.0862	0.0378

(c) RepRisk Events – Sun & Abraham (2021)

	CSR (KLD)			1(Foundation)			Donations (USD)			Pol. Contrib.		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
ATT	0.5354***	0.3475***	0.2898**	0.2529***	0.0042	0.0154**	1.907***	0.0272	0.2511***	0.1005***	0.0050	0.0088**
	(0.1657)	(0.1309)	(0.1404)	(0.0306)	(0.0058)	(0.0068)	(0.3221)	(0.0598)	(0.0724)	(0.0168)	(0.0032)	(0.0037)
Ref. Period	All	t=0	t=-1	All	t=0	t=-1	All	t=0	t=-1	All	t=0	t=-1
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year×GIC FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Observations	24,626	24,626	24,626	31,081	31,081	31,081	31,081	31,081	31,081	14,316	14,316	14,316
R^2	0.1657	0.6416	0.6417	0.4491	0.9226	0.9226	0.5077	0.9071	0.9072	0.6053	0.9181	0.9180

(d) RepRisk Events – Gardner (2022)

	CSR (KLD)		1(Foundation)		Donations (USD)		Pol. Contrib.	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ATT	0.6928*	0.4458**	0.2198***	0.0750***	2.020***	0.6191***	0.1052***	0.0573***
	(0.3600)	(0.1862)	(0.0230)	(0.0185)	(0.4848)	(0.2136)	(0.0269)	(0.0201)
Year×GIC FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	No	Yes	No	Yes	No	Yes	No	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	24,626	24,623	31,074	31,074	31,074	31,074	14,295	14,219
R^2	0.0084	0.0078	0.0530	0.0323	0.0485	0.0200	0.0576	0.0432

Electronic copy available at: https://ssrn.com/abstract=3143740

Table IA.15: Robustness — Alternative CSR Measures

Notes: This table presents robustness tests for the results summarized in Table 5 using an alternative measures of CSR. In both panels, the dependent variable in columns (1) and (2) is the ESG score from Thomson Reuters' Asset4 database. We further break down the ESG score into its three components (Environment, Social, and Governance) in columns (3)–(8). All dependent variables have been standardized to have a mean of zero and a standard deviation of one. "Years 0-1 Post" ("Years 0-4 Post") is an indicator variable that takes the value of one if a firm was affected by a data breach (Panel IA.15a) or RepRisk event (Panel IA.15b) in the current or previous year (previous four years), and zero otherwise. Data filters, control variables, and fixed effects in both panels are analogous to Table 5. Standard errors are clustered at the firm level and reported in parentheses. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

(a) Data Breaches

	ESG (A4)		E (A4)		S (A4)		G (A4)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Years 0-1 Post	.059**		.0579*		.0522		.0676***	
	(.029)		(.0337)		(.0354)		(.0246)	
Years 0-4 Post		.0689*		.0768*		.0561		.047
		(.0375)		(.0404)		(.0426)		(.0288)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year × GIC FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7396	7396	7390	7390	7396	7396	7396	7396
R^2	0.913	0.913	0.901	0.901	0.882	0.882	0.849	0.849

(b) RepRisk Events

	ESG (A4)		E (A4)		S (A4)		G (A4)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Years 0-1 Post	.0524**		.0948***		-.0404*		.0465***	
	(.021)		(.0231)		(.0224)		(.0166)	
Years 0-4 Post		.0565**		.111***		-.0574**		.0548***
		(.0253)		(.0283)		(.0265)		(.0199)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year × GIC FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	10202	10202	10190	10190	10202	10202	10202	10202
R^2	0.915	0.915	0.905	0.905	0.884	0.884	0.849	0.849

Electronic copy available at: https://ssrn.com/abstract=3143740

Table IA.16: Robustness — CSR Reaction to Data Breaches: Excluding Data Security and Privacy KLD Items

Notes: This table shows robustness tests for the results summarized in Table 5 with respect to the effect of data breaches on CSR scores. The dependent variable is the CSR score from KLD, constructed by excluding all data items that are potentially related to data or IT security. Similar to Table 5, "Years 0-1 Post" is an indicator variable that takes the value of one if a firm has disclosed a data breach in the current or previous year, and zero otherwise and "Years 0-4 Post" indicates whether a firm has disclosed a data breach within the past five years. "Treated" takes the value of one if a firm was ever affected by a data breach, and zero otherwise. Data breaches are included if the number of affected records is known and is at least 1,000. Firms are only included if there has ever been a data breach in their respective six-digit GIC industry. Controls include ln(Assets), ln(Assets)^2, and market leverage. Compustat variables have been Winsorized at the 5th percentiles. Year fixed effects, industry fixed effects (GIC), Year-by-industry fixed effects ("Yr×GIC FE"), and firm fixed effects are included as indicated. Standard errors are clustered at the firm level and reported in parentheses. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

