10.5465/AMPROC.2025.80bp

CORPORATE SOCIAL (IR)RESPONSIBILITY AND CYBERATTACKS: A SCREENING PERSPECTIVE

XIAOWEN TAN
Naveen Jindal School of Management
The University of Texas at Dallas
800 W Campbell Rd, Richardson, TX 75080

CUILI QIAN
Naveen Jindal School of Management
The University of Texas at Dallas

INTRODUCTION

For the past decades, research on corporate social responsibility (CSR) and irresponsibility (CSIR) has explored various consequences of CSR/CSIR on a firm and its stakeholders (Hericher & Bridoux, 2023; Shea & Hawn, 2019). Despite various stakeholders' responses to a firm's CSR/CSIR, the current research implicitly assumes stakeholders generally conform to social expectations and make corresponding responses to CSR or CSIR. However, limited attention has been paid to understanding how other audiences who are less conforming to social expectations, such as cybercriminals, view a firm's CSR/CSIR and whether they will make the same responses accordingly. In this paper, we aim to explore the research question of how a firm's CSR and CSIR activities would affect its likelihood of becoming a cyberattack target.

CSR is defined as "a business organization's configuration of principles of social responsibility, processes of social responsiveness, and policies, programs, and observable outcomes as they relate to the firm's social relationships" (Wood, 1991: 693). The general arguments are that CSR contributes to firm value creation and preservation through building moral value that enables a firm to garner support from stakeholders (Flammer, 2018; Wang & Qian, 2011) or to work as insurance-like protection that buffers the firm from adverse events (Godfrey, Merrill, & Hansen, 2009). On the other hand, CSIR is conceptualized as firms' actions that could negatively affect "an identifiable social stakeholder's legitimate claims" (Strike et al., 2006: 852). It, in turn, triggers negative reactions from stakeholders (Kölbel, Busch, & Jancso, 2017), even though it may improve firm performance in the short run by saving costs (Kim, Kim, & Qian, 2018). For example, a firm's CSIR leads to employees' moral emotions, such as anger, sympathy, and guilt, resulting in punitive behaviors towards the firms (Hericher & Bridoux, 2023).

Cybercrime is one of the most critical challenges firms face in the present business environment (Pearlson, Schwartz, Sposito, & Arbisman, 2022). According to the U.S. Department of Commerce[1], a cyberattack is defined as "any kind of malicious activity that attempts to collect, disrupt, deny, degrade, or destroy information system resources or the information itself." Reported cybercrimes have caused a $27.6 billion loss from 2018 to 2022.[2] For example, T-Mobile experienced a cyberattack in 2021, which resulted in a $500 million loss.[3] Given its importance, recent research has started to investigate the factors that expose firms to such crime. For instance, researchers find that a firm's size, value, and intangible assets positively lead to a firm's chance of being targeted (Kamiya, Kang, Kim, Milidonis, & Stulz,

10.5465/AMPROC.2025.80bp

2021), and firms' poor CSR performance and high status could enlarge the negative impact caused by cyberattacks (Han, Pollock, & Graffin, 2024; Kamiya et al., 2021).

We develop a framework for the impact of CSR and CSIR on cyberattacks, grounding our arguments in screening theory. Screening theory, as a mirror image of signaling theory, focuses on how outside stakeholders use screening devices to understand a firm's underlying quality when facing information asymmetry (Zhang, Shi, & Connelly, 2024). Screening devices are observable indicators related to unobservable qualities (Stiglitz, 1975). Cyberattacks are usually driven by instrumental and expressive motives (Nguyen, Kamada, & Ramakers, 2022; Willison & Warkentin, 2013). Instrumentally motivated attackers seek financial gains through data and information theft. Expressively motivated attackers, such as disgruntled current or former employees, primarily seek punishment in response to a perceived offense or disrespect.

