Do Software Companies Spread Cyber Risk?*

Giorgio Ottonello† Antonino Emanuele Rizzo‡

August 29, 2024

Abstract

We show that software companies are a systemic source of cyber risk in the economy due to software vulnerabilities that spread to customer firms throughout the digital supply chain. We introduce a novel database that connects vulnerability discoveries and related cyberattacks to software companies and their customer firms. Exposure to vulnerabilities i) increases the likelihood of cyberattacks and firm-level risk metrics and ii) negatively impacts customers' investment rates as well as sales growth. Market participants are slow to react to vulnerability announcements, likely overlooking the supply chain connections between software companies and their customers. Vulnerabilities are more harmful when coming from a software company with a larger market share and have aggregate effects at the industry level.

JEL-Classification: G24, G32
Keywords: cyber risk, software vulnerabilities, software companies, firm risk, digital supply chain

*We thank for useful comments Robert Battalio, Zhi Da, Nuri Ershain, Xu Qiping as well as seminar participants at the University of Notre Dame, Southern Methodist University, University of Illinois Urbana-Champaign, Nova SBE and conference participants at SGF 2024. This work was funded by Fundação para a Ciência e a Tecnologia (UID/ECO/00124/2013 and Social Sciences DataLab, Project 22209), POR Lisboa (LISBOA-01-0145-FEDER-007722 and Social Sciences DataLab, Project 22209) and POR Norte (Social Sciences DataLab, Project 22209). Any remaining errors are our own.

†NOVA School of Business and Economics (NOVA SBE), Email: giorgio.ottonello@novasbe.pt
‡ESADE Business School, Email:emanuele.rizzo@esade.edu

1 Introduction

Software companies are becoming increasingly important in today's digital economy. According to Software.org (2021), software companies provided in 2020 15.8 million U.S. jobs and contributed $933 billion in direct value-added GDP to the U.S. economy alone. Figure 1 illustrates the significant increase in the relevance of the software industry between 2006 and 2023. During this period, the market capitalization of software companies relative to the total U.S. stock market almost doubled, rising from 12% to 23%.

At the same time, another equally impressive trend has emerged: the growth of cybersecurity risk. Today, cyber risk is a top priority for both corporations and governments (Florackis et al. (2022); Jamilov, Rey, and Tahoun (2023)). Global damages from cybercrime are projected to reach $10.5 trillion by 2025, up from $3 trillion in 2015, according to Cybersecurity Ventures (2022). This escalating threat has driven significant regulatory initiatives, such as the European Union's GDPR and the United States' CISA. Our own estimates, illustrated in Figure 1, indicate that the number of cyberattacks on publicly traded U.S. companies has increased almost ninefold, rising from 10 to nearly 90 per year.

In this paper, we argue that the growth of software companies is a key driver of the rise in cybersecurity risk. Our hypothesis centers on the fact that the primary products of software companies—software programs—are frequently susceptible to software vulnerabilities (SVs). These are flaws within software products that hackers can exploit to launch cyberattacks on organizations utilizing such software. As software companies grow and their market share expands, a larger number of firms across the economy become exposed to these SVs, thereby increasing the likelihood of suffering a cyber attack.

To empirically assess our conjecture, we begin by compiling a novel database that links each software vulnerability discovered between 2006 and 2023 to its corresponding software company and, through it, to its customers. We further integrate this with a comprehensive list of cybersecurity attacks, which offers two significant advantages over what is currently available in the literature. First, by drawing from multiple sources, we compile a more exhaustive list of cyberattacks. Second, to the best of our knowledge, we are the first to identify all existing links between

each cyberattack and specific vulnerabilities in software products.

Armed with this data, we begin by studying the link between SVs and cybersecurity risk. Our baseline analysis utilizes a panel of software companies' customers, each of whom can be exposed to a SV through supply chain connections with the software company providing the vulnerable product. We document that exposure to a SV significantly increases the likelihood of suffering a cyberattack in the following quarter. Specifically, exposure to an additional software vulnerability predicts an increase in cyberattacks by 0.106 standard deviations (t-stat 4.41). Moreover, when we restrict the sample to cyberattacks that can be traced back to a specific software vulnerability used by the firm, the effect intensifies to 0.187 standard deviations (t-stat 5.75). In contrast, it is reassuring to observe that there is no predictability of cyberattacks that cannot be directly linked to SVs.

