Scherbina, Anna; Schlusche, Bernd

Working Paper

The Effect of Malicious Cyber Activity on the US Corporate Sector

AEI Economics Working Paper, No. 2023-06

Provided in Cooperation with:
American Enterprise Institute (AEI), Washington, DC

Suggested Citation: Scherbina, Anna; Schlusche, Bernd (2023) : The Effect of Malicious Cyber Activity on the US Corporate Sector, AEI Economics Working Paper, No. 2023-06, American Enterprise Institute (AEI), Washington, DC

This Version is available at:
https://hdl.handle.net/10419/280667

The Effect of Malicious Cyber Activity on the US Corporate Sector

Anna Scherbina
American Enterprise Institute, Brandeis University
Bernd Schlusche
Federal Reserve Board

AEI Economics Working Paper 2023-06
Updated June 2023

The Effect of Malicious Cyber Activity on the U.S. Corporate Sector

Anna Scherbina* — *Brandeis University*

Bernd Schlusche† — *Federal Reserve Board*

First draft: December 12, 2022
This draft: June 7, 2023

ABSTRACT

We compile a comprehensive dataset of adverse cyber events experienced by U.S. firms. We then categorize cyber incidents by their detrimental impacts on firms' assets and operations and show that firms suffer significant value losses across multiple cyber categories. These losses also spill over to economically linked firms, thereby amplifying the negative effect of malicious cyber activity on the economy. We additionally assemble a lexicon to identify from public sources firms that possess trade secrets, work on emerging technology or critical infrastructure projects, or have government and defense contracts, and show that such firms face a higher risk of a cyber incident.

JEL classification: G10, G12, G14, G17

Keywords: Cyberattacks, Malicious Cyber Activity, Cyber Threat Actors, Attack Vectors, Intellectual Property Theft, Trade Secrets, Critical Infrastructures, Emerging Technologies, Defense Contracts, Government Contracts, Spillover Effects, Economically Linked Firms, Corporate Transparency

*Address: Brandeis International Business School, 415 South Street, Waltham, MA 02453. E-mail: ascherbina@brandeis.edu.

†Address: Board of Governors of the Federal Reserve System, 20th Street and Constitution Avenue NW, Washington, DC 20551. E-mail: bernd.schlusche@frb.gov.

The views in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System or any other person associated with the Federal Reserve System. All remaining errors are our own. We thank seminar participants at Air Force Cyber College, Bentley University, Amrita University and Brandeis University for helpful comments. We are grateful for Refinitiv for making its data available to us for this project.

The Effect of Malicious Cyber Activity on the U.S. Corporate Sector

ABSTRACT

We compile a comprehensive dataset of adverse cyber events experienced by U.S. firms. We then categorize cyber incidents by their detrimental impacts on firms' assets and operations and show that firms suffer significant value losses across multiple cyber categories. These losses also spill over to economically linked firms, thereby amplifying the negative effect of malicious cyber activity on the economy. We additionally assemble a lexicon to identify from public sources firms that possess trade secrets, work on emerging technology or critical infrastructure projects, or have government and defense contracts, and show that such firms face a higher risk of a cyber incident.

JEL classification: G10, G12, G14, G17

Keywords: Cyberattacks, Malicious Cyber Activity, Cyber Threat Actors, Attack Vectors, Intellectual Property Theft, Trade Secrets, Critical Infrastructures, Emerging Technologies, Defense Contracts, Government Contracts, Spillover Effects, Economically Linked Firms, Corporate Transparency

1. Introduction

As the U.S. economy grew more reliant on information technology, malicious cyber activity has emerged as a new threat to corporations, government, and private citizens. Illegal activities, such as theft, sabotage, espionage, and equipment tampering, now can be committed in cyber space, where attribution is difficult and physical proximity is not required.

In this paper, we analyze the effect of malicious cyber activity on publicly traded firms. We show that cyber threats take many forms, from the more widely studied theft of personally identifiable information (PII) of a firm's customers and employees to the physical compromise of industrial control systems. For the purposes of this analysis, we construct, to the best of our knowledge, the most comprehensive dataset of publicly reported cyber events for the January 1999—January 2022 period by combining cyber events reported in four different datasets: the Refinitiv News Analytics (Refinitiv) dataset; the Privacy Rights Clearinghouse (PRC) dataset; the VERIS Community Database (VCDB); and the Capital IQ Key Developments (CapIQ) dataset.

Firms face unequal cyber threats that are determined by the nature of their business, network connections, and assets in place, and a firm's cyber risk can be inferred from publicly available information. For example, firms that collect PII may be targeted by criminals with profit motives and nation-states with espionage motives. Since retail firms are more likely to collect PII data of their customers, we find that high advertising expenditures, which are indicative of retail firms, significantly predict future cyber incidents. Availability of cash would put a firm at risk for a ransomware attack. We confirm that large cash holdings increase the likelihood of future cyber events. We also find that high profitability and growth opportunities help predict future malicious cyber events, which suggests that cyber threat actors target successful firms, possibly for industrial espionage.

News stories about a firm complement the accounting data in providing insights into why a firm may be targeted by cyber threat actors. We consider several business risk categories that can be inferred from news data. One source of risk involves having business connections with high-level targets, such as the U.S. government or the defense sector, that may position

a firm as a launching point for so-called "supply chain attacks," with the ultimate goal to breach government agencies or steal government secrets and government employee PII data. Sensitive projects present another source of risk. Against the backdrop of the intense competition for global dominance in the emerging technologies, such as artificial intelligence and autonomous driving vehicles, industrial espionage conducted through cyber means allows to make quick advances and to reduce research and development costs. Additionally, firms working on so-called critical infrastructures that are integral to the smooth functioning of the economy and the well-being of the citizens, such as the power grid, nuclear power plants, water treatment facilities, dams, etc., may be preemptively breached by nefarious actors such as terrorists or hostile nation-states intent on causing maximum disruption at some future time. Finally, having trade secrets and R&D projects, which can be stolen by cyber means, puts a firm at risk for intellectual property (IP) theft.

Using a variety of public documents, we compile a list of emerging technologies at the heart of the global economic competition; of critical infrastructure sectors, disruptions of which would cause wide-ranging economic and national-security consequences; of terms related to intellectual property and trade secrets; and of terms related to government and defense contracts. We then use the "bag of words" approach to identify from past news stories in Refinitiv those firms that have worked on emerging technology or critical infrastructure projects, had trade secrets or developed intellectual property, or had government or defense contracts in the previous year. Regression analysis confirms that these business aspects help predict next-year's adverse cyber events.

Our results show that firms suffer statistically and economically significant negative abnormal returns in response to the announcements of adverse cyber events. Furthermore, we find that prices react more quickly in the later part of the sample, implying that market participants have learned over time about the damage that such events can cause.

Guided by both practitioner classifications and by the commonality in the cyber events we observe in the dataset, we develop a taxonomy of cyber events based on the harms they inflict on the business. Whenever possible, we additionally classify the events by the identity of the

cyber threat actor and by the cyber intrusion vector. While PII theft is the most widespread type of adverse cyber event reported in public data due to regulatory reporting requirements, it is not the most damaging to firms' value. By far more damaging are ransomware, malware, and distributed denial of service (DDoS) attacks, all designed to disable IT systems and to interfere with the public and private access to corporate services, data, and websites.

The overall economic cost of malicious cyber activity for the U.S. economy is substantially larger than the sum of the individual costs incurred by the directly affected firms. Prior literature has shown that firm-specific shocks have the potential to affect other economically linked firms (e.g., Cohen and Frazzini (2008), Scherbina and Schlusche (2016), and Scherbina and Schlusche (2020)). Adverse cyber shocks have been shown to negatively affect industry peers of the focal firm. Kamiya, Kang, Kim, Milidonis, and Stulz (2021) show that firms in the same industry as the firm that experienced a cyber event also exhibit significantly negative price reactions. Similarly, Jamilov, Rey, and Tahoun (2021) show that stock prices of firms in the same country and industry as the focal firm also react negatively to the mentions of cyber terms in the focal firm's earnings conference calls. Studying the spillover effects of one of the most damaging cyberattacks, NotPetya, released in 2017, Crosignani, Macchiavelli, and Silva (2020) show that the effects propagated through the supply chains of the directly affected firms, and that the value losses suffered by the customers of these firms totalled at least four times the amount of value lost by the directly hit firms.

We also document a significant negative spillover effect to economically linked firms. We employ the methodology of Scherbina and Schlusche (2016) to identify economically linked firms for each firm that experienced a cyber event: Economically linked firms are identified as firms co-mentioned in the news with the cyber-affected firm over a rolling one-year window preceding the event. We show that economically linked firms experience a value loss equal to 44% of that experienced by the directly affected firms. Adding to the results in the papers mentioned above, we find that the spillover effect can travel across industries and is stronger for linked firms that are smaller than the directly hit firm.

Our results also highlight the problem of underreporting of adverse cyber events. It has been frequently argued that, despite the requirement by the Securities and Exchange Commission (SEC) that firms report all material corporate events publicly, including cyber incidents, there is widespread evidence of underreporting. For example, Amir, Levi, and Livne (2018) use the VCDB data to identify cyber events that were voluntarily disclosed by firms versus events that were initially reported by sources outside the firm and find that firms do not report all cyber events, and, specifically, they are likely to withhold disclosure of the more damaging incidents. The study further finds that firms are more likely to disclose when investors already have a strong suspicion that such incidents had occurred. Consistently, we document that the probability that a cyber event is being reported is increasing with outside investor scrutiny, as proxied by analyst coverage, news coverage, and institutional holdings of the firms' equity. Therefore, when estimating the overall cost of malicious cyber activity to the U.S. economy from available data, in addition to negative spillover effects, the issue of underreporting needs to be taken into account.

This paper contributes to the growing literature on the economic implications of the relatively new risk of malicious cyber activity.[1] A number of papers in that literature conduct event studies to analyze the impact of cyber events on firm value. Generally, with the data samples in prior work being heavily weighted towards instances of PII theft, these studies do not find overly strong reactions to cyber events in the relatively short event window that is uncontaminated by other news (see, e.g., Kvochko and Pant (2015), Hilary, Segal, and Zhang (2016), Akey, Lewellen, Liskovich, and Schiller (2018), and Hogan, Olson, and Angelina (2020)). In contrast, Kamiya, Kang, Kim, Milidonis, and Stulz (2021) analyze 165 adverse cyber events that were obtained from the PRC database over the 2005–2007 period and that were also covered in the Factiva news database (indicating that these are prominent events), also requiring that the affected firms had no other major news during the event study period, and find a more negative price reaction compared to other studies. Amir, Levi, and Livne (2018) find even larger market reactions to cyber events that are first reported by outsiders

[1] See, e.g., Kashyap and Wetherilt (2019) for a discussion of the emerging cyber risk.

rather than the firms themselves. Furthermore, Jamilov, Rey, and Tahoun (2021), Florackis, Louca, Michaely, and Weber (2022), and Jiang, Khanna, Yang, and Zhou (2022) show that cyber risk is a priced risk factor. Jamilov, Rey, and Tahoun (2021) construct a cybersecurity risk factor based on firms' exposure to shocks in the economy-wide level of cyber risk, which is measured from the aggregate mentions of cybersecurity-related terms in quarterly earnings calls. They estimate that the low-minus-high cyber-exposure-beta portfolio earns an annual return of -3.3%. Florackis, Louca, Michaely, and Weber (2022) measure firm-level exposure to cyber risk from the language of the risk factors section of the annual 10-K filings and likewise find that cyber risk is priced.

We additionally contribute to two other strands of literature. In compiling a lexicon of various categories of malicious cyber activity, a lexicon of emerging technologies and critical infrastructures, as well as lists of keywords related to intellectual property and trade secrets, and of keywords related to government and defense contracts, we also contribute to the literature that applies linguistic approaches to finance (e.g., Tetlock (2011) and Loughran and McDonald (2011)). Finally, in showing negative spillovers of cyber incidents to economically linked firms, we contribute to the literature on the network effects in finance (e.g., Barrot and Sauvagnat (2016) and Carvalho, Nirei, Saito, and Tahbaz-Salehi (2020)).

The rest of the paper is organized as follows: Section 2. describes the data used in the paper. Section 3. presents the empirical results, and Section 4. concludes.

2. Data

2.1 Data sources

We obtain data on adverse cyber events experienced by U.S. publicly traded firms from four different datasets: (1) the Refinitiv News Analytics dataset, (2) the Privacy Rights Clearinghouse dataset, (3) the VERIS Community Database, and (4) the Capital IQ Key Developments dataset. Below, we describe the four datasets in detail and outline the steps to combine them.

The first dataset, Refinitiv, includes news about publicly traded corporations from all major news sources, and the firms mentioned in the news stories are tagged with stock tickers. In addition to the news headlines and the date and time of the news announcements, Refinitiv provides metadata for the news stories, such as a newness score of a story, a relevance score for each firm mentioned in the news, a sentiment score (positive, negative, neutral) for each firm mentioned and the degree of confidence associated with the sentiment score, the number of words in the story, the number of sentences, as well as other quantitative measures. Refinitiv additionally classifies news stories into various topic categories, and a news item may have multiple topic categories.

From the Refinitiv news dataset, we construct a sample of stories that contain reports of firms experiencing adverse cyber incidents. We start by limiting the dataset to the news stories that center on the firm being mentioned rather than mentioning the firm in passing. Therefore, we drop all general news stories, that is, those with news topic categories that include the coverage of market and industry trends, analyst reports, research roundups, trade order imbalances, etc., as well as news stories that mention more than 10 firms because in such stories cybersecurity, if mentioned, is unlikely to be the first report of an incident affecting a particular firm.