Specifically, we argue that CSR, serving as screening devices for firms' discretionary resources, will increase the likelihood of firms becoming targets of instrumentally motivated cyberattacks. Hackers have incomplete information about firms' discretionary resources that determine firms' capability to pay[4], which directly affects their potential financial gains from the cyberattacks. Firms' CSR activities help instrumentally motivated hackers to screen target firms with more resources, given that firms' engagement in CSR indicates their financial well-being and discretionary resources (Seifert, Morris, & Bartkus, 2004). On the other hand, CSIR, serving as a screening device for lack of moral value, will increase the likelihood of firms becoming targets of expressively motivated cyberattacks. Firms with high levels of CSIR demonstrate a lack of moral value and generate potential harm to others, which leads to stakeholders' negative reactions (Shea & Hawn, 2019). Employees, in particular, have been found to react negatively to firm CSIR (Hericher & Bridoux, 2023). We argue that cyberattacks can be considered as one type of punishing behavior that employees use on firms' lack of moral standards.

Further, we identify the underlying channels of these two predictions by distinguishing external and internal attacks. While an external cyberattack is launched by individuals or organizations outside of the targeted firms, an internal cyberattack occurs when it is carried out by someone authorized to access the organization's network, systems, or data (D'Arcy, Adjerid, Angst, & Glavas, 2020; Kamiya et al., 2021). Specifically, we argue that external cybercriminals, who are rational and motivated by economic incentives (Hui, Kim, & Wang, 2017), may have higher instrumental motivations and would rely more on CSR as a screening device. In contrast, internal attackers are more likely to rely on CSIR as a screening device for the moral value of a firm.

To test the mechanisms, we identify two contingent factors that moderate the effects of CSR and CSIR on a firm's probability of being targeted, respectively. Stock return works as a signal to provide additional information about firms' resource richness (e.g., Qian, Lu, & Yu, 2019), which would weaken the positive impact of CSR as a screening device for external cyberattacks. On the other hand, internal cybercrimes reflect stronger expressive motivations of employees who identify more with the firm (Ashforth & Mael, 1989). Building on political ideology research (Chin, Hambrick, & Treviño, 2013), we argue that liberal employees will react even more negatively to CSIR to commit more cybercrimes. We test our hypotheses using cyberattacks targeting U.S. publicly listed firms between 2006 and 2019 and find results that largely support the hypotheses.

HYPOTHESES

10.5465/AMPROC.2025.80bp

Instrumentally motivated hackers aim for financial gains. To achieve this goal, they try to locate target firms with rich discretionary resources, which are relatively unobservable (Bourgeois, 1981; George, 2005). From the screening perspective, CSR helps reveal firms' discretionary resources (Bansal & DesJardine, 2014; DesJardine et al., 2021; Seifert et al., 2004; Slawinski & Bansal, 2015). Thus, CSR may work as an unintended signal for instrumentally motivated cybercriminals, leading to a higher probability of cyberattacks. Thus,

Hypothesis 1: A firm's level of CSR is positively related to its likelihood of becoming a cyberattack target.

Expressively motivated hackers seek target firms to vent their feelings toward them. After observing a firm's CSIR, individuals develop moral emotions towards the firm's lack of moral value. Such moral emotions experienced by expressively motivated hackers when observing a firm with a high level of CSIR can further lead to their punitive behaviors (Antonetti & Maklan, 2016). Thus,

Hypothesis 2: A firm's level of CSIR is positively related to its likelihood of becoming a cyberattack target.

External cyberattacks are initiated by hackers who are not affiliated with the firm, while internal hackers are firms' current or former employees. External hackers are mainly driven by instrumental motivations compared to internal hackers and face higher information asymmetry regarding a firm's discretionary resources than internal ones. Therefore, CSR, serving as a screening device, becomes more salient for external hackers who are primarily driven by instrumental motives and face information disadvantages about firms' resources. Thus,

Hypothesis 3: A firm's level of CSR is more positively related to its likelihood of becoming a target of an external cyberattack than that of an internal cyberattack.

Internal cyberattacks involve the theft or leakage of firms' data and information caused by current or previous employees. If employees observe irresponsible conduct in the firm that may harm stakeholders, their organization identification will decrease, which may eventually result in "sabotage, resistance, or aggressive behaviors aimed at the organization" (Vadera & Pratt, 2013: 177). Internal hackers who identify with the firm would experience even stronger emotions, which will strengthen their expressive motives and amplify punitive behaviors. Thus,

Hypothesis 4: A firm's level of CSIR is more positively related to its likelihood of becoming a target of an internal cyberattack than that of an external cyberattack.