Several additional tests provide strong empirical support to the interpretation that cyberattacks are driven by vulnerabilities identified within the software products. First, we document similar results in an alternative empirical setting based on a software company-level panel, where the dependent variable is the number of cyberattacks experienced by all customers of the software company. Analyses performed in this setting confirm that the emergence of a SV significantly increases the likelihood that customers of the company selling the vulnerable software suffer from cyberattacks. Second, we perform a falsification test that exploits the severity scores assigned to SVs, which evaluate the level of threat that the vulnerability poses to a company if it is successfully exploited by an attacker. In this test, we isolate SVs classified as minor and demonstrate that they do not affect the probability of cyberattacks. Lastly, using the block permutation method from Chetty, Looney, and Kroft (2009), we randomize the timing of vulnerabilities and demonstrate that the timing is crucial in determining the effect on the frequency of cyberattacks. Jointly, these tests help to ensure that our findings are not influenced by unobservable characteristics that might drive the matching between firms and software suppliers, thereby affecting the number of cyberattacks independently of SVs.

Next, we investigate the relationship between SVs and firm risk. If companies using software with vulnerabilities are more exposed to cyberattacks, we should observe a measurable impact

of SVs on standard risk metrics. To assess this hypothesis, we adopt again both the empirical setup based on a customer-level panel and the one using the software company-level panel. In both settings we regress one-quarter-ahead firm risk on a vulnerability dummy variable. We proxy firm risk using several measures: stock return volatility, idiosyncratic volatility, 95% Value at Risk (VaR), and second-order lower partial moment (LPSD). The first two proxies are standard measures of firm risk, capturing overall volatility and firm-specific risk, respectively. The latter two proxies are designed to capture tail risk, given that cyberattacks are typically extreme, low-frequency events.

Regardless of the proxy used, we find that being exposed to a SV increases firm risk in the following quarter. The magnitude of the increase ranges between 0.035 and 0.053 standard deviations in the firm-level sample and between 0.011 and 0.014 standard deviations in the software supplier-level sample, where we average the risk measures across all customer firms. The size of the effect is economically meaningful and comparable to effects documented in prior research for other types of risk.[1] Interestingly, the effect is highest for VaR and LPSD, consistent with the idea that cyber risk is best captured by measures of tail risk.

Given the significant impact on measures of firm risk, we next asses the economic relevance of SVs by studying their effect on firms' real activities. These tests are motivated by theoretical and empirical work that links increases in firm risk with decreases in firms' investment, as capital expenditures or R&D (see, e.g., Bloom, Bond, and Van Reenen (2007), Bloom (2009), Hassan et al. (2019)). Consistent with prior research, we find that the emergence of a SV, and the associated rise in firm risk, depresses firms' capital investment rate and R&D investment rate. Specifically, the discovery of an additional vulnerability leads to a decrease in both tangible and intangible investment rates by 0.044 and 0.043 standard deviations, respectively. In exploring whether an increase in cyber risk negatively impacts the overall firm business, we observe that the discovery of a SV is also followed by a decrease of sales growth. Finally, in response to the emergence of a SV, firms are more likely to hire a cybersecurity specialist company. This finding is interesting as it supports the interpretation that our measure indeed captures an increase in cybersecurity risk

[1] See for example the impact of political risk on firm stock volatility found by Hassan et al. (2019).

and shows how firms typically respond to this new type of threat.

From this set of empirical tests, we draw two main conclusions. First, SVs are a primary driver of cyber risk transmission among firms, significantly impacting overall company risk as well as depressing investments and sales. Second, software companies play a crucial role as catalysts of cyber risk due to the vulnerabilities present in their widely used products. A key question arising from our empirical evidence is how financial markets incorporate information on cyber risk stemming from SVs. This issue is particularly significant because cyber risk has only recently gained high relevance, making accurate measurement inherently challenging. Underscoring the importance of our question, Gomez Cram and Lawrence (2024) provide evidence that market participants have consistently undervalued software companies over the past 20 years. This suggests that the financial implications of SV-related cyber risks may not be fully appreciated by the market.

We begin by studying the short-term market reaction to the announcement of a SV affecting the company. In absence of frictions, market participants should immediately incorporate the fact that exposure to SVs leads to an increase in cyber risk and interpret it as negative news. The market reaction is measured through cumulative abnormal returns (CAR) in a 3 days window surrounding the event which includes the event day, the day before and the one after. We employ a different setting based on stacked cross-sections of firms, one for each vulnerability discovery. Each cross-section consists of the sample of companies exposed to the SV, as well as a control group of firms made of all the other companies in our data that are not affected by the SV. Independently from the model we use to construct CAR (market adjusted, CAPM, Fama and French 5 factor), we find no significant reaction to the SV announcement. This result indicates that, on average, market participants are not able to immediately understand the implications of SV exposure for software companies' customers.