Next, we find stories about cybersecurity incidents in two ways. First, we select news with the Refinitiv topics "Computer Crime / Hacking / Cybercrime" or "Cybersecurity." Second, in order not to miss other news that erroneously may not have been tagged with these two topic codes, we add news stories that we obtain by running keyword searches for cybersecurity-related terms obtained from: (1) a cybersecurity glossary compiled by Global Knowledge,[2] (2) unambiguously cybersecurity-specific terms defined by the Cybersecurity and Infrastructure Security Agency,[3] and (3) names of malicious code (malware) used by cyber threat actors, including malware that targets industrial infrastructure.[4] Finally, in

[2]Available at `https://www.globalknowledge.com/us-en/topics/cybersecurity/glossary-of-terms/#gref`.

[3]Available at `https://niccs.cisa.gov/about-niccs/cybersecurity-glossary`.

[4]Malware types and names are obtained from the following websites: (1) `https://us.norton.com/internetsecurity-malware-types-of-malware.html`, (2) `https://thehackernews.com/search/label/SCADA%20Malware`, and (3) `https://www.controlglobal.com/blogs/unfettered/`.

order to narrow down the news stories to the first mentions of an adverse cyber event experienced by a firm rather than, say, reports about a firm improving security after an incident or engaging in a new cybersecurity initiative, we keep only stories with a negative sentiment (specifically, stories with a Refinitiv sentiment score of "-1," while also requiring that this score is estimated with at least 75% confidence), and only those news that are classified as being highly relevant to the company mentioned (specifically news with a Refinitiv relevance score of "1").

The second data source is the PRC dataset,[5] which mainly collects information on PII breaches. Presently, PII breaches is the most widely reported category of corporate cybersecurity incidents because of strict notification requirements. The first such requirement is the Health Insurance Portability and Accountability Act (HIPAA), which was enacted by Congress in 1996 and mandates businesses in the healthcare industry to notify individuals when their private health information was breached.[6] Additionally, by now, all 50 states require that businesses notify individuals whose personally identifiable information has been breached, and the first such state-level data breach notification law was enacted in California in 2002.[7] For each observation of a data breach, PRC provides the name of the company that suffered the breach and the date the breach was made public. For some observations PRC provides additional information on the breach, such as the total number of records that were breached, whether the breach involved credit card records, whether a company insider was involved, etc.

VCDB is our third data source. It is a community-maintained dataset of cyber incidents that collects information on adverse cyber events experienced by entities in the public and private sectors.[8] The observations are collected from a number of public sources, such as media reports, press releases, the Department of Health and Human Services, and publicly available legal documents from various states and regions. In addition to cyber incident descriptions and the name of the entity affected, for some records VCDB provides information on the cy-

[5]This dataset is available at `https://privacyrights.org/data-breaches`.
[6]`https://www.hhs.gov/hipaa/for-professionals/breach-notification/index.html`.
[7]`https://www.itgovernanceusa.com/data-breach-notification-laws`.
[8]This dataset is available at `http://veriscommunity.net/vcdb.html`.

ber threat actor (which could be internal, e.g, manager, system administrator, executive, or external, e.g., nation-states, terrorists, competitors, activist groups, criminal organizations), the actor's motivation (e.g., espionage, grudge, ideology), the type of cyber compromise (e.g., skimming, exfiltration, surveillance), and the victims (e.g., patients, employees, customers).

The fourth source of cyber data is CapIQ. This dataset contains material news and events that can affect a firm's value. We search this dataset using the same cyber keywords we use for the Refinitiv search to identify cyber-related news.

Using these four datasets, we form a comprehensive dataset of unique cyber incidents experienced by publicly traded firms. While Refinitiv includes stock ticker symbols and CapIQ contains CUSIPs, the two other datasets do not provide a stock identifier. We match company names provided by these datasets (the "Company" variable in PRC and the "Victim ID" variable in VCDB) to the CRSP company name (COMNAM) variable using a fuzzy match; we then check these matches by hand, in order to link the PRC and VCDB records to the corresponding CRSP permno. In order to avoid having multiple records of the same cyber events reported across multiple databases, we enact the following procedure. If we encounter news about the same firm closely spaced in time (within five days of each other), we keep records from only one source using the following order of preference: (1) Refinitiv, (2) CapIQ, (3) VCDB, and (4) PRC.

The resulting dataset covers the January 1999—January 2022 period and is described in Table 1. As shown in Panel A, the dataset contains 1,843 unique observations of adverse cyber events,[9] and the firms are relatively large, with an average NYSE size decile of 7.95. Most of the sample (almost 60%) comes from Refinitiv; PRC is the second most important data source, providing 25% of all observations; and VCDB is the third most important data source, contributing 15% of observations to our final sample. Just under 3% of the data come from CapIQ, and this dataset covers slightly larger firms than the other three datasets, with an average NYSE size decile of 8.27.

[9]The number of unique companies affected by adverse cyber events is 790.

In addition to data on cyber events, we obtain balance sheet and income statement variables from the Merged CRSP/Compustat Database. Stock returns, the number of shares outstanding and prices are obtained from CRSP. Analyst coverage is measured as the number of analysts issuing annual earnings forecasts for the current fiscal year and is obtained from I/B/E/S. Finally, institutional ownership and the number of institutions holding shares of a firm are obtained from the Thomson Reuters Institutional (13f) Holdings dataset.

2.2 Categorizing adverse cyber events

2.2.1 Categorizing cyber compromise

Malicious cyber activity is a relatively new threat, and the government, industry and academia are still working on classifying cyber events into distinct categories. Some classifications are based on the attack vector (e.g., phishing, corporate insider, supply chain attack) while others are based on the motivation of the cyber threat actor (e.g., financially motivated, ideologically motivated, motivated by revenge, industrial espionage).[10] We propose categorizing cyber events based on the resulting harmful effect on the firm, as described below (the categories are presented in the order of prevalence in our data):

1. **PII breach**: Theft of personally identifiable information by cyber means.

2. **Security breach:** Unauthorized access to a corporate network, server, device or other IT asset. A security breach may result in many adverse consequences, such as data exfiltration, data deletion and corruption, and expanding the reach to other targets. Cyber incidents are classified into this category when no specific mention of a data breach or other known theft or damage is reported in the security breach news.

3. **Skimming/theft of funds:** Detection of skimming devices on card reading equipment and possible theft of funds by cyber means.

4. **Overt cyberattack:** Overt malware attack on the corporate IT infrastructure that adversely impacts the normal operations of a firm. In contrast to covert data and security breaches, the cyber threat actor does not make an effort to hide the intrusion and the resulting negative consequences are easily observed in real time.

5. **Ransomware attack:** Attack by a ransomware malicious code that renders a firm's data and files inaccessible by encrypting them until a ransom is paid.

[10] See, for example, Shevchenko, Jang, Malavasi, Peters, Sofronov, and Trück (2021) for a discussion of various proposed cyber taxonomies.

6. **Cyber lawsuit:** A lawsuit filed against a firm by corporate customers or partners negatively affected by a cyber incident or a cyber vulnerability. News about an adverse cyber event or a cyber vulnerability are often widely reported for the first time when a lawsuit is filed.

7. **IP theft:** Theft of corporate intellectual property by cyber means.

8. **Security flaw:** Discovery of a security flaw that exposes a firm to a network intrusion or data theft and is typically reported by cybersecurity researchers or competing firms. Such news do not indicate that the vulnerability was already exploited, only that it is now in the public domain and may be exploited in the future.

9. **Malware infection:** Discovery of malicious software on the corporate IT network. A malware infection can be premeditated or accidental, and the consequences may range from a system slowdown to data theft and equipment destruction. Ransomware is a type of malware, and an incident is classified in this category when no mention of ransomware or a ransom demand is made.

10.**DDoS (Distributed Denial-of-Service) attack:** malicious action to disrupt access to connected online devices and websites.

11.**Control systems compromise:** Compromise of software and equipment responsible for controlling industrial processes.

While cyber intrusion categories 1, 2, 3, 7 and 11 are covert and equivalent to theft or sabotage, categories 4 and 10 are overt attacks.[11] Unless the information on cyber categories are available from the PRC or VCDB datasets, all classifications are achieved with keyword searches. Sample headlines for each of these categories are provided in Appendix Table A1.

Panel B of Table 1 shows the breakdown of cyber events in our dataset by category, as well as by data source. Given that firms are widely mandated to report PII breaches, this is the most prevalent cyber event category, containing almost 40% of all observations, and roughly 55% of these observations are contributed by the PII-breach-dominated PRC dataset. The "Unclassified" group contains cyber incidents that do not fall into any of the above categories either because it contains an uncommon type of cyber compromise or, alternatively, not enough detail is provided in the headline to classify the incident. The unclassified events make up 12% of the sample, and almost all observations in that category, 97%, are obtained

[11] A destructive cyberattack on Sony Pictures Entertainment (SPE), later attributed to North Korea, is an example of an overt attack (see. e.g., `https://www.fbi.gov/news/press-releases/update-on-sony-investigation`).

from Refinitiv.[12] Control systems compromise is the least prevalent category in our dataset, however, incidents in this category are severely underreported and appear to be more prevalent among private firms (see, e.g., Weiss, Stephens, and Miller (2022)).

2.2.2 Identifying cyber threat actors and attack vectors

The severity of a cyber compromise may depend on the identity of the cyber threat actor and the vector of the attack. Nation-state-sponsored actors are thought to be the most technically sophisticated. While in most cases, attribution is difficult or impossible, for some records, the VCDB dataset provides information on the cyber actors who perpetrated the incident. The three categories are: "internal actors,", "external actors," and "partner actors," with further subtypes. For example, for external actors, the subtypes are "nation-state," "state-sponsored or affiliated group," "terrorist," "competitor," "activist," etc. Similarly, for some records, the PRC dataset reports whether the breach is attributed to a corporate insider. Based on this information, we are able to classify some cyber actors as "corporate insiders" or "nation-states," with other categories being significantly less prevalent. For the observations in our final dataset that are lacking this information from the VCDB and PRC datasets, whenever possible, we classify cyber threat actors into "nation-state" or "corporate insider" categories based on keyword searches of the news headlines. The U.S. government has made efforts to identify nation-state-sponsored cyber actors, and we use these identifications for our keyword searches.[13] Additionally, major cybersecurity companies use their own unique names for the so-called advanced persistent threat (APT) cyber actors, which are most likely nation-states, and we use these names in our keyword searches as well.[14] Appendix Table A2 presents examples of headlines corresponding to "nation-state" and "corporate insider"

[12] Consider two examples of cyber headlines in the "Unclassified" category, both obtained from Refinitiv: (1) "(NAV) announced Monday that it has learned of a credible potential cybersecurity threat to its information technology system on May 20." and (2) "Brief-Essex Property Trust reports cyber-incident on computer networks."

[13] See, e.g. `https://crsreports.congress.gov/product/pdf/IF/IF11718`.

[14] See, e.g., `https://www.crowdstrike.com/blog/meet-crowdstrikes-adversary-of-the-month-for-june-mustang-panda/`.

threat actors. Panel C of Table 1 shows that 3.15% of incidents in our dataset are attributed to nation-states and only few incidents (0.11%) are attributed to corporate insiders.

Whenever possible, we also classify attack vectors into "internal" and "supply chain attack" vectors. We search news headlines as well as cyber event descriptions in the VCDB and PRC datasets for relevant keywords. Additionally, if the threat actor was classified as "partner actors" in the VCDB dataset, this indicates that the attack vector was a "supply chain" attack. Given that an external actor can arrange an internal attack vector, we are careful not to assume that an internal attack vector corresponds to a "corporate insider" cyber threat actor. However, whenever a perpetrator is classified as a corporate insider, we assume that the attack vector is internal. Appendix Table A3 presents examples of headlines corresponding to "supply chain" and "internal" attack vectors. Panel D of Table 1 shows that for 4.99% of incidents, an internal attack vector was identified (this is significantly higher than the percentage of corporate insider threat actors since this analysis is largely based on keyword searches). Only few incidents (0.22%) are attributed to supply chain attacks.

The remainder of Table 1 reports additional summary statistics. Panel E presents the distribution of cyber incidents by industry (using the 12 industry classifications from Kenneth French's website) and shows that most incidents occurred in the "Business Equipment" and "Money" industries. Cyber event probabilities in the last column are obtained by scaling the number of incidents by the number of firms in the industry, and it shows that the probability of a firm experiencing a cyber incident is the highest in the "Business Equipment," "Shops," and "Telecom" sectors. Panel F shows that the number of cyber events increases over time but declines slightly after 2018. This is likely a feature of our dataset since the PRC dataset stops in 2020 and VCDB has lower coverage after 2018 and no coverage after 2020.

2.3 Identifying firms' business risks for a cyber compromise

We use the Refinitiv news data to identify the aspects of a firm's business that may increase its cyber risk. We focus on three such risk categories. First, we hypothesise that R&D activities and the possession of intellectual property and trade secrets would raise the risk

that a firm is targeted for IP theft and industrial espionage. The second source of risk is working on government and, more narrowly, defense contracts. The government ties would make a firm a potential infiltration target from which to launch supply chain attacks on government agencies.[15] Alternatively, government contractors may themselves be targeted for the theft of intellectual property connected to government interests and of employee PII data for espionage purposes. The third source of risk is related to the nature of a firm's projects. Work on critical infrastructures, the disruption of which would lead to large-scale negative consequences, would raise the risk of a cyber compromise by terrorists, ransom-seeking criminals, and hostile nation-state actors.[16] And developing new technologies at the heart of the international rivalry for technological superiority would raise the risk of cyber espionage.

We identify these three aspects of a firm's business through the corresponding Refinitiv news topic codes,[17] which we supplement with our own keyword searches. For the search for government contractors, the keywords we use contain full names and abbreviations for all U.S. government agencies, combined with the linguistic equivalents of the word "contract." (We exclude news about a firm losing or failing to obtain a government contract.) We assume that the category "defense contract" is a subset of the "government contract" category, and for finding news about defense contracts, we use the full names and abbreviations of all military, intelligence, and national security government agencies. For the keywords pertaining to critical infrastructures, we use the list of the critical infrastructure sectors from the Department of Homeland Security (DHS).[18] Finally, to identify highly competitive emerging technologies we use the U.S. Government Critical and Emerging Technologies List[19] and the list compiled by Graham, Klyman, Barbesino, and Yen (2021).