In Hypothesis 3, we predict that the screening effect of CSR is salient to external hackers, which affects the potential return for a cyberattack. Stock return serves as a valuable indicator of a firm's resources. If attackers can acquire additional information about firms' resources through their stock return, their reliance on CSR as a screening device decreases accordingly. Thus,

Hypothesis 5: A firm's stock return weakens the positive relationship between its level of CSR and the likelihood of becoming an external cyberattack target.

10.5465/AMPROC.2025.80bp

Our Hypothesis 4 predicts that CSIR could trigger internal cyberattacks from employees because of a firm's lack of moral value. We consider employees' political ideology as a moderator. Liberal investors, as well as liberal CEOs, tend to value resource equality and fairness, emphasizing more social engagement (Chin et al., 2013; DesJardine, Shi, & Westphal, 2024). When a firm's employees are more liberal, they may have even stronger negative reactions to the firm's irresponsible activities. Thus,

Hypothesis 6: Employees' liberalism strengthens the positive relationship between a firm's level of CSIR and its likelihood of becoming an internal cyberattack target.

METHODS

Data and Sample

We tested our hypotheses using cyberattacks targeting U.S. public firms between 2006 and 2019. We identified data breaches using the database provided by the Privacy Rights Clearinghouse (PRC). We thus identified 155 malicious cyberattacks across 97 firm-years involving 78 publicly traded U.S. firms. Table 1 presents a distribution of the 97 cyberattacks occurring by industry (two-digit Standard Industrial Classification codes) and year. Excluding missing values, our final sample consists of 97 cyberattacks occurring of 3,877 firms with 29,385 observations.

Measures

Cyberattack is our dependent variable, which is set to 1 if a firm experienced at least one cyberattack in a given year and 0 otherwise. *External cyberattack* occurs when the attack is initiated by hackers outside of the targeted firm. It is measured as 1 if a firm experienced at least one hacking cyberattack in a given year and 0 otherwise. *Internal cyberattack* equals 1 if a firm experienced at least one insider fraud cyberattack in a given year and 0 otherwise.

Following previous research, we measured CSR based on six dimensions from the MSCI data: community, diversity, employee relations, environment, product, and human rights (Jia et al., 2020; Qian et al., 2023). We calculated *CSR strength* as the total number of strengths across the six dimensions to measure CSR. Similarly, CSIR is the *CSR concern,* measured by the total number of concerns across the six dimensions.

Moderators. Stock return was calculated as the change in share price from the beginning to the end of the fiscal year plus dividends, divided by the beginning share price (Harrison, Thurgood, Boivie, & Pfarrer, 2020).

To measure the employees' political ideology of a firm, we used all the political donations made by firm employees to either Republican or Democratic recipients (Gupta et al., 2017). We used those donation records to create four indicators of organizational liberalism following the approach used by Gupta et al. (2017). We then averaged the four indicators to calculate *liberalism* as a measure of employees' political ideology.

Control Variables. We controlled *Firm size*, *Firm age*, and a dummy variable that equals 1 if the firm is included in the *S&P 500* index. We also controlled for total *intangible assets*, *ROA, cash flow*, *leverage ratio*, and *Tobin's Q*. We have also included year- and industry-fixed effects to control the unobserved time and industry factors.

10.5465/AMPROC.2025.80bp

Estimation Strategy

We applied the probit random-effect model to estimate the hypotheses given that our dependent variable is the probability of a cyberattack[5]. We did not include firm fixed effects or use a conditional probit/logistic regression because they would lead to the drop of all firms that have not been attacked. Moreover, most of the attacked firms only experienced a cyberattack once in our sample, which leads to our dependent variables having limited variance within individuals. We clustered standard errors by firms.