The lack of market reaction to the discovery of SVs can be explained by three possible reasons: i) SVs may not significantly impact companies, leading to no market reaction; ii) market participants may be unaware of cyber risks; iii) market participants may react slowly to information about SVs. The explanations i) and ii) are less likely. First, we have documented that SVs are followed by an increase of the probability of cyberattacks as well as lower investments and sales

growth. Second, market-based firm-level risk measures increase on a quarterly basis following a SV discovery. We conduct additional tests to assess the empirical support for explanation iii). We examine the impact of SVs on firm stock returns over longer horizons, based on the premise that if information is incorporated slowly, a market reaction to the discovery of a SV should be detectable later. Using the same empirical setting employed for firm-level risk variables, we analyze one-quarter-ahead firm-level risk-adjusted stock returns as the dependent variable. Our findings indicate a negative and significant stock market reaction to the discovery of SVs at the quarterly frequency. The economic magnitude ranges between -0.6% and -0.8%, depending on the risk adjustment method used. We conclude that market participants are slow to incorporate information about cyber risks originating from the discovery of SVs.

While our results strongly support the slow incorporation of information following SVs discoveries, they are silent about the frictions that prevent a more timely market reaction. We argue that market participants fail to immediately take into account the supply chain linkages between software companies and their customers. This is consistent with the findings of Cohen and Frazzini (2008), who show that investors struggle to incorporate customer-supplier linkages in stock prices. In support of this explanation, we provide evidence that market participants quickly react to information on cyber risk when there are no supply chain linkages to be considered.

In the first step, we analyze the stock market reaction of software suppliers to the discovery of a vulnerability in the software they produce and sell. In this scenario, market participants do not need to consider supply chain linkages between the vulnerable product and its customers, as the vulnerability is directly associated with its producer. SVs are undeniably negative news for software companies, as they need to invest resources in fixing the problem and their reputation is likely adversely affected. Consistent with the proposed mechanism, we find that market participants react immediately to the discovery of a SV when supply chain links are not a factor. The cumulative abnormal return (CAR) of software companies within the three-day window around the event is negative and significant, ranging between -0.22% and -0.27%, depending on the risk adjustment. Consistently, when examining software companies' stock returns at the quarterly horizon, the coefficient is negative but not significant. Overall, our results suggest that the negative

news is incorporated into stock prices immediately.

Second, we analyze the short-term stock market reaction of customer firms following the occurrence of a cyberattack that originates from a SV. This type of event is useful for studying the frictions that cause the slow incorporation of information, as it is directly related to cybersecurity risk but there is no need to consider customer-supplier links to identify the affected company. In this case, we observe a negative and significant effect on the cumulative abnormal return (CAR), ranging between -1.08% and -1.28%, depending on the asset pricing model used to construct CAR. Overall, our results are consistent with market participants failing to immediately incorporate the effect of SVs on companies using the flawed software because they struggle to account for supplier-customer links in a timely manner. Consequently, they only manage to do so with a delay, resulting in the slow incorporation of information about cyber risk stemming from SVs.

In the final part of the paper, we explore the idea that software companies can be a systemic source of cybersecurity risk for other firms in the economy. Recent events, such as the CrowdStrike global IT crash of July 2024, have stimulated the discussion about the systemic nature of software companies (Welburn (2024)).[2] Intuitively, if the size and market share of software companies influence the transmission of cybersecurity risk, any effect we document should be more pronounced when the SV originates from a larger software company with a more significant market share. This is because a larger customer base using vulnerable software means more firms can be attacked by exploiting that vulnerability, leading to an aggregate increase in expected cyberattacks across all customers. At the individual customer firm level, using vulnerable software from a larger software company may result in a higher probability of being a cyberattack victim. Ceteris paribus, hackers obtain a higher payoff by exploiting a vulnerability that affects a large number of companies compared to one that affects only a few. Consequently, a vulnerability in more popular software attracts more hackers attempting to exploit it, thereby increasing the likelihood of cyberattacks.

To explore the role of software supplier market share, we repeat our tests by adding the interaction between the vulnerability dummy variable and a variable capturing the software company's

[2]Further, in April 2024 the U.S. government has instructed the Cybersecurity and Infrastructure Security Agency (CISA) to identify and categorize certain critical infrastructure entities as Systemically Important Entities, with software companies being among those. See here for details.

market share. The latter is defined as the ratio of the cumulative market capitalization of software companies' customers to the total market capitalization of customers in our data. We find that a relatively larger software company strengthens the predictive relationship between vulnerabilities and cyberattacks. At the software company level, when aggregating the cyberattacks of all its customers, a one standard deviation increase in the software supplier's market share amplifies the effect of a vulnerability by 0.015 standard deviations. This is a substantial increase, considering that the baseline impact of SVs is 0.031 standard deviations. The effect is only driven by cyberattacks directly linked to the vulnerabilities of the software supplier and is absent when using cyberattacks not due to vulnerabilities as a dependent variable. A regression at the customer level yields the same results. Consistent with the findings on cyberattacks, we also show that SVs originating from larger software companies have a stronger impact on customer firms' risk measures and real effects, both at the customer level and when aggregating risk across all of a software company's customers. These findings suggest that the recent growth of software companies documented in Figure 1 can help explain the steep increase in cyberattacks over the same period. From a policy perspective, regulators should consider that a more concentrated software industry can significantly increase cyber risk in the economy and devise tools to better manage this trend.