[15] See, e.g., Ghadge, Weiß, Caldwell, and Wilding (2020) for a discussion of cyber risks in supply chains.

[16] See, e.g., Haber and Zarsky (2018) and Krauss (2018) for a discussion of cyber vulnerability of the critical infrastructure and industrial control systems.

[17] Consider a few examples of the Refinitiv topic codes that correspond to these three business risks, such as topic codes for "intellectual property," "contract wins," "defense," "military procurement," "power stations," "water utilities," "semiconductors," "autonomous vehicles," "machine learning," and "quantum computing."

[18] Available at `https://www.cisa.gov/critical-infrastructure-sectors`.

[19] `https://www.whitehouse.gov/wp-content/uploads/2022/02/02-2022-Critical-and-Emerging-Technologies-List-Update.pdf`.

Appendix Table A4 provides examples of news stories indicating (1) firm's work on a government contract; (2) more narrowly, a defense contract; (3) work on critical infrastructure or emerging technology projects; and (4) possession of trade secrets or intellectual property. The indicator variable for each business risk category described above is set to 1 if a firm has at least one news story in the appropriate category over the prior 12 months and 0 otherwise.

Panel G of Table 1 compares the characteristics of firms that have experienced at least one adverse cyber event in a given year to those that have not. Variable definitions are provided in Appendix Table A5. Since we rely on Refinitiv news searches to identify business risks, the sample for this table includes only firms that were covered by Refinitiv in a given year.

The table shows that cyber-affected firms are different in many respects from unaffected firms. Given that larger firms are a more lucrative target than smaller firms and that firms try to withhold information about cyber incidents from the public, it is not surprising that variables associated with size and a more transparent information environment, as captured by analyst coverage, the number of Refinitiv news stories, the number of funds that hold the company stock, and institutional ownership, are all significantly higher for the sample of firms that experienced an adverse cyber event in a particular year relative to those that did not. Additionally, cyber-affected firms have significantly higher asset intangibility (though a lower R&D/Assets ratio). These firms are also more profitable, as indicated by a significantly higher ROA, and have higher growth opportunities, as indicated by a higher Tobin's Q and a lower Book/Market ratio. Additionally, cyber-affected firms spend more on advertising, indicating that they are likely to be consumer-facing and, therefore, collect customer PII. Finally, as we hypothesised, cyber-affected firms are more likely to be government or defense contractors, possess trade secrets, and work on emerging-technology or critical infrastructure projects. However, the incremental contributions of each of these characteristics on cyber risk need to be assessed in a multivariate regression setting.

3. Empirical Results

3.1 Predicting adverse cyber events

To assess which firm attributes increase its risk of a cyber incident, we estimate a linear probability model (LPM) predicting at least one cyber incident occurring at any time during the year as a function of firm characteristics that are publicly known at the end of the previous year. Specifically, we estimate the following model:

$$P(Y_{i,t} = 1|X_{i,t-1}) = \beta X_{i,t-1} + \Theta_t + \epsilon_{it} \tag{1}$$

where $Y_{i,t}$ equals 1 if a firm experiences at least one adverse cyber event in year t and 0 otherwise, and $X_{i,t-1}$ are explanatory variables available for each firm i as of December 31 of year $t-1$. (For balance sheet data, we assume that it takes three months after the firm's fiscal year end for the information to become public.) The model is estimated with year fixed effects, Θ_t,[20] and standard errors are clustered by firm and year.

We use a number of predictive variables. All else equal, we expect larger firms to make more lucrative targets since they possess more PII of their customers and employees as well as more intellectual property and trade secrets due to their larger scale. Therefore, firm size is one of the explanatory variables. Additionally, given the widespread reluctance of firm management to report cyber incidents (see, e.g., Amir, Levi, and Livne (2018)) the attention of large sophisticated market participants may increase the likelihood that a cyber incident is uncovered and publicly reported. Hence, we use additional predictors that proxy for investor attention: analyst coverage, institutional ownership, the number of investment funds that hold a firm's stock, and news coverage.

Compared to business-facing firms, customer-facing firms make more attractive targets for PII theft as they collect personal information of their customers. Additionally, they would

[20]We use the LPM specification because, given the use firm fixed effects in some model specifications, a Probit or Logit model specification will produce inconsistent estimates of β since the number of firm fixed effects to be estimated increases with the number of firms, N, while the number of time periods, T, is fixed (see, e.g., Wooldridge (2010) for a discussion of the incidental parameter problem). However, we checked that in specifications without firm fixed effects, Probit regressions produce results of a similar economic magnitude.

suffer larger losses from DDoS attacks if they sell their products and services online. We use advertising intensity as a proxy for customer-facing firms.

Additionally, we include three variables intended to capture whether a firm owns intellectual property that can be stolen through cyber means—R&D intensity, asset intangibility, and an indicator variable for whether a firm had news about trade secrets or intellectual property in the past year. We also use the indicator variables described in the previous section for whether a firm is a government contractor, defense contractor, or works on critical infrastructure or emerging technology projects.

Having cash on hand makes a firm an attractive target for ransomware attacks, and we, therefore, add cash holdings as another explanatory variable. Highly profitable firms and firms with good growth potential may be targeted for industrial espionage, and we include several variables intended to capture profitability and growth opportunities (ROA, Book/Market and Tobin's Q). In model specification (1), we include industry dummies for the Durable Goods, Finance, and Retail industries. These three industries may be more vulnerable to cyber threats due to the prevalence of valuable IP in the "Durable Goods" industry, easy access to cash in the Finance industry, and a wealth of PII data in both "Finance" and the "Retail" industries. In all other specifications, industry dummies are included, unless firm dummies is included instead.

Given that many cyber news stories are obtained from Refinitiv, and the business risk indicators rely on the Refinitiv stories as well, a regression that includes the full sample of firms may show a mechanical positive relation between the cyber event probability and the indicator variables derived from the Refinitiv dataset. In order to avoid this problem, we limit the sample of firms to those that are covered by Refinitiv in years t and $t-1$. Details of the explanatory variable definitions and calculations are provided in Appendix Table A5, and the sample covers the January 1999—December 2022 period.

Regression results and descriptive tables are presented in Table 2. Regressions in Panel A of the table predict an occurrence of any type of cyber event in the following year; Panel B reports sample statistics; Panel C presents statistics on cyber event probabilities and the

business risk dummies discussed above, by industry; and Panel D reports a correlation table for the explanatory variables.

The regression results show that the indicator variables for working on government and defense contracts, being involved in critical infrastructure or emerging technology projects, and having trade secrets and intellectual property all significantly increase the odds that a firm will experience an adverse cyber event next year. This result is consistent with our hypothesis that these aspects of a firm's business increase its cyber risk.

The estimated economic magnitudes of these business-related risks are large. When it comes to the effect of government contracts, the regression specifications with industry dummies (specifications (3) and (4)) indicate that having a government contract in the past year increases the probability of a cyber event in the next year by 2.1 percentage points. This represents an increase of 175% relative to the unconditional sample average probability of 1.2 percentage points per year (reported in Panel C). Specifications (7) and (8) show that having a defense contract increases the probability of a cyber event by roughly the same magnitude, that is, 2 percentage points. The economically large estimates on the government and defense contract indicators are particularly notable since they should be muted by the cybersecurity requirements that the government imposes on its contractors.[21] Furthermore, the specifications with industry dummies show that working on critical infrastructure and emerging technology projects increases the likelihood of an adverse cyber event in the following year by 0.5–0.6 percentage points (or by 42%–50% relative to the unconditional sample average probability), while having trade secrets increases this probability by 0.7 percentage points (or by 58% relative to the sample average).

Specifications with firm fixed effects (models (5), (6), (9) and (10)) show slightly lower but mostly still highly significant regression coefficients. Relative to not having a government contract, having one increases a firm's probability of a cyber compromise in the following year by 1.3 percentage points (which represents an increase of 108% relative to the sample

[21] See, e.g., the May 2021 Executive Order on Improving the Nation's Cybersecurity for recent initiatives on government contractor cybersecurity requirements (`https://www.whitehouse.gov/briefing-room/presidential-actions/2021/05/12/executive-order-on-improving-the-nations-cybersecurity/`).

average). For defense contracts, the estimated coefficient is slightly higher, 1.5 percentage points, amounting to a 125%-increase relative to the sample average. Similarly, when a firm obtains a critical infrastructure or an emerging technology project, the probability of a cyber event in the following year goes up by 0.4 percentage points (a one-third increase). Finally, the coefficients on the trade secrets and IP indicator variables are positive but statistically insignificant.

When it comes to other firm characteristics, the table shows that firm size, as well as firm age, are significant positive predictors of a cyber compromise. Profitability and growth opportunities, captured by ROA and Tobin's Q, are also significantly positively associated with future cyber incidents. Considering the effect of specific industries, model (1) shows that, after controlling for other firm characteristics, the "Retail" industry classification is a significant predictor of a future cyber incident, likely because retail firms are more prone to PII theft and point-of-sale skimming,[22] types of compromise that are prevalent in our dataset.

Specifications with industry fixed effects show significantly positive coefficients on advertising expenditures, asset intangibility, cash holdings, as well as proxies for investor attention—analyst coverage, news coverage, and the number of investment funds that hold the stock; the latter result suggests that investor attention increases corporate transparency when it comes to cyber incident reporting. While the coefficients on institutional ownership are negative in the regressions without firm fixed effects, they turn positive in specifications with firm fixed effects. In the cross-section of firms, concentrated institutional ownership may improve firm governance (e.g., Gillan and Starks (2002) and Chung and Zhang (2011)) and, consequently, cybersecurity investments, thereby decreasing the risk of cyber incidents; however, at the firm level, in the year following increases in institutional ownership, the rate of cyber incident reporting may increase due to higher investor scrutiny. The coefficient on the indicator variable of whether a firm experienced a cyber incident in the previous year is significantly positive in the regression specifications with industry dummies but is insignifi-

[22]See, e.g., `https://www.fbi.gov/how-we-can-help-you/safety-resources/scams-and-safety/common-scams-and-crimes/skimming`.

cant in the firm fixed effects specifications. This suggests that some firms have multiple cyber incidents in our dataset, however, at the firm level, a cyber incident in the current year does not increase the risk of a cyber incident in the following year, likely because closely-spaced in time cyber intrusions are not as profitable to the perpetrators. Firm fixed effects regressions further show that firms are more likely to experience a future cyber compromise following increases in R&D spending.

Our findings regarding firm characteristics that help predict adverse cyber events are consistent with those reported in Kamiya, Kang, Kim, Milidonis, and Stulz (2021). That paper analyzes instances of data breaches reported in PRC over the 2005–2017 period and also finds that the probability of a firm experiencing a data breach in the following year significantly increases in firm size, Tobin's Q, and asset intangibility.

We also consider separately two subsets of cyber incidents, namely only those attributed to nation-state actors and only those likely related to industrial espionage (specifically, cyber incidents classified as "IP theft" or "security breach"). The results are reported in Appendix Table A6, Panels A and B, respectively and are discussed in Appendix Section A1. The results show that the business risk indicators also help predict cyber events in these two subsets.

Panel C of Table 2 shows that the probability of a firm experiencing at least one cyber event in a given year. The panel shows that the risk of cyber incidents varies across industries, aligning with the results of Table 1, Panel E, which tabulates the *number* of cyber events by industry. Firms in the "Business Equipment," "Shops," and "Telecommunications" industries have probabilities of 2% or higher, while the average probability across all firms in the sample is only 1.2%.

Industries also differ in the business risks that would predispose them to cyber threats. A higher proportion of firms in the "Business Equipment" industry have government contracts and possess trade secrets and intellectual property. Firms in the "Utilities" sector are naturally most likely to work on critical infrastructure and emerging technology projects.[23]

[23]For the purpose of having within-industry variation, we do not automatically assume that all utilities firms work on critical infrastructure projects, instead, we put into this category all firms that work on water

Defense contracts are more prevalent among firms in the "Health" industry. Some sectors, such as "Shops," have notably low probabilities across all business risk indicators, though some may find it surprising that these probabilities are not closer to zero. A possible explanation is that firms across most industries have by now adopted emerging technologies, such as machine learning and computer vision. Moreover, even non-technological companies posses trade secrets and intellectual property.[24] And while defense contractors are predominately technology-oriented firms, the defense industry also contracts with low-tech firms (see, e.g., the October 10, 2018, headline: "US Foods wins $453 mln U.S. defense contract -Pentagon."). Finally, the last column of Panel E shows the industry distribution of firms in our sample, which is comprised of firms at the intersection of three datasets: CRSP, Compustat and Refinitiv. "Money" is the most prevalent industry, containing 27% of firms in our sample, while "Durables" is the least prevalent, with only 2% of firms in our sample belonging to this industry.

In sum, the results in this section show that firms face an unequal level of cyber threat. Firms with supply chain links to the government, and those that work on highly competitive new technologies, critical infrastructure projects, and have intangible assets and high growth opportunities face higher cyber threats.

3.2 Event studies

In this section, we document a negative price reaction to announcements of adverse cyber events that is not reversed in the near future. Cumulative abnormal returns (CARs) between one day before the event until τ days after the event are calculated as the sum of daily abnormal returns: $CAR_i[-1, \tau] = \Sigma_{t=-1}^{\tau} AR_{it}$, where the daily abnormal returns (AR) are calculated using the 1-, 3- and 4-factor models (Fama and French (1992) and Carhart (1997)). The results are presented in Table 3.

treatment or renewable, nuclear and hydroelectric energy and energy storage projects, which are inherently dangerous when compromised or which face intense international competition.

[24]Consider, for example, the July 28, 2015 headline, "The TJX Companies, Inc. to Acquire Off-Price Australian Retailer Trade Secret <TJX.N>."