	Dependent Variable: Norm CSR (KLD) (excl. data security items)									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Years 0-1 Post	.331**	.253*	.238*	.197	.14					
	(.155)	(.143)	(.141)	(.143)	(.122)					
Years 0-4 Post						.496***	.461***	.437***	.407***	.355***
						(.15)	(.147)	(.145)	(.147)	(.131)
Treated	.447***	.17	.109	.119		.332***	.056	.000622	.0134	
	(.116)	(.117)	(.11)	(.112)		(.119)	(.12)	(.114)	(.116)	
Controls	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	No	No	Yes	Yes	Yes	No	No
GIC FE	No	No	Yes	No	No	No	No	Yes	No	No
Yr × GIC FE	No	No	No	Yes	Yes	No	No	No	Yes	Yes
Firm FE	No	No	No	No	Yes	No	No	No	No	Yes
Observations	23275	23152	23152	23137	22738	23275	23152	23152	23137	22738
R^2	0.034	0.068	0.110	0.155	0.603	0.037	0.071	0.112	0.157	0.604

Electronic copy available at: https://ssrn.com/abstract=3143740

Table IA.17: Robustness — Coefficient Stability with Oster Bounds

Notes: As further robustness tests for the results presented in Tables 2, 3, 4, 5, and 6, this table tests the "coefficient stability" of our main estimates using the method proposed in Oster (2019). In all four panels, β^* represents the "lower bound" of a coefficient if there existed proportionate selection on unobservables that was equally important as the controls included in our model (i.e., $\delta = 1$). δ^* is a critical statistic that indicates how stable the "controlled" estimate is, i.e., how much variation would have to be explained by unobservables relative to observables in the model for the estimated coefficient of interest to be equal to zero (i.e., $\beta = 0$). Negative values of δ^* indicate that the coefficient increases in magnitude when covariates are included. (Absolute) values of δ^* that are greater than one are considered "robust" by Oster (2019). Data filters, dependent variables, controls, and fixed effects are similar to the corresponding Tables 2, 3, 4, 5, and 6. Standard errors are clustered at the firm level and reported in parentheses. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

(a) Data Breaches – Years 0-1

	CSR (KLD)		Donations (M.USD)		1(Foundation)		Pol. Contrib.		1(IT Sec.)		M/B	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Years 0-1 Post	.34**	.18	1.38***	.859***	.144***	.108***	.0907**	.0603	.0756**	.0603*	-.748***	-.856***
	(.15)	(.137)	(.481)	(.323)	(.0308)	(.0274)	(.042)	(.0378)	(.0316)	(.0311)	(.15)	(.151)
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Treated Coef.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Yr × GIC FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Specification	Uncontr.	Contr.	Uncontr.	Contr.	Uncontr.	Contr.	Uncontr.	Contr.	Uncontr.	Contr.	Uncontr.	Contr.
$\beta^*\|\delta = 1$		.122		.636		.0913		.0462		.0506		-.903
$\delta^*\|\beta = 0$		2.99		3.68		5.86		3.69		5.82		-20.3
R2	.0381	.168	.0626	.309	.07	.245	.106	.464	.13	.292	.0212	.274

(b) Data Breaches – Years 0-4

	CSR (KLD)		Donations (M.USD)		1(Foundation)		Pol. Contrib.		1(IT Sec.)		M/B	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Years 0-4 Post	.542***	.424***	1.23***	.907***	.209***	.192***	.0852**	.0755**	.144***	.127***	-.558***	-.599***
	(.149)	(.145)	(.384)	(.285)	(.0315)	(.0304)	(.0361)	(.0316)	(.0326)	(.0315)	(.175)	(.165)
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Treated Coef.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Yr × GIC FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Specification	Uncontr.	Contr.	Uncontr.	Contr.	Uncontr.	Contr.	Uncontr.	Contr.	Uncontr.	Contr.	Uncontr.	Contr.
$\beta^*\|\delta = 1$		.377		.767		.184		.0711		.115		-.622
$\delta^*\|\beta = 0$		7.54		5.92		15.5		9.77		9.09		-29.7
R2	.0404	.17	.063	.31	.0722	.248	.107	.466	.131	.292	.0212	.274

Electronic copy available at: https://ssrn.com/abstract=3143740

Continued...