Next, we study the aggregate implications of SVs. If software companies were systemic source of cyber risk, their impact should be economically meaningful also at the aggregate level. Specifically, we test whether our results hold at the industry level. Ex-ante, this is not clear. While we document that customers using vulnerable software experience increased risk and lower stock returns at the quarterly frequency, these adverse effects might be counterbalanced by positive reactions from competitors, potentially nullifying the aggregate effect. Additionally, if a company is hit by a cyberattack, competitors might respond by strengthening their cybersecurity measures, thereby reducing overall industry risk. We repeat our main tests at the industry level, using the Fama and French 49 industry classifications. We find that the number of discovered SVs affecting customers in an industry strongly predicts future cyberattacks on firms within that industry. Similarly, an increase in discovered SVs leads to an increase in industry-level risk and lower stock returns at the quarterly frequency. The magnitudes of these effects are largely in line with those

documented at the firm level. We conclude that SVs have significant aggregate implications and can substantially alter the exposure to cyber risk across entire industries.

We contribute to the growing literature on cybersecurity and finance. Recent papers have produced measures of cybersecurity risk at the firm level (Jamilov, Rey, and Tahoun (2023), Florackis et al. (2022)). Others have instead focused on the effect of cyberattacks on firms (Crosignani, Macchiavelli, and Silva (2023)), the financial system (Duffie and Younger (2019), Eisenbach, Kovner, and Lee (2022), Kotidis and Schreft (2022), Eisenbach, Kovner, and Lee (2023)), or government institutions (Curti et al. (2023)). Our paper differs from this line of work as we approach the question from a new angle: we study the origins of cybersecurity risk. In this respect, we identify software companies (through SVs) as an important source of cybersecurity risk for firms. Our evidence indicates that SVs propagate cybersecurity risk through the digital supply chain network of the software company, even to customers that belong to different industries. In that respect, we also contribute to the literature that examines how supply chains can serve as a medium through which risks are transmitted in the economy (Hertzel et al. (2008), Barrot and Sauvagnat (2016), Carvalho et al. (2021)). We add to this literature by introducing another important dimension of risk that propagates though supplier-customer links: cyber risk.

Further, our paper relates to the literature that studies the diffusion of new technologies and its real and financial effect.[3] Within this strand of the literature, Gomez Cram and Lawrence (2024) study the diffusion of the software industry and show that market participants have persistently undervalued software companies in the last decades. Our contribution lies in showing that the diffusion of new technologies (in this case software) can be an important contributor to the growth of novel types of risk for all firms in the economy.

Finally, we contribute to the literature that studies the diffusion of information in financial markets. A large number of papers have documented (theoretically and empirically) that investors do not incorporate news timely in market prices, leading to slow diffusion of information.[4] Cohen

[3]The literature is extremely vast. For real effects, see for example Autor, Levy, and Murnane (2003), Acemoglu and Restrepo (2020), Akerman, Gaarder, and Mogstad (2015), Bloom et al. (2021). For the impact on financial markets, see Pástor and Veronesi (2009), Kogan et al. (2017), Ward (2020).

[4]Merton (1987), Hong and Stein (1999), Peng and Xiong (2006), Hong, Torous, and Valkanov (2007), DellaVigna and Pollet (2007).

and Frazzini (2008) demonstrate that this effect is particularly strong when there are supply chain linkages to be taken into account. Our paper provides an important application of this phenomenon to cybersecurity risk, by showing that investors neglect it when they need to consider the supply chain linkages between software companies and their customers.

The remainder of the paper is structured as follows. Section 2 provides some institutional background on SVs. Section 3 describes the data and the measures we use in analysis. Section 4 presents the main empirical findings. Section 6 concludes.

2 Institutional Background

This section provides some institutional background on SVs as well as the role that software companies play in this context. First, we discuss the process that brings to the public disclosure of SVs. This is important as in our data we observe a SV once it is publicly disclosed. Second, we describe how SVs are still harmful after they are disclosed and security patches are available. This discussion motivates our empirical setup, which is based on measuring the effect of SVs after they are disclosed to the public.