Panel A presents CARs for different event windows, and different factor-model specifications. The left-hand-side of the table uses the entire sample, and the right-hand-side, only a more recent subperiod, January 2014–January 2022. CARs are significantly negative for all event windows and models used to compute abnormal returns, and there is evidence of underreaction as CARs become more negative in longer event windows. There is no evidence of return reversals. Moreover, the reaction to adverse cyber news is stronger and quicker in the later subperiod, which indicates that investors grew more aware over time about the negative consequences of cyber intrusions and interferences. The magnitudes we observe are economically large: Firms lose between 0.58% and 0.89% of value already in the three-day window around the announcement, and over 27 trading days ([-1,+25] event window), firms lose between 0.94% and 1.43% of value, depending on the abnormal return model and the sample period.

These magnitudes are comparable to price reactions to cyber news found in related papers. Kamiya, Kang, Kim, Milidonis, and Stulz (2021) study 165 announcements of major data breaches that appeared both in PRC and the Factiva news dataset and find an average four-factor-adjusted CAR of -0.79% in the [-1,+1] event window, and of -1.12% in the [-1,+5] window. Akey, Lewellen, Liskovich, and Schiller (2018) also use PRC and observe that the more records were breached the more negative CARs become, but CARs tend to be less negative for firms with high corporate social responsibility (CSR) scores, which suggests that investing in CSR helps cushion against the reputational damage of data breaches. Hogan, Olson, and Angelina (2020) obtain cyber events collected by Advisen Ltd., a data provider that collects these events from public sources. Their dataset covers the January 1990–April 2019 period, and the authors limit the sample to observations that affect more than 99 people. The dataset contains 3,992 observations that primarily include instances of PII theft (2/3 of all observations). They estimate an average CAR of -0.19% in the [-1,+1] event window and of -0.25% in the [-1,+5] event window (both statistically significant). Iyer, Simkins, and Wang (2020) also use the PRC dataset and study bond price reactions to the news of data breaches. They analyze the impact of 277 data breach events that occurred in the 2005–2016

time period for which bond trading data are available and find that bonds lose on average 1.30% of value in the one-month window after the news of a breach. Finally, Jamilov, Rey, and Tahoun (2021) analyze the effect of cyber terms being mentioned in firms' earnings calls and find that such mentions are also associated with negative abnormal returns during the week of the call, and each additional mention of a cyber keyword is found to reduce weekly returns by 0.22% (4.3 basis points per day over 5 trading days).

Panel B shows CARs over a 3-day event window [-1,+1] by NYSE market capitalization breakpoints, and the results show that the negative effect is stronger for smaller firms: CARs for firms in the smallest four NYSE size deciles range from −1.39% to −1.29% over the entire sample and from −2.25% to −2.15% over the January 2014—January 2022 subperiod. CARs become less negative as stocks increase in size, such that for the largest NYSE size decile CARs are between −0.30% and −0.32% for the entire sample period and −0.36% for the later subperiod. The likely explanation for the negative relation between value loss and firm size is that smaller firms are more concentrated on a single line of business and that business could be severely damaged by a cyber event, while larger firms tend to be more diversified across projects and lines of business.

Table C presents price reactions by cyber event category, starting from unclassified events and then continuing in order from the most to the least prevalent category in the data. The table shows that DDoS and non-ransom malware incidents, which result in a loss of value of over 2% over a 3-day window, are the most damaging events. Ransomware attacks, with a 4-factor CAR of −1.39%, come third in terms of the resulting value loss, and cyberattacks and security breaches come fourth and fifth, respectively, with 4-factor CARs of −1.15% and −0.58%, respectively. The price reactions to PII theft is not as large in magnitude but nonetheless statistically significant. It is not surprising that prices do not react as strongly to PII theft as to other types of cyber compromise. Absent large regulatory penalties and customer boycotts in response to an inadequate protection of personal information, breached PII is largely an externality for a firm. The average price reaction to cyber events that could not be classified is highly statistically significant, with an average CAR of −1.51%.

Finally, the price reactions to cyber incidents classified as Skimming/theft of funds, Cyber lawsuit, IP theft, and Security flaw are statistically insignificant. It may be surprising that the value loss in response to IP theft is insignificant since a small single-product firm can loss its entire livelihood when its IP is stolen (consider the effect of IP theft from American Superconductor Corporation).[25] Since IP theft is nonstandard news, it may take time to process its implications.

The information processing speed may be quicker for firms with a more intensive news coverage. We check if this is the case by rerunning the results in Panels A and C using the Refinitiv cyber news sample only. Cybersecurity breaches that are reported in Refinitiv also generate more investor attention and are likely to be larger in scale. The results, reported in Appendix Table A7, show that the Refinitiv-only sample of cyber events is indeed accompanied with more immediate and more negative CARs, which are estimated to be up to -1.8% in the [-1,+25] event window in the January 2014—January 2022 sample period. As for cyber news categories, three most negative price reactions are estimated for the cyber incidents classified in the "DDos" (-3.6%), "Cyber lawsuit" (-2.1%), and "Malware (non-ransom)" (-2.1%) categories. The reaction to the news of PII theft is estimated to be -1.0%, and it ranks eighth out of 11 total cyber categories in terms of severity, behind (1) DDoS, (2) cyber lawsuit, (3) non-ransom malware, (4) theft of funds by cyber means, (5) unclassified, (6) ransomware, and (7) overt cyberattack categories.

Panel D presents price reactions to cyber compromises by industry and shows that the largest value loss is experienced by firms in the "Chemicals," "Business Equipment," and "Other" industries. Price reactions are also significantly negative in the "Money" industry. Firms in these industries are more reliant on information technology and thus have a wider attack surface and also possess intangible assets, such as intellectual property and PII data, leading to larger losses in the event of a cyber compromise.

Finally, Panel E analyzes the determinants of the magnitude of cyber-induced value losses. Specifically, we regress 4-factor CARs computed over a [-1,+1] event window on character-

[25] `https://www.justice.gov/opa/pr/chinese-company-sinovel-wind-group-convicted-theft-trade-secrets`.

istics of the cyber incident and the business risks of the firms discussed earlier. When it comes to business risks, the results show that the indicator for critical infrastructure and emerging-technology projects is a consistently positive predictor of the price reaction. The likely explanation is that the effects of IP theft experienced by firms that work on emerging technologies are hard to quantify immediately, and the market may underreact to the immediate news of IP theft.[26] Moreover, security breaches of critical infrastructure firms so far have not resulted in any large-scale damaging effects. The table furthermore shows that incidents attributed to internal attack vectors reduce the magnitude of the negative price effect, likely because such observations are primarily dominated by instances of small-scale PII theft obtained from the PRC dataset.

3.3 Negative spillover effects

In this section, we investigate whether adverse cyber events affect economically linked firms that were not directly targeted. We use the methodology of Scherbina and Schlusche (2016) to identify linked firms from the news data. Specifically, we use the Refinitiv news dataset to identify firms that are co-mentioned in the same news story as the firm that experienced a cyber event over a one-year rolling window that ends two days prior to the event. To ensure that the news stories truly focus on the firms being mentioned rather than other general topics, we require that a news story mentions exactly two firms and that both firms are mentioned with the highest relevance score of 1.[27] Furthermore, to make sure that the firms are not mentioned as competitors, we require that both firms have the same news sentiment score. In the main specification, we additionally require that the news stories from which linked firms are identified cannot be classified as cyber news, based on the news topics and keywords we use to identify cyber news. The reason for removing prior cybersecurity co-mentions is to avoid potential instances of common cyber vulnerabilities being explicitly

[26] This result is consistent with the average CAR for the IP theft category being insignificant in Panel C of the table.

[27] The types of linkages identified through the news data include partnerships, parent/subsidiary relations, customer/supplier relations, cross-investments, common customer, common exposure to legal and regulatory risks, reliance on common resources, infrastructure and production factors, the use of common technology, etc. (see Table A3 in Scherbina and Schlusche (2016) for a full list).

discussed in the news, thereby potentially creating a common cybersecurity sentiment that may explain the future return co-movement. Finally, we require that economically linked firms did not experience a cyber incident themselves during the event window.

The sample of linked stocks is described in Table 4, Panel A. The panel shows that for 1,448 cyber events (or 79% of the sample), we are able to identify economically linked stocks. In that sample, each focal firm (the firm that experienced the cyber event) has, on average, 13.84 linked firms, and the focal and linked firms were co-mentioned 2.51 times, on average, over the past year. Economically linked firms tend to be smaller than focal firms, likely because, as discussed earlier, larger firms represent a more attractive target and also have a larger attack surface. Finally, 40% of linked firms belong to the same industry as the focal firm based on the shared 12-industry classification we use in the paper.

With thus identified linked stocks, we study spillover effects emanating from adverse cyber news of the directly affected firm using the event study methodology described earlier. Panel B of Table 4 presents the event study results over event windows of various lengths that start on the day of the event, day 0.[28] As before, CARs are calculated using the 1-, 3-, and 4-factor models, with standard errors clustered by the month of the event. The left side of the table uses the entire January 1999–January 2022 sample period, and on the right side, the sample is restricted to the January 2014–January 2022 subperiod.

The results in Panel B show significant negative spillovers to economically linked firms; the linked-stock price reaction amounts to about 22%–36% of the magnitude of the price reaction of the focal firm (shown in Panel B of Table 3) over a 10-day event window, and this ratio increases to 41%–44% over the [0,+25] event window. Moreover, the comparison between the left and right sides of the table shows that in the later subperiod, linked stocks react to the cyber news of focal firms with a shorter delay: CARs start being significant already in the [0,+5] window in the later subperiod vs. only in the [0,+10] window for the entire sample period. This suggests that the market learned to process spillover effects from

[28]We do not start the event window on day −1 as we did with the event studies for the directly affected firms in order to allow for the news to be publicly known. The results are similar when we start the event window on day −1.

cyber shocks more quickly over time. The longer event windows show that the initial price reactions are not reversed but rather prices continue to decline over time, which suggests an initial underreaction rather than an overreaction to cyber compromises of linked firms.

In order to keep the event window relatively short and uncontaminated by other events but still allow sufficient time for prices of linked stocks to react, the subsequent panels of the table use a [0,+5] event window, unless noted otherwise. (Moreover, unless noted otherwise, subsequent panels use the January 2014–January 2022 subperiod since price reactions are not statistically significant over this event window in the entire sample.)

Panel C shows that the negative spillover effects to linked firms are robust to various alternative sample construction choices. The first specification does *not* exclude cyber news from the set of co-mentions used to identify linked firms. The results are very close in magnitude to the main-specification results. In the second specification, linked stocks are identified as in the main specification, but the rolling identification window is reduced from 12 to six months. CARs are similar in magnitude, but the statistical significance is somewhat weaker because of the smaller number of linked stocks identified this way. In the third specification, we identify linked stocks in the same way as in the main specification but exclude both linked and focal firms that make earnings announcements within the [0,+5] event window, and in the fourth specification, we remove both linked and focal stocks that had any news in the Refinitiv news dataset over the event window. Both specifications show CARs that are even more negative than those in the main specification; a possible explanation is that in that subsample investors are not distracted from the negative cyber news by other news that are, on average, of neutral sentiment. The last specification considers only linked stocks that belong to a different industry than the focal firm and also shows more negative CARs than the main specification. A possible reason is that linked firms identified in other industries are more likely to be linked in a way that is conducive to supply chain attacks, such as a partnership, parent/subsidiary relation, an M&A, or a supply chain link.[29]

[29]What types of linkages put firms at a higher risk for negative spillover effects is a topic for future research.

Panel D analyzes negative spillover effects by cyber event category and shows that linked stocks react significantly negatively to security breach news and cyber events in the unclassified category, earning CARs amounting to 36 and 19 percentage points, respectively, of the CARs earned by the focal firms in that event window and time period.[30] The CARs are largely negative but insignificant across other categories. It is not surprising that linked stock prices react strongly to the security breach news since unauthorized intrusions into corporate networks may be used for supply chain attacks on linked firms, for example by stealing log-in credentials or inserting malicious code into trusted software.[31]

Panel E presents 4-factor CARs by the industry of the linked firm. The table shows that adverse cyber events tend to have significant negative spillover effects on economically linked stocks belonging to the following five industries: "Business Equipment," "Energy," "Money," "Non-Durables," and "Telecommunications." Firms in these industries possess valuable IP and PII data that can be stolen via a supply-chain cyber intrusions. Moreover, these are technologically savvy industries, and the technological reliance may create concerns about shared vulnerabilities.

Finally, Panel F reports the results of regressing 4-factor CARs on dummy variables indicating characteristics of linked firms relative to the directly affected firm as well as additional controls. The indicator variable "Smaller" takes a value of "1" if the linked firm is in a lower NYSE decile than the focal firm and "0" otherwise. The "Same industry" indicator is set to "1" for linked firms that are in the same industry as the cyber-affected firm, using the Fama-French 12-industry classification, and "0" otherwise. We additionally include the business risk indicator variables used earlier, calculated separately for both the cyber-affected and linked firms. (These four indicator variables are set to one when a firm had a government contract, a defense contract, a project in critical infrastructure or emerging technologies, and IP or trade secret news over the past year, respectively, and are set to zero otherwise.)

[30] For reference, over this event window and time period, focal firms earn statistically significant four-factor CARs equal to -0.82% in the security breach category and -1.70% in the unclassified category.

[31] One example of a supply chain attack is the breach of Target with the credentials stolen from its HVAC vendor (`https://money.cnn.com/2014/02/06/technology/security/target-breach-hvac/index.html`).

Finally, when available, we include indicator variables for the characteristics of the cyber incidents, such as whether the cyber threat actor was a nation-state or whether an internal attack vector was used. In most specifications, linked firms that are smaller than the directly affected firm show larger negative reactions, likely because they are more dependent on the cyber-affected firm. All other explanatory variables are statistically insignificant.