RESULTS

Our findings reveal that firms with higher *CSR strength* have approximately double the likelihood (2.420% vs 1.163%) of experiencing cyberattacks compared to those with lower CSR strength, supporting Hypothesis 1. Similarly, firms with higher *CSR concerns* face a 95.33% increased likelihood of cyberattacks (2.424% vs 1.241%), supporting Hypothesis 2. While the study found no significant difference in how CSR strength affects internal versus external cyberattacks (rejecting Hypothesis 3), it did find that CSR concerns specifically increase the likelihood of internal cyberattacks but not external ones (supporting Hypothesis 4). Additionally, the research found that stock returns moderate the relationship between CSR strength and external cyberattacks, with the effect being weaker for firms with higher stock returns (supporting Hypothesis 5), but found no support for the hypothesized moderating effect of employee liberalism on the relationship between CSR concerns and internal cyberattacks (rejecting Hypothesis 6).

Robustness Analyses

Matched sample analysis. We matched each targeted firm to a non-targeted control firm in the same year (the year the targeted firm experienced the first cyberattack in our sample) and the same two-digit SIC code. The propensity score is calculated using the logit regression of *Cyberattack* on *S&P 500*, *firm age*, and *ROA*. The results are consistent with our main findings.

Inclusion of MSCI rating as an exogeneous shock. We follow Chatterji and Toffel (2010) and Cheng, Ioannou, and Serafeim (2014) and utilize the initiation of MSCI rating of a firm's CSR/CSIR activities as a sudden increase in information disclosure. The results are in line with our prediction.

CONTRIBUTIONS

We contribute to the CSR and CSIR literature by uncovering cybercriminals' reactions to firms' activities as an unintended consequence. Building on screening theory, we highlight CSR and CSIR as distinct screening devices affecting targeting likelihood. While prior research explored CSR/CSIR impacts on stakeholders like consumers and analysts, identifying effects on cybercriminals is crucial given their significant organizational and societal impact. This study joins recent discussions about CSR/CSIR costs (DesJardine, Marti, & Durand, 2021). CSR generally benefits firms, yet our study reveals hackers interpret it as indicating strong financial standing and resource availability. Ironically, this positive impression makes these firms more attractive targets as potential benefits increase, introducing additional costs to CSR activities.

We also contribute to cybersecurity literature by distinguishing external and internal hackers and examining how CSR/CSIR affects instrumental and expressive motivations. Information about firm resources attracts instrumentally motivated hackers, while moral values signal opportunities for expressively motivated hackers (Agarwal, Ghosh, Ruan, & Zhang, 2024; Yue, Wang, & Hui, 2019). We broaden the field by revealing how strategic issues shape cyberattack vulnerability, moving beyond traditional technological prevention focus. While previous research emphasized prevention methods like infrastructure investment and security policies (Bulgurcu, Cavusoglu, & Benbasat, 2010), our study demonstrates how firm-level strategies can position organizations as cyberattack targets.

Moreover, our findings have several practical implications. When making decisions about CSR or CSIR levels, practitioners should be aware of the additional risks brought by potential cyberattacks. The strategies for CSR/CSIR and cybersecurity should not be viewed in isolation but need to be considered together. When pushing firms to reveal more information about their CSR/CSIR practices, policymakers should be aware that this could potentially make these firms a target for cyberattacks.

ENDNOTES

1. https://csrc.nist.gov/glossary/term/cyber_attack (Accessed on July 31 2024).

2. https://www.ic3.gov/Media/PDF/AnnualReport/2022_IC3Report.pdf (Accessed on July 31 2024).

3. https://www.cnet.com/tech/mobile/t-mobile-agrees-to-500m-settlement-in-massive-data-breach/ (Accessed on July 31 2024).

4. Despite firms' reluctance to admit that they pay off hackers, studies have shown that victim's payment in cybercrime is adopted by a significant portion of firms. For example, research on the Bitcoin ecosystem finds that from 2013 to mid-2017, the market for payments to hackers had a minimum worth of USD 12,768,536 (Paquet-Clouston, Haslhofer, & Dupont, 2019). Another study interviewed 41 victims of cybercrimes and found that eight among them paid hackers, and 11 considered paying (Connolly & Borrion, 2022). https://www.wsj.com/articles/companies-remain-reluctant-to-admit-paying-off-hackers-e1688946 (Accessed on August 25 2024).

5. We replicated each probit regression model in our main analysis and found similar results.

REFERENCES AVAILABLE FROM THE AUTHORS