2.1 Timing of Public Disclosure of SVs

A SV is a flaw in an application or operating system, which can be used with security impact. SVs pose a high risk to users of the affected application because cybercriminals race to exploit these vulnerabilities to cash in on their schemes. There are three broad groups of actors that can discover SVs. The first one are the software companies themselves, who test their products for abnormal behavior to locate the vulnerability before an attack is launched. The second group are independent security researchers, who cooperate with vendors and scan their software searching for flaws. Finally, a malicious hacker may be the first to discover the vulnerability. This is the worst case scenario, since there is no way to guard against the exploit before an attack happens.

After a vulnerability is disclosed, several considerations influence the decision of when to make it public. On the one hand, software companies may prefer that SVs are disclosed only to them-

selves initially and made public only after the patches are introduced. Damages to reputation and brand image can be one reason for that, as disclosing SVs can tarnish a software vendor's reputation and erode customer trust. At the same time there are legal and regulatory concerns, because disclosing SVs could potentially expose vendors to legal liabilities, especially if a vulnerability results in data breaches or other forms of harm to users. Indeed, information on a SV can give attackers who were not otherwise sophisticated enough to find the problem on their own the very information they need to exploit a security hole in a computer or system and cause harm. On the other hand, public disclosure of security information enables informed consumer choice and may incentivize vendors to repair vulnerabilities and build more secure products. Moreover, even if a patch is not available, disclosure may still be preferable to inform other researchers, security experts, and the broader community about the vulnerability. Cybersecurity professionals and enterprises whose sensitive data or systems is at risk may try to find other ways to mitigate or eliminate the threat.

The debate regarding the publication of information about SVs is still ongoing and it is one of the most vigorous public policy discussions in the security field. The importance of clear vulnerability disclosure rules is testified by the several policies emerging in recent years. In the US, for example, it is the Cybersecurity and Infrastructure Security Agency (CISA) that is responsible for coordinating public disclosure of newly identified cybersecurity vulnerabilities, while in the European Union it is the European Union Agency for Cybersecurity (ENISA) that does a similar job. However, while vulnerabilities can be discovered anywhere in the world a software is available and actively used, there are no globally accepted standards for disclosing SVs.

Whatever the set of rules that apply, a SV can be disclosed even if a patch does not exist. The CISA policy, for example, is to "disclose vulnerabilities as early as 45 days after the initial attempt to contact the vendor is made regardless of the availability of a patch or update".[5] In this respect, a particularly harmful type of SVs is so-called zero-day vulnerability. This is an application flaw for which there is no solution provided from software companies and the vulnerability is being actively exploited by malicious actors.

[5] CISA vulnerability disclosure process.

2.2 Are SVs Harmful after Disclosure?

Regardless of the existence of a patch, a vulnerability can be dangerous for the users of an application years after its discovery (e.g., August, Niculescu, and Shin (2014), August et al. (2022)). The US Department of Homeland Security has estimated that approximately 80 percent of successful security breaches involve unpatched software (Microsoft report). The key issue is that few companies are able to properly patch all vulnerabilities to which they are exposed. For example Tenable, a cybersecurity company, reports that only 10% of organizations addressed all their open vulnerabilities within a year of first assessment (Tenable report). There are at least three reasons why companies remain vulnerable despite patches are available.

First, the sheer number of patches a company should be aware of and apply is monumental. On average, a company uses around 110 applications, a figure that has been steadily growing since 2015 (Bettercloud report) Hence, IT specialists must monitor on average 110 applications for the updates and, when necessary, patch them. This issue is exacerbated by the fact that many organizations do not have full and complete inventories of all applications and software they are using. Moreover, the process of patching a single vulnerability is time consuming, as companies need to navigate the various processes involved in the patch and test it to ensure it does not interfere with the company's operations.[6]

Second, and related, many companies lack the staff and the expertise to patch all vulnerabilities in a proper way. Today's extremely complex IT systems impose a heavy burden on company's IT specialists that need to understand the vast network dependencies of each application to apply a patch effectively.

Finally, there are situations where patching simply is not possible. For example, a legacy software that a company requires for day to day operations may no longer be supported by the manufacturers. Perhaps they discontinued the product or are no longer in business. Either way, patches simply are not available. Another common issue is that patching requires that systems to be stopped and then rebooted to allow patches to be installed. This means that the services provided by that system are going to experience some downtime. Downtime costs are often very

[6]On average it takes 12 days for a company IT teams to coordinate for applying a patch across all company's devices (Ponemon Institute report).

high for companies. For some of them, even a small amount of downtime is not an option (e.g., healthcare sector).