The results presented in this section illustrate that malicious cyber activity inflicts a larger damage to the economy than the losses experienced by the directly affected firms, and that some categories of malicious cyber activity lead to larger negative spillovers than others. These results confirm previously reported findings that the costs of adverse cyber events can spread beyond the firms that experienced them. Analyzing the effects of the NotPetya cyberattack, which was a destructive malware attack perpetuated by a nation-state and which disrupted normal operations of the directly affected firms, Crosignani, Macchiavelli, and Silva (2020) find that the negative effects propagated downstream to customer firms, causing customers to report significantly lower profits. Florackis, Louca, Michaely, and Weber (2022) analyze the effects of the SolarWinds hack, a supply-chain attack on SolarWinds Corporation also perpetuated by a nation-state, in which the cyber threat actor compromised the company's Orion software through which it was able to gain access to the networks, systems and data of SolarWinds' customers. The authors identify the 38 most important SolarWinds customers and show that these firms experienced significantly negative CARs around the report of the hack. Kamiya, Kang, Kim, Milidonis, and Stulz (2021) use the PRC dataset comprised primarily of PII data breaches over the 2005—2017 time period and define industry peers as firms within the same four-digit SIC codes; they find that industry peers react significantly negatively to the news of an adverse cyber event experienced by a peer firm. The magnitude of this reaction amounts to, on average, 61% of the CAR of the cyber-affected firm. Instead of cyber incidents, Jamilov, Rey, and Tahoun (2021) consider the effect of cyber terms being mentioned in earnings calls of industry peers. They define industry peers as firms that are headquartered in the same country as the cyber-affected firm and that also operates in the same six-digit NAICS industry. They find that when a

cyber-affected firm has a cyber term mentioned during its earnings call, industry peers also react significantly negatively, even though they themselves did not experience earnings calls with cyber terms mentioned in the same week. The magnitude of the industry peer reaction amounts to 64% of the magnitude of the reaction of the cyber-affected firm.

Complementing these results, we show that economically linked firms of smaller size than the directly hit firm suffer greater negative consequences and that losses from malicious cyber activity also spill over across industries. The fact that the spillover effects cross industry boundaries suggests that these effects are not driven by investor sentiment but are rather explained by economic inter-dependencies and common cyber vulnerabilities. We additionally investigate various categories of adverse cyber events and show that some have more capacity than others to inflict widespread damage. In future research, it would be useful to identify the types of firm linkages that are most conducive to transmitting various types of cyber shocks through the firm networks.

4. Conclusion

In this paper, we investigate how malicious cyber activity affects publicly traded firms. We construct a comprehensive dataset of adverse cyber events by combining four different data sources: the Refinitiv News Analytics dataset, the Capital IQ Key Developments dataset, the VERIS Community Database, and the Privacy Rights Clearinghouse data breach dataset. One contribution of this paper is to introduce a taxonomy of adverse cyber events that are based on the damages inflicted on the firm, such as theft of personally identifiable information; security breaches that allow hackers access to corporate IT systems and devices; card skimming or financial theft; overt cyberattacks that impair the normal functioning of corporate IT systems; ransomware attacks that freeze access to IT systems or data until a ransom is paid; cyber lawsuits resulting from cyber vulnerabilities or losses incurred as a result of prior cyber intrusions; theft of intellectual property; discovery of a security flaw in corporate products or systems; non-ransom malware attacks; DDoS attacks that hinder

access to corporate servers and websites; and control systems compromise, with the latter being rare in our dataset.

We show that while PII theft is the most widely reported cyber event due to regulatory requirements, it is not the most damaging for firm value and also not the cyber compromise that causes the largest negative spillover effects. The largest firm value losses are associated with DDoS, ransomware attacks and overt cyberattacks, while security breaches results in the largest spillover effects. The four datasets that we use focus their collection efforts on different types of cyber categories. For example, the widely used PRC dataset is dominated by incidents of PII theft. In order to correctly assess the economic damage from malicious cyber activity, it is, therefore, important to use all data to construct a representative dataset of the types of cyber incidents suffered by the corporate sector.

Our second contribution is to identify firms' business risks that attract the attention of cyber threat actors. The risk of malicious cyber activity varies across firms and depends on the nature of a firm's assets, operations, and business connections. We find that firms that have government and defense contracts, and that engage in critical infrastructure and emerging technology projects, as well as firms that own intellectual property and trade secrets, face a higher risk of experiencing future cyber incidents. This risk also increases in the amount of cash on a firm's balance sheet, possibly due to a higher ability to make ransom payments, and in advertisement spending, likely because consumer-facing firms also collect PII information on their customers that can be stolen by cyber means. Finally, all else equal, higher scrutiny appears to increase the likelihood that a cyber incident is reported.

Our third contribution is to show that losses from cyber incidents spread beyond directly affected firms and spill over to economically linked firms. The magnitude of the spillover effect amounts up to 44% of the value loss experienced by the directly affected firms and is especially large when the linked firms are smaller than the directly hit firms. Moreover, spillover effects also affect economically linked firms that operate in a different industry than the targeted firm, indicating that they are likely explained by shared vulnerabilities and economic inter-dependencies rather than investor sentiment.

Our results suggest that when assessing the economy-wide impact of malicious cyber activity, it is important to take into account the negative network effects. The challenge for future work lies in quantifying the cost of malicious cyber activity to the U.S. economy after properly accounting for underreporting and the network effects of cyber compromises.

References

Akey, Pat, Stefan Lewellen, Inessa Liskovich, and Christopher M. Schiller, 2018, Hacking corporate reputations, *Corporate Law: Law & Finance eJournal*.

Amir, Eli, Shai Levi, and Tsafrir Livne, 2018, Do firms underreport information on cyber-attacks? evidence from capital markets, *Review of Accounting Studies* 23, 1177–1206.

Barrot, Jean-Noël, and Julien Sauvagnat, 2016, Input specificity and the propagation of idiosyncratic shocks in production networks, *Quarterly Journal of Economics* 131, 1543–1592.

Carhart, Mark M., 1997, On persistence in mutual fund performance, *Journal of Finance* 52, 57–82.

Carvalho, Vasco M., Makoto Nirei, Yukiko U. Saito, and Alireza Tahbaz-Salehi, 2020, Supply chain disruptions: Evidence from the great east japan earthquake, *Quarterly Journal of Economics* 136, 1255–1321.

Chung, Kee H., and Hao Zhang, 2011, Corporate governance and institutional ownership, *Journal of Financial and Quantitative Analysis* 46, 247–273.

Cohen, Lauren, and Andrea Frazzini, 2008, Economic links and predictable returns, *Journal of Finance* 63, 1977–2011.

Crosignani, Matteo, Marco Macchiavelli, and André Silva, 2020, Pirates without borders: The propagation of cyberattacks through firms' supply chains, *Journal of Financial Economics* 147, 432–448.

Fama, Eugene F., and Kenneth R. French, 1992, The cross-section of expected stock returns, *Journal of Finance* 46, 427–466.

Florackis, Chris, Christodoulos Louca, Roni Michaely, and Michael Weber, 2022, Cybersecurity Risk, *Review of Financial Studies* 36, 351—407.

Ghadge, Abhijeet, Maximilian Weiß, Nigel D. Caldwell, and Richard Wilding, 2020, Managing cyber risk in supply chains: a review and research agenda, *Supply Chain Management* 25, 223–240.

Gillan, Stuart L, and Laura T Starks, 2002, Institutional investors, corporate ownership and corporate governance, *United Nations University Discussion Paper*.

Graham, Allison, Kevin Klyman, Karina Barbesino, and Hugo Yen, 2021, *The Great Tech Rivalry: China vs the U.S.*Belfer Center for Science and International Affairs, Harvard Kennedy School.

Haber, Eldar, and Tal Zarsky, 2018, Cybersecurity for infrastructure: A critical analysis, *Florida State University Law Review* 44, 515–577.

Hilary, Gilles, Benjamin Segal, and May H. Zhang, 2016, Cyber-risk disclosure: Who cares?, Working paper.

Hogan, Karen M., Gerard T. Olson, and M. Rejoice Angelina, 2020, A comprehensive analysis of cyber data breaches and their resulting effects on shareholder wealth, *Accounting Technology & Information Systems eJournal*.

Iyer, Subramanian R., Betty J. Simkins, and Heng Wang, 2020, Cyberattacks and impact on bond valuation, *Finance Research Letters* 33, 101–215.

Jamilov, Rustam, Hélène Rey, and Ahmed Tahoun, 2021, The anatomy of cyber risk, Working paper.

Jiang, Hao, Naveen Khanna, Qian Yang, and Jiayu Zhou, 2022, The cyber risk premium, Working paper.

Kamiya, Shinichi, Jun-Koo Kang, Jungmin Kim, Andreas Milidonis, and Rene Stulz, 2021, Risk management, firm reputation, and the impact of successful cyberattacks on target firms, *Journal of Financial Economics* 139, 719–749.

Kashyap, Anil K., and Anne Wetherilt, 2019, Some principles for regulating cyber risk, *AEA Papers and Proceedings* 109, 482–87.

Krauss, Clifford, 2018, Cyberattack shows vulnerability of gas pipeline network, *New York Times*, April 4, 2018. Available at: `https://www.nytimes.com/2018/04/04/business/energy-environment/pipeline-cyberattack.html`.

Kvochko, Elena, and Rajiv Pant, 2015, Why data breaches don't hurt stock prices, *Harvard Business Review* March 31.

Loughran, Tim, and Bill McDonald, 2011, When is a liability not a liability? Textual analysis, dictionaries, and 10-ks, *Journal of Finance* 66, 35–65.

Scherbina, Anna, and Bernd Schlusche, 2016, Economic linkages inferred from news stories and the predictability of stock returns, Working paper.

———, 2020, Follow the leader: Using the stock market to uncover information flows between firms, *Review of Finance* 24, 189–225.

Shevchenko, Pavel V., Jiwook Jang, Matteo Malavasi, Gareth Peters, Georgy Sofronov, and Stefan Trück, 2021, Quantification of cyber risk – risk categories and business sectors, Working paper.

Tetlock, Paul C., 2011, All the news that's fit to reprint: Do investors react to stale information?, *Review of Financial Studies* 24, 1481–1512.

Weiss, Joseph, Rob Stephens, and Nadine Miller, 2022, Control system cyber incidents are real—and current prevention and mitigation strategies are not working, *Computer* 55, 128–137.

Wooldridge, Jeffrey M, 2010, *Econometric analysis of cross section and panel data* (MIT press).

Table 1
Sample description

This table presents descriptive statistics for the dataset of adverse cyber events used in this paper. The sample period is January 1999–January 2022.

Panel A: Cyber events by dataset

Dataset	Number of observations	% of all observations	Avg. NYSE size decile
Refinitiv	1,053	57.14%	7.96
PRC	464	25.18%	7.89
VCDB	274	14.87%	7.98
Capital IQ	52	2.82%	8.27
All data	1,843	100.00%	7.95

Panel B: Cyber event categories

Category	% of all data	Obtained from: Refinitiv	PRC	VCDB	CapIQ
PII breach	39.88%	26.12%	54.56%	17.41%	1.90%
Unclassified	12.05%	96.85%	0.90%	2.25%	0.00%
Security breach	11.01%	93.10%	1.97%	4.93%	0.00%
IP theft	9.39%	55.49%	26.59%	12.72%	5.20%
Overt cyberattack	8.84%	94.48%	0.61%	4.91%	0.00%
Skimming/theft of funds	6.62%	26.23%	7.38%	66.39%	0.00%
Ransomware	3.74%	94.20%	0.00%	4.35%	1.45%
Security flaw	3.53%	78.46%	1.54%	20.00%	0.00%
Cyber lawsuit	2.77%	45.10%	0.00%	0.00%	54.90%
Malware (non-ransom)	1.41%	100.0%	0.00%	0.00%	0.00%
DDoS attack	0.71%	69.23%	0.00%	30.77%	0.00%
Control syst. compromise	0.05%	100.0%	0.00%	0.00%	0.00%

Panel C: Distribution of cyber events by threat actor

Threat actor	Percentage
Undetermined	96.74%
Nation-state	3.15%
Corporate insider	0.11%

Panel D: Distribution of cyber events by attack vector

Attack vector	Percentage
Undetermined	94.79%
Internal	4.99%
Supply chain	0.22%

Panel E: Distribution of cyber events by industry

Industry	% of all firms	% of all cyber events	Cyber event probability
BusEq	12.28%	26.01%	2.33%
Chems	1.39%	0.72%	0.57%
Durbl	1.61%	2.05%	1.39%
Enrgy	3.24%	0.83%	0.28%
Hlth	6.97%	4.26%	0.67%
Manuf	6.08%	4.21%	0.76%
Money	37.52%	22.41%	0.66%
NoDur	3.47%	2.77%	0.88%
Other	16.46%	17.99%	1.20%
Shops	6.41%	13.23%	2.26%
Telcm	2.60%	4.76%	2.01%
Utils	1.97%	0.77%	0.43%

Panel F: Distribution of cyber events by year

Year	Percentage
1999	0.16%
2000	0.16%
2001	0.16%
2002	0.05%
2003	0.81%
2004	0.71%
2005	1.47%
2006	3.04%
2007	4.23%
2008	2.55%
2009	2.39%
2010	4.56%
2011	5.32%
2012	5.32%
2013	7.92%
2014	10.36%
2015	7.16%
2016	7.05%
2017	8.74%
2018	9.22%
2019	5.97%
2020	5.81%
2021	6.67%
2022	0.16%

Panel G: Firm characteristics as of year-end (only firms receiving Refinitiv news coverage in a given year)

	Mean for firms in past 12 mo.:			
	Cyber-affected	Unaffected	Difference	t-value
Analyst coverage	14.692	4.897	9.796	(27.11)
News coverage	629.817	125.961	503.857	(16.53)
News sentiment	0.165	0.173	-0.008	(-1.09)
Number of funds	549.632	122.854	426.778	(23.54)
Inst. ownership	0.621	0.425	0.196	(8.32)
Adv. Exp./Assets	0.013	0.008	0.005	(5.71)
R&D Exp./Assets	0.028	0.048	-0.020	(-10.45)
Intangibility	0.819	0.741	0.078	(12.14)
Cash/Assets	0.179	0.187	-0.008	(-1.53)
Firm size ($, mil.)	168,448.04	12,215.31	156,232.73	(10.98)
Firm age	24.899	15.585	9.314	(16.30)
ROA	0.076	-0.077	0.152	(22.41)
Book/Market	0.479	0.549	-0.071	(-4.56)
Tobin's Q	2.084	1.913	0.171	(3.35)
In prior 12 months, news of:				
Govt. contracts	0.088	0.018	0.069	(7.99)
Defense contracts	0.061	0.012	0.049	(6.66)
Trade secrets	0.152	0.030	0.121	(10.99)
Critical infr. & emerging tech. projects	0.267	0.080	0.188	(13.78)

Table 2
Predictors of adverse cyber events

This table presents the results of a linear probability model regressions explaining next year's adverse cyber events, which are coded as an indicator variable taking the value of 1 if a firm experienced such cyber event during the next calendar year and 0 otherwise. The explanatory variables are described in the Appendix, and all accounting variables are considered to be publicly known three months after the fiscal year end. The sample period is January 1999–January 2022. Standard errors are clustered by firm and year, and t-statistics are reported in parentheses.