(c) RepRisk Events – Years 0-1

	CSR (KLD)		Donations (M. USD)		1(Foundation)		Pol. Contrib.		Reg. News Sent.		M/B	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Years 0-1 Post	.653***	.479***	1.84***	1.41***	.268***	.208***	.0964***	.0731***	-.0128**	-.021***	-.123	-.508***
	(.0828)	(.0752)	(.251)	(.216)	(.0194)	(.0189)	(.0176)	(.0148)	(.00532)	(.0052)	(.111)	(.107)
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Treated Coef.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Yr × GIC FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Specification	Uncontr.	Contr.	Uncontr.	Contr.	Uncontr.	Contr.	Uncontr.	Contr.	Uncontr.	Contr.	Uncontr.	Contr.
$\beta^*\|\delta = 1$		.395		1.16		.165		.0599		-.0254		-.642
$\delta^*\|\beta = 0$		4.18		4.21		3.72		4.03		-6.04		-4.08
R2	.0549	.185	.106	.262	.132	.257	.141	.476	.0505	.126	.0223	.267

(f) RepRisk Events – Years 0-4

	CSR (KLD)		Donations (M. USD)		1(Foundation)		Pol. Contrib.		Reg. News Sent.		M/B	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Years 0-4 Post	.649***	.507***	1.57***	1.25***	.246***	.199***	.0867***	.0694***	-.0111**	-.0164***	-.213*	-.516***
	(.0797)	(.0745)	(.226)	(.199)	(.0208)	(.0204)	(.0189)	(.0162)	(.00544)	(.00528)	(.114)	(.11)
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Treated Coef.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Yr × GIC FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Specification	Uncontr.	Contr.	Uncontr.	Contr.	Uncontr.	Contr.	Uncontr.	Contr.	Uncontr.	Contr.	Uncontr.	Contr.
$\beta^*\|\delta = 1$		.438		1.07		.165		.0594		-.0191		-.62
$\delta^*\|\beta = 0$		4.88		4.89		4.21		4.61		-8.3		-5.41
R2	.0558	.187	.102	.26	.131	.257	.138	.476	.0505	.125	.0223	.267

Electronic copy available at: https://ssrn.com/abstract=3143740

Table IA.18: Robustness — Propensity Score Matching

Notes: This table presents propensity score matching (PSM) tests analogous to our results in Tables 2, 3, 4, 5, and 6, for data breaches (Panels IA.18a and IA.18b) and RepRisk events (Panels IA.18c and IA.18d). Specifically, we match each treated firm to its $k = 10$ nearest neighbors within the same industry (GIC 6-digit) and year, using propensity scores based on CSR (KLD) score, E-Index, G-Index, log(Assets), squared log(Assets), Leverage, log(MtB), ROA, and Sales, each observed in the year before the reputation shock. We exclude firms outside of the common support based on a caliper of 0.15. "Years 0-1 Post" ("Years 0-4 Post") indicates that the firm experienced a reputation shock in the current or previous year (previous four years). "Treated" takes the value of one if a firm was ever affected by a reputation event, and zero otherwise. We include event-, event-by-firm, and time ('t') fixed effects (capturing the year relative to the event year) as indicated. Data filters, dependent variables, and control variables are similar to the corresponding Tables 2, 3, 4, 5, and 6. Standard errors are clustered at the firm level and reported in parentheses.*, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

(a) Data Breaches – Years 0-1

	CSR (KLD)		Donations		1(Foundation)		Pol. Contributions		IT Sec. (0/1)		M/B	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Post 1 × Treated	.0823	.0862	.506**	.324*	.164***	.132***	.0247**	.0419***	.0602*	.0659*	-.25**	-.2*
	(.105)	(.0882)	(.241)	(.187)	(.0306)	(.028)	(.0115)	(.0132)	(.0332)	(.0333)	(.108)	(.104)
Treated	.0105		-.0907		.0333		.00912		.0025		.688***	
	(.101)		(.206)		(.0313)		(.011)		(.0248)		(.16)	
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Event FE	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Event × Firm FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
t FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5188	5177	5039	5036	5039	5036	2543	2533	4552	4546	5527	5524
R^2	0.205	0.695	0.433	0.794	0.298	0.817	0.617	0.880	0.301	0.642	0.283	0.784