To sum up, software vulnerabilities are an increasing threat to companies that is hard to contain. Consistent with the anecdotal evidence, our paper will demonstrate that the number of vulnerabilities a company is exposed to is an important source of cybersecurity risk.

3 Data and Variable Construction

3.1 Data

We obtain accounting data and stock prices of U.S. firms for the period 2006-2023 from the CRSP-Compustat merged dataset. Data on supply chains is obtained from the FactSet Revere Supply Chain Relationship database.[7] We will need supply chain information to link products affected by software vulnerabilities to the customers using them, as detailed in the next section. Our final sample of U.S. firms comes from the intersection of CRSP-Compustat merged and FactSet Revere Supply Chain Relationship. To be included in the sample, a company needs to appear in both databases at any given point in time. The customer-level sample contains 81,125 observations, pertaining, on average, to 1,431 firms every year and 3,875 unique customers. In Table 1, Panel A we present summary statistics for the main customer-level variables used in the paper.

Next, we collect the list of known SVs from two different sources. The first one is the list provided by the Cybersecurity and Infrastructure Security Agency (CISA). The second is based on the information available on the Zero Day Initiative (ZDI) website.[8] Both datasets provide information on the date in which the SV was first published; the Common Vulnerabilities and Exposures (CVE) identifier; an assessment of the vulnerability risk, through the Common Vul-

[7] FactSet Revere collects customer and supplier relationship information from primary public sources, such as SEC 10-K annual filings, investor presentations, and press releases. Its multiple sources imply that FactSet Revere is less affected by the limitation of datasets that rely on information from the Statement of Financial Accounting Standard (SFAS) No.131, which requires firms to disclose only the existence and sales to principal customers representing more than 10% of total firm revenues (see e.g., Adelino et al. (2023)).

[8] CISA (ZDI) data is downloaded from https://www.cisa.gov/known-exploited-vulnerabilities-catalog (https://www.zerodayinitiative.com/advisories/published/).

nerability Scoring System Version 3.0 (CVSS v3.0); and finally, the identity of the vendor of the product affected by the vulnerability (i.e, the software company). The vulnerabilities in our sample affect at least once the software products of 66 different software companies which we can identify in Factset Revere.

To identify firms in our sample that are victims of cyberattacks, assemble a novel database based on several data sources. The first is the Privacy Rights Clearinghouse (PRC) data. PRC is a nonprofit organization that collects information on data breaches and other cybersecurity incidents from publicly available, government-maintained data sources. PRC categorizes incidents into hacking attacks and data breaches originating from other causes. The second source is the VERIS Community Database (VCDB). VCDB provides a repository of widely collected public incidents and offers a classification of security incidents similar to that of PRC. The first two sources have been also used by other papers in the literare (e.g, Florackis et al. (2022)). We supplement this data with information retrieved from Hackmageddon, a website specializing in publishing cyberattack data and newspaper repositories.[9] Our final list of cyberattack-firm pairs includes 785 events and is far more comprehensive than what has been used in the literature so far. For example, Jamilov, Rey, and Tahoun (2023) cover a similar sample period and has in total 295 firm-cyberattack pairs.

We manually identify public US corporations in our list of cyberattacks. For each cyberattack affecting a public US firm, we manually search for information about the root cause of the attack. Whenever possible, we link a cyberattack to a specific software vulnerability and label the attack as vulnerability-related. In other cases, we classify the attack as phishing, DDoS, credential stuffing, malware, or miscellaneous. We are able to trace back cyberattacks to vulnerabilities in the software distributed by 43 unique software suppliers. We are also able to identify 100 unique customers affected by cyberattacks.

Last, we define our general sample of software companies, independently from whether they are hit by a vulnerability or not. First, we identify the list of 4-digit SIC and NAICS codes where at least one of our 66 software software companies affected by vulnerabilities operates. Next, we

[9]More details on PRC data are https://privacyrights.org/. Info on VCDB data is https://github.com/vz-risk/VCDB. For Hackmageddon, see https://www.hackmageddon.com/.

select all firms in the CRSP-Compustat database that operate within both one of the SIC 4-digit industries and one of the NAICS codes in our list. The resulting supplier-level sample consists of 26,660 observations. On average, we observe 445 software companies each year and 1,090 unique software suppliers from 2006 through 2023. The number of software companies in our sample slightly exceeds the total number of publicly listed software companies identified by Gomez Cram and Lawrence (2024) from April 2003 through December 2022, which is 1,030. This discrepancy arises from the differing objectives behind the data collection in the two studies. Gomez Cram and Lawrence (2024) focus on publicly traded companies where software is the primary source of revenue, aiming to measure the value of software with precision. In contrast, our sample includes any company involved in software production, regardless of whether it constitutes their primary revenue source, thereby encompassing any company whose products potentially generate software vulnerabilities. Table 1, Panel B reports summary statistics for the main supplier-level variables used in the paper.