Panel A: Predicting adverse cyber events

Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Govt. contracts	.	.	0.021	0.021	0.013	0.013	.	.	.	.
	.	.	(5.23)	(5.30)	(3.82)	(3.82)	.	.	.	.
Defense contracts	.	.	.	.	.	.	0.020	0.020	0.015	0.015
	.	.	.	.	.	.	(3.09)	(3.13)	(2.48)	(2.48)
Critical infr. &	.	.	0.005	0.005	0.004	0.004	0.006	0.006	0.004	0.004
emerging tech. proj.	.	.	(3.30)	(3.30)	(1.75)	(1.76)	(3.30)	(3.30)	(1.75)	(1.76)
Trade secrets	.	.	0.007	0.007	0.002	0.002	0.007	0.007	0.002	0.002
	.	.	(2.42)	(2.49)	(0.95)	(0.95)	(2.49)	(2.56)	(0.98)	(0.98)
Cyber_{t-1}	0.194	0.190	0.188	0.188	0.008	0.008	0.188	0.188	0.008	0.008
	(9.48)	(9.42)	(9.47)	(9.45)	(0.46)	(0.46)	(9.46)	(9.45)	(0.46)	(0.46)
An. Coverage	0.002	0.001	0.001	0.002	-0.001	-0.001	0.001	0.002	-0.001	-0.001
	(1.86)	(0.89)	(0.96)	(2.50)	(-0.40)	(-0.55)	(0.90)	(2.47)	(-0.41)	(-0.55)
News coverage	0.007	0.006	0.005	0.005	-0.000	-0.000	0.005	0.005	-0.000	-0.000
	(4.53)	(4.20)	(3.56)	(3.40)	(-0.13)	(-0.16)	(3.63)	(3.48)	(-0.04)	(-0.07)
No. funds	0.002	0.002	0.002	.	-0.000	.	0.002	.	-0.000	.
	(2.29)	(2.62)	(2.56)	.	(-0.45)	.	(2.61)	.	(-0.44)	.
Inst. ownership	-0.012	-0.013	-0.012	-0.009	0.010	0.010	-0.013	-0.009	0.010	0.010
	(-3.86)	(-3.89)	(-3.76)	(-4.02)	(5.23)	(5.57)	(-3.83)	(-4.09)	(5.24)	(5.58)
Adv./Assets	.	0.042	0.055	0.056	-0.012	-0.013	0.053	0.053	-0.012	-0.012
	.	(1.98)	(2.52)	(2.53)	(-0.26)	(-0.28)	(2.41)	(2.43)	(-0.25)	(-0.27)
R&D/Assets	.	0.001	-0.002	-0.003	0.021	0.020	-0.002	-0.004	0.021	0.020
	.	(0.13)	(-0.48)	(-0.75)	(5.19)	(5.27)	(-0.57)	(-0.85)	(5.17)	(5.24)
Intangibility	.	0.006	0.005	.	0.006	.	0.005	.	0.006	.
	.	(3.00)	(2.21)	.	(0.95)	.	(2.34)	.	(0.96)	.
Cash/Assets	.	0.013	0.013	0.014	-0.000	0.001	0.013	0.014	-0.000	0.001
	.	(3.67)	(3.68)	(4.07)	(-0.04)	(0.27)	(3.64)	(4.05)	(-0.04)	(0.28)
Firm size	0.004	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005
	(8.55)	(8.84)	(8.70)	(8.71)	(3.45)	(3.43)	(8.67)	(8.68)	(3.46)	(3.44)
Firm age	0.001	0.001	0.001	0.001	.	.	0.001	0.002	.	.
	(1.29)	(2.76)	(2.38)	(2.93)	.	.	(2.46)	(3.03)	.	.
ROA	0.002	0.003	0.003	0.003	0.013	0.013	0.003	0.003	0.013	0.013
	(1.91)	(2.67)	(2.48)	(2.10)	(8.13)	(8.16)	(2.51)	(2.12)	(8.12)	(8.16)
Book/Market	.	0.000	0.001	0.000	0.001	0.001	0.001	0.000	0.001	0.001
	.	(0.64)	(0.71)	(0.58)	(1.80)	(1.67)	(0.72)	(0.58)	(1.84)	(1.71)
Tobins Q	0.002	0.002	0.002	0.002	0.001	0.001	0.002	0.002	0.001	0.001
	(6.99)	(5.42)	(5.49)	(5.95)	(2.46)	(2.48)	(5.49)	(5.95)	(2.47)	(2.48)
Dur. goods ind.	0.002	.	.	.	.	.	.	.	.	.
	(1.35)	.	.	.	.	.	.	.	.	.
Financial ind.	0.002	.	.	.	.	.	.	.	.	.
	(1.03)	.	.	.	.	.	.	.	.	.
Retail industry	0.011	.	.	.	.	.	.	.	.	.
	(3.02)	.	.	.	.	.	.	.	.	.
Year dummy	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Ind. dummy	N	Y	Y	Y	N	N	Y	Y	N	N
Firm dummy	N	N	N	N	Y	Y	N	N	Y	Y
Obs.	81473	81473	81473	81473	81473	81473	81473	81473	81473	81473
Adj. RSq.	0.065	0.069	0.071	0.070	0.007	0.007	0.070	0.070	0.007	0.007

Panel B: Explanatory variables

	Mean	Median	25th pct.	75th pct.	Std. dev.
Analyst coverage	5.007	2.000	0.000	7.000	6.760
News coverage	131.638	76.000	30.000	149.000	242.849
Number of funds	127.663	62.000	9.000	159.000	207.968
Inst. ownership	0.427	0.381	0.032	0.766	0.386
Adv. Exp./Assets	0.008	0.000	0.000	0.002	0.025
R&D Exp./Assets	0.048	0.000	0.000	0.032	0.116
Intangibility	0.742	0.870	0.589	0.973	0.290
Cash/Assets	0.187	0.083	0.024	0.257	0.236
Firm size ($, mil.)	14,065.162	777.325	169.519	3,377.779	103,813.50
Firm age	15.690	12.000	5.000	22.000	14.154
ROA	-0.075	0.068	-0.017	0.137	0.415
Book/Market	0.548	0.467	0.223	0.793	0.684
Tobins Q	1.914	1.325	1.023	2.094	1.695

Panel C: Cyber events and business risks, by industry

	Annual	Business risks				
Industry	cyber event probability	Govt. contracts	Defense contracts	Critical infr. & emerging tech. proj.	Trade secrets	% of all obs.
BusEq	0.020	0.294	0.129	0.039	0.049	16.02%
Chems	0.004	0.147	0.097	0.023	0.009	2.08%
Durbl	0.011	0.110	0.082	0.016	0.025	1.93%
Enrgy	0.004	0.078	0.033	0.026	0.007	4.47%
Hlth	0.006	0.219	0.205	0.020	0.011	12.15%
Manuf	0.008	0.174	0.056	0.021	0.048	7.66%
Money	0.012	0.051	0.016	0.009	0.004	26.79%
NoDur	0.008	0.044	0.047	0.007	0.004	3.69%
Other	0.011	0.101	0.030	0.030	0.025	12.34%
Shops	0.020	0.055	0.025	0.012	0.009	7.17%
Telcm	0.022	0.241	0.101	0.036	0.023	3.12%
Utils	0.005	0.291	0.016	0.040	0.013	2.57%
All sample	0.012	0.142	0.070	0.022	0.020	100.00%

Panel D: Correlation table

	Govt. contr.	Mil. contr.	CI or NT proj.	Trade secrets	Past cyber events	Analyst coverage	No. news	No. funds	Inst. own.	Adv. exp.	R&D. exp.	Intang.	Cash hold.	Firm size	Firm age	ROA	B/M	Tobin's Q
Govt. contr.	1.000	0.482	0.205	0.093	0.088	0.074	0.178	0.058	0.044	-0.022	0.015	0.027	-0.001	0.072	0.082	0.047	-0.011	0.003
		(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(0.8770)	(<.0001)	(<.0001)	(<.0001)	(0.0011)	(0.4310)
Mil. contr.	0.482	1.000	0.153	0.073	0.060	0.059	0.152	0.049	0.042	-0.030	-0.011	0.025	-0.023	0.076	0.068	0.039	0.004	-0.012
	(<.0001)		(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(0.0011)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(0.1897)	(0.0002)
CI or NT proj.	0.205	0.153	1.000	0.198	0.087	0.166	0.339	0.116	0.092	-0.028	0.158	0.072	0.152	0.102	0.054	0.072	-0.045	0.089
	(<.0001)	(<.0001)		(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
Trade secrets	0.093	0.073	0.198	1.000	0.078	0.086	0.212	0.058	0.032	-0.002	0.209	0.084	0.178	0.029	0.013	-0.011	-0.046	0.107
	(<.0001)	(<.0001)	(<.0001)		(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(0.4974)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(0.0001)	(0.0008)	(<.0001)	(<.0001)
Past cyber	0.088	0.060	0.087	0.078	1.000	0.109	0.159	0.093	0.057	0.022	-0.016	0.031	-0.003	0.151	0.055	0.041	-0.009	0.009
	(<.0001)	(<.0001)	(<.0001)	(<.0001)		(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(0.4095)	(<.0001)	(<.0001)	(<.0001)	(0.0061)	(0.0036)
An. cov.	0.074	0.059	0.166	0.086	0.109	1.000	0.420	0.787	0.688	0.062	-0.049	0.063	-0.017	0.481	0.133	0.318	-0.051	0.131
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)		(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
No. news	0.178	0.152	0.339	0.212	0.159	0.420	1.000	0.344	0.309	0.063	0.004	0.027	-0.005	0.478	0.207	0.158	0.005	0.034
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)		(<.0001)	(<.0001)	(<.0001)	(0.2463)	(<.0001)	(0.1061)	(<.0001)	(<.0001)	(<.0001)	(0.1639)	(<.0001)
No. funds	0.058	0.049	0.116	0.058	0.093	0.787	0.344	1.000	0.759	0.029	-0.109	0.050	-0.097	0.439	0.240	0.204	-0.062	0.066
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)		(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
Inst. own.	0.044	0.042	0.092	0.032	0.057	0.688	0.309	0.759	1.000	0.037	-0.064	0.121	-0.035	0.280	0.185	0.204	-0.036	0.067
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)		(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
Adv. exp.	-0.022	-0.030	-0.028	-0.002	0.022	0.062	0.063	0.029	0.037	1.000	-0.024	0.052	0.047	-0.064	-0.006	0.119	-0.033	0.114
	(<.0001)	(<.0001)	(<.0001)	(0.4974)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)		(<.0001)	(<.0001)	(<.0001)	(<.0001)	(0.0666)	(<.0001)	(<.0001)	(<.0001)
R&D exp.	0.015	-0.011	0.158	0.209	-0.016	-0.049	0.004	-0.109	-0.064	-0.024	1.000	0.216	0.614	-0.395	-0.167	-0.270	-0.120	0.395
	(<.0001)	(0.0011)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(0.2463)	(<.0001)	(<.0001)	(<.0001)		(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
Intang.	0.027	0.025	0.072	0.084	0.031	0.063	0.027	0.050	0.121	0.052	0.216	1.000	0.349	-0.120	-0.077	0.185	0.213	0.189
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)		(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
Cash hold.	-0.001	-0.023	0.152	0.178	-0.003	-0.017	-0.005	-0.097	-0.035	0.047	0.614	0.349	1.000	-0.435	-0.253	-0.057	-0.074	0.464
	(0.8770)	(<.0001)	(<.0001)	(<.0001)	(0.4095)	(<.0001)	(0.1061)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)		(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
Firm size	0.072	0.076	0.102	0.029	0.151	0.481	0.478	0.439	0.280	-0.064	-0.395	-0.120	-0.435	1.000	0.289	0.205	0.034	-0.281
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)		(<.0001)	(<.0001)	(<.0001)	(<.0001)
Firm age	0.082	0.068	0.054	0.013	0.055	0.133	0.207	0.240	0.185	-0.006	-0.167	-0.077	-0.253	0.289	1.000	0.122	0.016	-0.144
	(<.0001)	(<.0001)	(<.0001)	(0.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(0.0666)	(<.0001)	(<.0001)	(<.0001)	(<.0001)		(<.0001)	(<.0001)	(<.0001)
ROA	0.047	0.039	0.072	-0.011	0.041	0.318	0.158	0.204	0.204	0.119	-0.270	0.185	-0.057	0.205	0.122	1.000	0.292	0.062
	(<.0001)	(<.0001)	(<.0001)	(0.0008)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)		(<.0001)	(<.0001)
B/M	-0.011	0.004	-0.045	-0.046	-0.009	-0.051	0.005	-0.062	-0.036	-0.033	-0.120	0.213	-0.074	0.034	0.016	0.292	1.000	-0.307
	(0.0011)	(0.1897)	(<.0001)	(<.0001)	(0.0061)	(<.0001)	(0.1639)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)		(<.0001)
Tobin's Q	0.003	-0.012	0.089	0.107	0.009	0.131	0.034	0.066	0.067	0.114	0.395	0.189	0.464	-0.281	-0.144	0.062	-0.307	1.000
	(0.4310)	(0.0002)	(<.0001)	(<.0001)	(0.0036)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	

Table 3
Stock price reaction to cyber events

This table presents results for stock price reactions to cyber incident announcements. Cumulative abnormal returns (CARs), reported in %, are based on three abnormal return specifications: (1) market-adjusted, (2) adjusted for the Fama-French three factors, and (3) four-factor adjusted (Fama-French three factors and the Carhart momentum factor), as specified above each column. Unless specified otherwise, the sample period is January 1999–January 2022. Standard errors are clustered by firm and year, and t-statistics are reported in parentheses.