(b) Data Breaches – Years 0-4

	CSR (KLD)		Donations		1(Foundation)		Pol. Contributions		IT Sec. (0/1)		M/B	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Post 4 × Treated	.276**	.315***	.236	.094	.207***	.152***	.0095	.0352**	.0398	.0746**	-.255**	-.191*
	(.11)	(.0993)	(.253)	(.189)	(.0354)	(.0313)	(.0125)	(.0147)	(.0358)	(.0349)	(.114)	(.11)
Treated	-.0389		-.225		.032		.0113		-.0105		.731***	
	(.0987)		(.212)		(.0316)		(.011)		(.0251)		(.157)	
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Event FE	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Event × Firm FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
t FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7515	7513	7361	7361	7361	7361	3797	3794	6737	6735	8044	8044
R^2	0.213	0.605	0.449	0.797	0.329	0.836	0.604	0.857	0.295	0.607	0.316	0.753

Electronic copy available at: https://ssrn.com/abstract=3143740

(c) RepRisk Events – Years 0-1

	CSR (KLD)		Donations		1(Foundation)		Pol. Contributions		News Sent.		M/B	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Post 1 × Treated	.147***	.184***	.0557	.0841	.0435***	.0331***	.0297***	.027***	-.0333**	-.0199	-.182***	-.168***
	(.0456)	(.0427)	(.0746)	(.0552)	(.0124)	(.00789)	(.00791)	(.00611)	(.0133)	(.0132)	(.0616)	(.0585)
Treated	-.249***	0	-.0164	0	.00652	0	.00779	0	.021*	0	.314***	0
	(.0555)	(.)	(.113)	(.)	(.0286)	(.)	(.00693)	(.)	(.0126)	(.)	(.0913)	(.)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Event FE	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Event × Firm FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
t FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	17556	17511	15925	15922	15925	15922	7406	7364	16033	16000	18363	18360
R^2	0.176	0.610	0.378	0.862	0.281	0.905	0.558	0.889	0.122	0.409	0.243	0.770

(d) RepRisk Events – Years 0-4

	CSR (KLD)		Donations		1(Foundation)		Pol. Contributions		News Sent.		M/B	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Post 4 × Treated	.242***	.273***	.0732	.148***	.0347***	.0374***	.0241***	.0241***	-.0299**	-.0201	-.21***	-.194***
	(.0466)	(.0432)	(.0712)	(.0495)	(.0128)	(.00802)	(.00785)	(.00679)	(.0129)	(.0126)	(.062)	(.0559)
Treated	-.297***	0	-.0512	0	-.00207	0	.00539	0	.0175	0	.334***	0
	(.0551)	(.)	(.111)	(.)	(.0275)	(.)	(.00729)	(.)	(.0127)	(.)	(.0889)	(.)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Event FE	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Event × Firm FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
t FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	21970	21925	20317	20317	20317	20317	9575	9554	20495	20432	23265	23265
R^2	0.181	0.562	0.346	0.834	0.288	0.890	0.551	0.869	0.117	0.405	0.229	0.736

Electronic copy available at: https://ssrn.com/abstract=3143740

european corporate governance institute

about ECGI

The European Corporate Governance Institute has been established to improve *corporate governance through fostering independent scientific research and related activities.*

The ECGI will produce and disseminate high quality research while remaining close to the concerns and interests of corporate, financial and public policy makers. It will draw on the expertise of scholars from numerous countries and bring together a critical mass of expertise and interest to bear on this important subject.

The views expressed in this working paper are those of the authors, not those of the ECGI or its members.

www.ecgi.global

european corporate governance institute

ECGI Working Paper Series in Finance

Editorial Board

Editor	Mike Burkart, Professor of Finance, London School of Economics and Political Science
Consulting Editors	Renée Adams, Professor of Finance, University of Oxford Franklin Allen, Nippon Life Professor of Finance, Professor of Economics, The Wharton School of the University of Pennsylvania Julian Franks, Professor of Finance, London Business School Mireia Giné, Associate Professor, IESE Business School Marco Pagano, Professor of Economics, Facoltà di Economia Università di Napoli Federico II
Editorial Assistant	Asif Malik, Working Paper Series Manager

www.ecgi.global/content/working-papers

european corporate governance institute

Electronic Access to the Working Paper Series

The full set of ECGI working papers can be accessed through the Institute's Web-site (www.ecgi.global/content/working-papers) or SSRN:

Finance Paper Series	http://www.ssrn.com/link/ECGI-Fin.html
Law Paper Series	http://www.ssrn.com/link/ECGI-Law.html

www.ecgi.global/content/working-papers