3.2 Measuring exposure to SVs

For each of the 66 software companies in our sample whose product is hit by a vulnerability, we retrieve the list of customers from Factset Revere. We measure exposure to SVs by defining any firm that is a customer of one of the 66 software companies at the time of a SV discovery as being exposed to that vulnerability. In Table IA2 in the Internet Appendix we report the statistics on SVs for each of the software companies involved. The table reports the total number of vulnerabilities affecting each software company, the number of unique customers that we are able to link to each supplier through the supply chain, and the total market capitalization of these customers. The average number of vulnerabilities by supplier is around 14, with a minimum of 1 and a maximum of 332 for Microsoft. Suppliers in our list are linked on average to 58 customers, with a minimum of 1 and a maximum of 742 for Microsoft.

To validate our measure, we present In Table 2 the correlation between exposure to SVs and lagged firm-level covariates. In Panel A, we use an indicator as the dependent variable, which takes the value 1 if a firm is exposed to a SV in one of the software products it uses. In the first three

columns, we consider a list of standard firm-level characteristics: market capitalization, book-to-market ratio, profitability, leverage, and the cumulative return over the previous 12 months. We report the results for regression models without fixed effects, with firm fixed effects and with both firm and time fixed effects. Interestingly, none of the firm-level controls exhibit a significant and consistent relationship with the emergence of a software vulnerability (SV) across all regression models. In Column (4), we introduce the variable *Supplier MktShare*, calculated as the average market share of the firm's software suppliers. Notably, this variable is positively and significantly correlated with the likelihood of exposure to a vulnerability. One possible explanation is that malicious hackers are more motivated to identify vulnerabilities in products sold by suppliers with larger market shares, as these vulnerabilities could potentially impact a greater number of customers. This larger expected impact increases the potential payoff for hackers. The role of supplier market share is further examined in Section 5.1. Finally, we include a variable that counts the number of successful cyberattacks against the company in the previous year. This variable does not show a correlation with the likelihood of a vulnerability emerging, which is reassuring. It helps to rule out the possibility of reverse causality, where the emergence of a software vulnerability could be attributed to a successful hacking attack against the firm that uses the software.

In Panel B of Table 2, we repeat the same analysis using a panel measured at the software company level. The dependent variable is an indicator that takes the value of 1 if a software product sold by the software company is affected by a vulnerability. None of the observable characteristics we consider have a significant correlation with the likelihood of the discovery of a software vulnerability. Overall, the empirical evidence in Table 2 suggests that firm-level observables do not consistently predict SVs.

4 Empirical Results

This section presents the main results of the paper. In 4.1, we start by studying the effect of SVs on cyber risk, measured through the probability of cyber attacks. We do so in a standard panel regression, both at the customer firm level and aggregated at the level of the software company. Next, 4.2 uses the same empirical settings to investigate the effect of vulnerabilities on general

proxies of firm risk. Section 4.3 studies the real effects of SVs. Finally, in 4.4 we investigate how financial markets incorporate information on cyber risk stemming from SVs.

4.1 Occurrences of cybersecurity incidents

In the first part of our analysis, we aim to determine if SVs constitute a source of cybersecurity risk. To do this, we examine whether SVs increase the likelihood of future cyberattacks. If that was the case, a firm impacted by a SV should face a higher probability of experiencing a cyberattack compared to a company that is not affected. We test this conjecture in standard panel regressions at the customer firm level. The sample used in this setting includes firms in the intersection of the Factset Revere dataset and CRSP-Compustat merged, as previously described in the Data section.[10] We estimate the model:

$$CyberAttacks_{c,t+1} = \alpha + \beta_1 Vulnerability_{c,t} + \gamma X_{c,t} + \epsilon_{c,t+1} \tag{1}$$

where c indexes customers and t quarters. We use as dependent variable the number of cyber-attacks that hit the software user firm c in quarter $t+1$. We further decompose the dependent variable into cyberattacks related and unrelated to SVs and use each component as dependent variable. In all columns, the main independent variable is an indicator with value 1 if customer c is exposed to a vulnerability coming from one of its software suppliers in quarter t. $X_{c,t}$ is a set of controls that includes market capitalization, book-to-market, ROA, previous 12-month return, leverage, and the number of cybersecurity incidents that the firm c experienced over the previous year. Including occurrences of past cybersecurity attacks against a firm is important, and it is meant to control for a simple reverse causality story: the vulnerability in a software emerges as a consequence of a successful hacking attack against the firm that used that software. Under this alternative scenario, it is the cyberattack that causes the vulnerability, and not the other way around. By controlling for cyberattacks occurred in the past year we make sure these incidents cannot explain our results. α indicates different sets of fixed effects, including year-quarter and

[10] In Table IA5 in the appendix, we show that our baseline results are robust to adopting a stacked DiD estimation (see e.g., Baker, Larcker, and Wang (2022)). We describe this setting in more details in Section 4.1.1.

firm fixed effects. Standard errors are clustered at the customer level. For ease of interpretation we standardize the dependent variables.