Panel A: Abnormal returns calculated over different event windows

	CAR (%)					
Event	Entire sample			Jan 2014–Jan 2022		
window	Mkt-adj.	FF-adj.	4-factor-adj.	Mkt-adj.	FF-adj.	4-factor-adj.
$[-1, 0]$	-0.412	-0.413	-0.424	-0.589	-0.597	-0.628
	(-4.25)	(-4.43)	(-4.45)	(-4.94)	(-5.30)	(-5.47)
$[-1, +1]$	-0.610	-0.581	-0.584	-0.886	-0.856	-0.875
	(-4.94)	(-4.97)	(-4.86)	(-5.48)	(-5.68)	(-5.66)
$[-1, +5]$	-0.603	-0.620	-0.598	-0.885	-0.959	-0.961
	(-3.90)	(-4.06)	(-3.89)	(-4.68)	(-5.12)	(-5.14)
$[-1, +10]$	-0.757	-0.822	-0.785	-1.016	-1.120	-1.135
	(-3.59)	(-4.04)	(-3.83)	(-3.73)	(-4.30)	(-4.37)
$[-1, +15]$	-0.927	-0.953	-0.909	-1.288	-1.419	-1.403
	(-3.70)	(-3.97)	(-3.76)	(-3.82)	(-4.40)	(-4.32)
$[-1, +20]$	-0.937	-0.966	-0.890	-1.324	-1.426	-1.406
	(-3.10)	(-3.40)	(-3.10)	(-3.08)	(-3.55)	(-3.51)
$[-1, +25]$	-1.007	-1.034	-0.943	-1.302	-1.425	-1.393
	(-3.10)	(-3.36)	(-3.04)	(-2.85)	(-3.32)	(-3.24)

In the panels that follow, the event window is [-1,+1]

Panel B: By market capitalization

	CAR (%)					
NYSE size	Entire sample			Jan 2014–Jan 2022		
decile	Mkt-adj.	FF-adj.	4-factor-adj.	Mkt-adj.	FF-adj.	4-factor-adj.
1 through 4	-1.286	-1.330	-1.390	-2.148	-2.181	-2.252
	(-2.51)	(-2.54)	(-2.54)	(-3.20)	(-3.18)	(-3.18)
5 through 7	-0.503	-0.415	-0.341	-1.038	-0.914	-0.887
	(-1.37)	(-1.19)	(-1.19)	(-2.33)	(-2.18)	(-2.18)
8	-0.945	-0.958	-0.931	-1.000	-1.042	-1.079
	(-2.44)	(-2.52)	(-2.52)	(-2.33)	(-2.57)	(-2.57)
9	-0.743	-0.648	-0.707	-0.934	-0.740	-0.818
	(-3.25)	(-3.05)	(-3.05)	(-3.02)	(-2.64)	(-2.64)
10	-0.324	-0.297	-0.302	-0.358	-0.364	-0.361
	(-3.21)	(-3.32)	(-3.32)	(-2.67)	(-3.03)	(-3.03)

Panel C: By cyber category

Cyber category	4-factor-adj. CAR (%)
Unclassified	-1.512[a]
	(-3.29)
PII breach	-0.258[c]
	(-1.76)
Security breach	-0.583[b]
	(-2.04)
Skimming/theft of funds	-0.328
	(-0.75)
Overt cyberattack	-1.146[c]
	(-1.78)
Ransomware	-1.388[b]
	(-2.52)
Cyber lawsuit	-0.618
	(-0.71)
IP theft	0.013
	(0.04)
Security flaw	0.101
	(0.30)
Malware (non-ransom)	-2.170[b]
	(-2.06)
DDoS attack	-2.827[a]
	(-2.83)

Panel D: By industry

Industry	4-factor-adj. CAR (%)
BusEq	-0.847^{a}
	(-4.48)
Chems	-1.797^{b}
	(-2.25)
Durbl	0.804
	(1.77)
Enrgy	-0.272
	(-0.70)
Hlth	0.694
	(0.88)
Manuf	0.107
	(0.25)
Money	-0.320^{b}
	(-2.15)
NoDur	-0.490
	(-1.22)
Other	-1.499^{a}
	(-3.13)
Shops	-0.273
	(-1.26)
Telcm	-0.305
	(-0.88)
Utils	0.141
	(0.19)

Panel E: Regressions of 4-factor CARs on business risks and characteristics of cyber incidents

	Model								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Intercept	-0.636^{a}	-0.616^{a}	-0.789^{a}	-0.581^{a}	-0.762^{a}	-0.754^{a}	-0.660^{a}	-0.838^{a}	-0.831^{a}
	(-4.74)	(-4.79)	(-4.57)	(-3.95)	(-4.31)	(-4.30)	(-5.15)	(-4.60)	(-4.59)
Govt. contracts	0.340^{c}				0.205			0.203	
	(1.82)				(1.11)			(1.09)	
Defense contracts		0.320				0.154			0.171
		(1.37)				(0.64)			(0.72)
CI and emerging			0.504^{b}		0.568^{b}	0.582^{b}		0.578^{b}	0.589^{b}
tech. projects			(2.44)		(2.37)	(2.40)		(2.42)	(2.45)
IP and trade				-0.013	-0.293	-0.284		-0.296	-0.290
secrets				(-0.05)	(-1.02)	(-0.99)		(-1.04)	(-1.02)
Nation-state							0.325	0.213	0.214
actor							(0.75)	(0.51)	(0.52)
Internal attack							1.293^{a}	1.322^{a}	1.326^{a}
vector							(2.64)	(2.73)	(2.74)
Adj. R^2	0.001	0.000	0.002	0.000	0.003	0.003	0.003	0.006	0.006

Table 4
Market reaction of economically linked stocks

This table presents the analysis of the market reaction of economically linked stocks to the cyber incident news of focal firms. In the main specification, economically linked stocks are identified for each firm that has experienced an adverse cyber event as stocks that were co-mentioned in the same news story and with the same sentiment as the focal firm over the prior 12-month period that ends two days before the cyber news. The news stories used for identifying linked stocks must mention exactly two firms and, unless specified otherwise, stories about cyber-related topics are excluded. We require that the linked firms did not themselves experience a cyber incident starting 1 trading day before the cyber event and ending at the end of the event window. Cumulative abnormal returns are calculated based on 1-, 3-, and 4-factor models. The sample period is January 1999–January 2022, unless specified otherwise. Standard errors are clustered by firm and year, and t-statistics are reported in parentheses.

Panel A: Descriptive statistics on the sample of linked stocks

Number of cyber events	1,448
Avg. number of linked stocks per focal firm	13.84
Avg. number of past co-mentions	2.51
Avg. mkt. cap. of linked stocks ($, mil.)	65,932
Avg. mkt. cap. of focal stocks ($, mil.)	111,725
Fraction of linked firms that are, relative to the cyber-affected firm:	
- In a smaller NYSE size decile:	42.98%
- In a larger NYSE size decile:	24.59%
- In the same industry:	40.29%

Panel B: Cumulative abnormal returns calculated over various event windows

	CAR (%)					
Event	Entire sample			Jan 2014 –Jan 2022		
window	Mkt-adj.	FF-adj.	4-factor-adj.	Mkt-adj.	FF-adj.	4-factor-adj.
$[0, +1]$	-0.010	-0.015	-0.004	-0.066	-0.055	-0.063
	(-0.29)	(-0.50)	(-0.14)	(-1.43)	(-1.36)	(-1.60)
$[0, +5]$	-0.102	-0.084	-0.068	-0.224	-0.189	-0.200
	(-1.92)	(-1.83)	(-1.40)	(-3.37)	(-3.34)	(-3.62)
$[0, +10]$	-0.225	-0.224	-0.171	-0.464	-0.408	-0.405
	(-2.98)	(-3.37)	(-2.38)	(-4.80)	(-4.67)	(-4.74)
$[0, +15]$	-0.332	-0.339	-0.284	-0.531	-0.507	-0.491
	(-3.66)	(-4.10)	(-3.39)	(-4.65)	(-4.71)	(-4.76)
$[0, +20]$	-0.430	-0.456	-0.391	-0.610	-0.594	-0.587
	(-3.86)	(-4.48)	(-3.82)	(-4.20)	(-4.43)	(-4.48)
$[0, +25]$	-0.480	-0.488	-0.418	-0.644	-0.581	-0.568
	(-3.81)	(-4.15)	(-3.69)	(-3.89)	(-3.76)	(-3.92)

In the panels that follow, the event window is [0,+5]

Panel C: Alternative specifications and robustness checks

CAR (%)					
Entire sample			Jan 2014–Jan 2022		
Mkt-adj.	FF-adj.	4-factor-adj.	Mkt-adj.	FF-adj.	4-factor-adj.
Do not remove cyber stories when identifying linked firms					
-0.098	-0.076	-0.060	-0.215	-0.177	-0.187
(-1.86)	(-1.66)	(-1.23)	(-3.26)	(-3.10)	(-3.37)
Identify linked firms over rolling 6-month period					
-0.096	-0.066	-0.051	-0.225	-0.163	-0.171
(-1.42)	(-1.07)	(-0.81)	(-2.80)	(-2.30)	(-2.44)
Remove focal or linked firms with earnings announcements in [0,+5] window					
-0.110	-0.097	-0.073	-0.222	-0.197	-0.202
(-1.96)	(-2.03)	(-1.44)	(-3.18)	(-3.40)	(-3.58)
Remove linked firms with any news in [0,+5] window					
-0.107	-0.090	-0.074	-0.232	-0.197	-0.208
(-2.05)	(-1.96)	(-1.52)	(-3.57)	(-3.51)	(-3.80)
Linked stocks limited to stocks in different industry than focal firm					
-0.114	-0.101	-0.099	-0.263	-0.228	-0.234
(-1.73)	(-1.71)	(-1.69)	(-3.05)	(-3.10)	(-3.26)

In the panels that follow, the time period is Jan 2014–Jan 2022

Panel D: By cyber category of the focal firm

Cyber category	4-factor-adj. CAR(%)
Unclassified	-0.320[a]
	(-3.14)
PII theft	-0.124
	(-1.28)
Security breach	-0.290[b]
	(-2.17)
Skimming/theft of funds	0.076
	(0.22)
Overt cyberattack	-0.088
	(-0.61)
Ransomware	-0.069
	(-0.32)
Cyber lawsuit	-0.458
	(-1.47)
IP theft	-0.306
	(-1.61)
Security flaw	-0.275
	(-1.58)
Malware (non-ransom)	0.146
	(0.31)
DDoS attack	-0.840
	(-1.29)

Panel E: By industry of linked firms

Industry	4-factor-adjusted CAR(%)
BusEq	-0.148[c]
	(-1.85)
Chems	0.281
	(0.44)
Durbl	0.190
	(0.81)
Enrgy	-0.621[b]
	(-2.00)
Hlth	0.174
	(0.73)
Manuf	-0.239
	(-1.37)
Money	-0.190[b]
	(-2.30)
NoDur	-0.421[b]
	(-1.98)
Other	-0.272
	(-1.55)
Shops	-0.245
	(-1.24)
Telcm	-0.271[c]
	(-1.85)
Utils	-0.605
	(-1.31)

Panel F: Regression of 4-factor CARs on the linked and focal stock features and characteristics of cyber incidents

	Model								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Intercept	-0.105	-0.234	-0.135	-0.206	-0.214	-0.174	-0.185	-0.192	-0.203
	(-1.71)	(-3.26)	(-1.76)	(-1.75)	(-1.81)	(-1.56)	(-1.65)	(-1.76)	(-1.85)
Characteristics of linked-firms relative to focal firm:									
Smaller	-0.197		-0.191	-0.199	-0.197	-0.175	-0.167	-0.181	-0.173
	(-1.83)		(-1.77)	(-1.87)	(-1.85)	(-1.61)	(-1.53)	(-1.68)	(-1.60)
Same industry		0.093	0.077	0.067	0.075	0.066	0.069	0.060	0.063
		(0.82)	(0.67)	(0.58)	(0.65)	(0.58)	(0.60)	(0.52)	(0.55)
Business aspects of linked firms:									
Govt. contracts						-0.189		-0.188	
						(-1.46)		(-1.45)	
Defense contracts							-0.065		-0.063
							(-0.44)		(-0.43)
CI & emerg. tech.						0.031	0.025	0.028	0.022
projects						(0.25)	(0.20)	(0.22)	(0.17)
IP and trade secrets						0.111	0.100	0.111	0.099
						(0.96)	(0.87)	(0.96)	(0.87)
Business aspects of the focal firm:									
Govt. contracts				-0.204					
				(-1.60)					
Defense contracts					-0.100				
					(-0.74)				
CI & emerg. tech.				0.170	0.145				
projects				(1.35)	(1.13)				
IP and trade secrets				0.028	-0.003				
				(0.24)	(-0.03)				
Characteristics of cyber compromise:									
Nation-state actor								0.281	0.281
								(1.15)	(1.15)
Internal attack								0.021	0.024
vector								(0.10)	(0.11)
Adj. R^2	0.030	0.006	0.034	0.066	0.046	0.050	0.042	0.070	0.062

Appendix

A1 Predicting subsets of cyber events

We rerun model (1) to predict two subsets of cyber events, only those that are attributed to nation-state actors and only those that can be thought of as an industrial-espionage-related cyber intrusions (specifically, cyber incidents classified as "IP theft" or "security breach"). The results are reported in the Appendix Table A6.

Panel A of the table predicts only the cyber events attributed to the nation-state actors (the indicator variable is set to one if a firm experienced at least one cyber event attributed to a nation-state in the following year and to zero otherwise). (The unconditional probability of a firm experiencing at least one cyber incident attributed to a nation-state actor is 0.05 percentage points per year.) In the regression specification with industry fixed effects, the coefficient on the government contracts indicator variable is significantly positive, and the coefficients on the indicator variable for critical and emerging technology projects is significantly positive in all regression specifications; the estimated regression coefficient of 0.1 percentage points implies that having projects in critical infrastructure and emerging technologies increases the likelihood of a cyber incidents attributed to nation-states by 100%.

Panel B considers only a subset of cyber espionage events, those that are classified in the "IP theft" and "security breach" categories. (The unconditional probability of a firm experiencing at least one such cyber espionage cyber incident is 0.03 percentage points per year.) The specifications with industry fixed effects show positive significant coefficients on dummies for government and defense contracts, critical infrastructure and emerging-technology projects and trade secrets. Likely because of the very low prevalence of such events, in both panels, the coefficients in the firm-fixed-effects specifications are statistically insignificant.

Table A1
Categories of adverse cyber incidents
This table provides explanations and sample headlines for each category of cyber events.

Category	Explanation	Sample headline
PII breach	Theft of personally identifiable information by cyber means	Apple's website for developers was accessed by unauthorized parties Registered developer names, mailing addresses, and email addresses may have been accessed on Thursday, July 18. (source: Veris)
Security breach	Breach of corporate IT systems and devices	Sears Holdings Corp <shld.o> is "actively reviewing" its systems to determine if it has been victim of security breach. (source: Refinitiv)
Skimming/ theft of funds	Detection of skimming devices on card reading equipment, possible theft of funds	Self-checkout skimmer at Walmart. (source: Veris)
Overt cyberattack	Overt malware attack on corporate IT infrastructure that adversely impacts normal operations of the firm	Patrick Industries - cyberattack impacted certain of co's administrative & production servers & resulted in disruption of operations. (source: Refinitiv)
Ransomware	Malware attack on corporate IT infrastructure that renders it inoperable until ransom is paid	Brief-DXC identifies ransomware attack on part of its xchanging environment. (source: Refinitiv)
Cyber lawsuit	A lawsuit filed against a firm by companies or persons negatively affected by cyber incidents or vulnerabilities	Amazon's ring cameras are vulnerable to hackers, lawsuit in U.S. claims. (source: Refinitiv)
IP theft	Theft of corporate intellectual property by cyber means	A former senior software architect at Microsoft has been arrested and charged over allegedly stealing Windows 8 trade secrets. (source: Veris)
Security flaw	Discovery of a security flaw that exposes a firm to a network intrusion of data theft	Relay Medical Corp. addresses the recent cybersecurity vulnerability in surveillance and security cameras market. (source: Refinitiv)
Malware (non-ransom)	Discovery of malicious code on the company IT network	ABB's Swedish computer network hit by worm. (source: Refinitiv)
DDoS attack	Distributed denial-of-service attack, which is a malicious action to disrupt access to connected online services and sites	eBay <ebay.o> says hacker attack shuts down site. (source: Refinitiv)
Control systems compromise	Compromise of software and equipment responsible for controlling industrial processes	Ukraine's telecommunications system has come under attack, with equipment installed in Russian-controlled Crimea used to interfere with the mobile phones of members of parliament, the head of Ukraine's SBU security service said on Tuesday. (source: Refinitiv)

Table A2
Classifying cyber threat actors

This table provides sample headlines for identifying cyber threat actors.

Actor	Explanation	Sample headline
Nation-state	Malicious cyber actors located abroad and directed by their government	American Airlines, Sabre said to be hit in hacks backed by China - Bloomberg (source: Refinitiv)
Corporate insider	An employee of the company with internal access to corporate systems	Finance employee stole PII on patients from database and gave to co-conspirator for fraud. (source: Veris)

Table A3
Classifying cyber attack vectors

This table provides sample headlines for identifying cyber attack vectors.

Vector	Explanation	Sample headline
Supply chain	Compromise launched through a provider or partner with access to corporate IT systems	Symantec confirms ASUS software supply chain attack. (source: Refinitiv)
Internal	Compromise occurred through internal systems	Internal actor downloaded a trove of company documents about 40 gigabytes over a four-year period, including code. (source: Veris)

Table A4
Identifying firms' business risks

This table provides sample headlines for a firm's business activities that may attract the interest of cyber threat actors.

Classification	Explanation	Sample headline
Government contracts	A firm has a contract with a U.S. or foreign government agency	Leidos Awarded $365 Million Department of Energy Research Support Contract.
Military contracts	A firm has a contract with a defense or intelligence agency in the U.S. or abroad	IMMURON LTD - STUDIES COMMISSIONED BY US DEPARTMENT OF DEFENSE TO EVALUATE TRAVELAN'S ABILITY TO NEUTRALISE PATHOGENIC BACTERIA
Critical infrastructure sectors and new tech. projects	A firm is involved in work on critical infrastructure sectors or new-technology projects	QUANTUM-SI, A PIONEER IN SEMICONDUCTOR CHIP-BASED PROTEOMICS, TO COMBINE WITH HIGHCAPE CAPITAL ACQUISITION CORP.
Trade secrets and IP	Possession of intellectual property or trade secrets	Meridian Waste Solutions' Attis Innovations Executes License for Proprietary Biofuel Process Technology

Table A5
Variable definitions

This table provides a detailed description of the variables used in the analysis. All variables are computed as of the end of the prior month or year. We assume that accounting variables become publicly known three months after the annual report is published.

Advertising expenditure/Assets. Advertising expendisture scaled by total assets (*xad/at*). Source: Compustat.

Analyst coverage. Number of analysts issuing current year's earnings-per-share forecast (*numest*). Source: I/B/E/S.

Book-to-market. Book value of common equity (*ceq*) divided by the market value of common equity (*prcc_f* × *csho*). Source: Compustat.

Cash. Cash holdings (*ch*). Source: Compustat.

Critical infrastructure and emerging technology projects. Indicator variable set to 1 if the firm had stories about critical infrastructure or new-technology projects in the Refinitiv dataset in prior year and 0 otherwise. Source: Refinitiv.

Durable goods industries. Indicator variable set to 1 for firms with SIC codes between 3400 and 3999 and 0 otherwise. Source: Compustat.

Financial industry. Indicator variable set to 1 for firms with SIC codes between 6000 and 6999 and 0 otherwise. Source: Compustat.

Firm age. Number of years since the firm first appeared in Compustat dataset. Source: Compustat.

Firm size. Natural logarithm of the total market value of the firm (*at* − *ceq* + *prcc_f* × *csho). Source: Compustat.*

Government contracts. Indicator variable set to 1 if the firm had government contract-related stories in the Refinitiv dataset in prior year and 0 otherwise. Source: Refinitiv.

Intangibility. Asset intangibility, measured as 1−total property, plant and equipment (*ppent*)/total assets (*at*). Source: Compustat.

Institutional ownership. Total number of shares held by institutions that own more than 5% of a firm's equity divided by the number of shares outstanding (*shrout*). Sources: Thomson-Reuters 13F, CRSP.

Market capitalization. Natural logarithm of the market value of common equity (*abs(prc)* × *shrout*). Measured as of June of previous year. Source: CRSP.

Defense contracts. Indicator variable set to 1 if the firm had defense or intelligence contract-related stories in the Refinitiv dataset in prior year and 0 otherwise. Source: Refinitiv.

News coverage. Total number of news stories that mentioned the firm over the prior 12 months. Source: Refinitiv.

Number of funds. Total number of funds that own more than 5% of a firm's equity. Source: Thomson-Reuters 13F.

R&D expenditures/Assets. R&D expenditures (*xrd*); missing values set to 0. Source: Compustat.

ROA. Operating income before depreciation (*oibdp*) divided by total assets (*at*). Source: Compustat.

Retail industry. Indicator variable set to 1 for firms with SIC codes between 5200 and 5999 and 0 otherwise. Source: Compustat.

Tobin's Q. Total assets (at) − common/ordinary equity (ceq) + market value of equity ($prcc\_f \times csho$) to total assets (at). Source: Compustat.
Trade secrets. Indicator variable set to 1 if the firm had stories about trade secrets or intellectual property in the Refinitiv dataset in prior year and 0 otherwise. Source: Refinitiv.

Table A6
Predictors of cyber events attributed to nation-states and likely espionage

This table presents the results of a linear probability model explaining the subset of cyber events: events attributed to nation-states in Panel A and likely espionage-related intrusions in Panel B. All accounting variables are considered to be publicly known three months after the fiscal year end. Models (1), (2), (3), and (4) correspond to models (3), (5), (7) and (9) of Table 2 Panel A, respectively. The sample period is January 1999—January 2022. Standard errors are clustered by firm and year, and t-statistics are reported in parentheses.

Panel A: Predicting cyber events attributed to nation-states

Model	(1)	(2)	(3)	(4)
Govt. contracts	0.003^b	0.002	.	.
	(2.02)	(1.25)	.	.
Defense contracts	.	.	0.004	0.003
	.	.	(1.60)	(0.84)
Critical infr. &	0.001^b	0.001^b	0.001^b	0.001^b
emerging tech. proj.	(2.31)	(2.18)	(2.18)	(2.17)
Trade secrets	0.001	-0.000	0.001	-0.000
	(1.38)	(-0.24)	(1.40)	(-0.24)
Controls	Y	Y	Y	Y
Year dummy	Y	Y	Y	Y
Ind. dummy	Y	N	Y	N
Firm dummy	N	Y	N	Y
Obs.	81473	81480	81473	81480
Adj. RSq.	0.010	0.001	0.010	0.001

Panel B: Predicting cyber espionage events (events classified as IP theft or security breach)

Model	(1)	(2)	(3)	(4)
Govt. contracts	0.004^b	-0.000	.	.
	(2.41)	(-0.18)	.	.
Defense contracts	.	.	0.006^c	0.003
	.	.	(1.77)	(0.82)
Critical infr. &	0.002^b	0.002	0.003^c	0.0000
emerging tech. proj.	(2.00)	(0.20)	(1.95)	(0.15)
Trade secrets	0.003^b	-0.001	0.004^b	-0.001
	(2.04)	(-0.72)	(2.05)	(-0.75)
Controls	Y	Y	Y	Y
Year dummy	Y	Y	Y	Y
Ind. dummy	Y	N	Y	N
Firm dummy	N	Y	N	Y
Obs.	81473	81480	81473	81480
Adj. RSq.	0.025	0.002	0.025	0.002

Table A7
Stock price reaction to cyber events, Refinitiv news sample only

This table presents results for stock price reactions to cyber incident announcements that contain only announcements in the Refinitiv dataset. Cumulative abnormal returns (CARs), reported in %, are based on three abnormal return specifications: (1) market-adjusted, (2) adjusted for the Fama-French three factors, or (3) adjusted for the Fama-French three factors and the momentum factor, as specified above each column. Unless specified otherwise, the sample period is January 1999—January 2022. Standard errors are clustered by firm and year, and t-statistics are reported in parentheses.

Panel A: Abnormal returns calculated over different event windows

	CAR (%)					
Event	Entire sample			Jan 2014–Jan 2022		
window	Mkt-adj.	FF-adj.	4-factor-adj.	Mkt-adj.	FF-adj.	4-factor-adj.
$[-1, 0]$	-0.734	-0.737	-0.768	-0.777	-0.788	-0.833
	(-5.46)	(-5.76)	(-5.92)	(-5.13)	(-5.48)	(-5.73)
$[-1, +1]$	-1.076	-1.029	-1.063	-1.157	-1.124	-1.161
	(-6.03)	(-6.04)	(-6.13)	(-5.47)	(-5.57)	(-5.65)
$[-1, +5]$	-1.036	-1.074	-1.075	-1.146	-1.225	-1.242
	(-4.65)	(-4.84)	(-4.87)	(-4.56)	(-4.89)	(-4.97)
$[-1, +10]$	-1.288	-1.409	-1.434	-1.190	-1.308	-1.289
	(-3.58)	(-4.10)	(-4.17)	(-3.86)	(-4.37)	(-4.32)
$[-1, +15]$	-1.563	-1.688	-1.669	-1.687	-1.876	-1.871
	(-4.38)	(-4.95)	(-4.85)	(-3.88)	(-4.56)	(-4.49)
$[-1, +20]$	-1.498	-1.645	-1.593	-1.577	-1.774	-1.766
	(-3.60)	(-4.20)	(-4.07)	(-3.05)	(-3.68)	(-3.66)
$[-1, +25]$	-1.481	-1.648	-1.591	-1.591	-1.836	-1.831
	(-3.34)	(-3.87)	(-3.75)	(-2.96)	(-3.62)	(-3.60)

Panel B: By cyber category
Abnormal returns are computed over event window [-1, +1]

Cyber category	4-factor-adj. CAR (%)
Unclassified	-1.589[a]
	(-3.40)
PII theft	-1.041[a]
	(-3.62)
Security breach	-0.645[b]
	(-2.10)
Skimming/theft of funds	-1.743
	(-1.20)
Overt cyberattack	-1.226[c]
	(-1.80)
Ransomware	-1.423[b]
	(-2.49)
Cyber lawsuit	-2.072
	(-1.25)
IP theft	-0.114
	(-0.24)
Security flaw	0.481
	(1.25)
Malware (non-ransom)	-2.052[c]
	(-1.88)
DDoS	-3.609[a]
	(-2.81)