The results are presented in Table 3, Panel A. Column (1) shows that the coefficient on $Vulnerability_{c,t}$ is positive and highly statistically significant. A firm affected by an additional vulnerability through its suppliers experiences an increase in the number of hacking attacks in the next quarter by 0.11 standard deviations. When we focus on the sample of cyberattacks that we can directly link to software vulnerabilities, the effect becomes stronger. Being exposed to one more vulnerability increases the number of hacking attacks in the next quarter by 0.19 standard deviations. In the last column, we focus on cyberattacks whose causes we cannot trace back to software vulnerabilities. This is useful to address a plausible concern: companies with poor internal cybersecurity controls, or with low awareness of cybersecurity risks, may be matched with suppliers with similarly poor controls, or low awareness. The products of these suppliers are plausibly more likely to present vulnerabilities. As a result, the documented effects would be driven by this matching rather than by vulnerabilities per se. Interestingly, the last column, which focuses on cyberattacks whose causes we cannot trace back to software vulnerabilities, shows a coefficient that is not statistically significant and of a smaller magnitude. Since these different attacks may still indicate inadequate internal controls or attention to cyber risks, the absence of an effect speaks against the alternative interpretation of the results based on the endogenous customer-supplier matching.

In Panel B of Table 3 we perform the same analysis in a panel at the software company level. The goal of this second analysis is to verify whether the customer-level results also hold at the supplier level. This a natural setting for investigating whether the cybersecurity risk affecting software customers indeed originates from their suppliers. Using our sample of software suppliers, we estimate the following:

$$CyberAttacks_{s,t+1} = \alpha + \beta_1 Vulnerability_{s,t} + \gamma X_{s,t} + \epsilon_{s,t+1} \tag{2}$$

where the unit of observation is a supplier s in quarter t. The dependent variables in this panel are the total number of cybersecurity incidents of any type, related to vulnerabilities, or due to

causes different from vulnerabilities that hit the software supplier s customers in quarter $t+1$. The set of controls $X_{s,t}$ includes market capitalization, book-to-market, ROA, previous 12-month return, leverage, and the number of cybersecurity incidents that the software supplier s customers experienced over the previous year. α indicates different sets of fixed effects, including year-quarter and software company fixed effects. Standard errors are clustered at the supplier level.

Coefficients in Panel B show that the customer-level results documented in Panel A hold at the supplier level. When the customers of software supplier s are exposed to an additional vulnerability, the probability of cyberattacks increases by 0.038 standard deviations. The effect is coming solely cyberattacks directly linked to vulnerabilities in columnn 2, while it is absent for attacks due to different causes in column 3.

4.1.1 Falsification tests

In this section, we report the results of two falsification tests to support our interpretation of the results that software vulnerabilities are a first-order driver of cybersecurity risk. Broadly, the falsification tests aim to mitigate the concern that unobserved factors (instead of SVs) are the driver of our main results.

In the first falsification test, we re-estimate equations (1) and (2) including only minor vulnerabilities when constructing $Vulnerability_{i,t}$. Minor vulnerabilities are defined as those vulnerabilities with a Common Vulnerability Scoring System (CVSS) below 7, and thus identified as medium or low risk. Focusing on these minor vulnerabilities is useful to address some of the alternative explanations based on unobserved factors. The underlying rationale is that a vulnerability, even if minor, could still be a signal of poor internal cybersecurity controls, or low awareness of cybersecurity risks. Therefore, if these unobserved factors were behind our results, we should still observe that minor vulnerabilities predict cybersecurity incidents. In contrast, since a vulnerability is minor when its estimated impact on the number of future hacking attacks is low, under our interpretation of the results we should not observe any significant relation betwen minor vulnerabilities and future attacks. Consistent the latter, Table IA4, shows that minor vulnerabilities have no impact on the number of cyberattacks in the next quarter.

With the second falsification test, we again address the possibility that unobservable characteristics drive the matching between firms and software suppliers and affect the number of cyberattacks independently from software vulnerabilities. If that was the case, the timing of a vulnerability discovery would not matter, as the firm-supplier pair would have a higher number of cyber attack unconditionally. We employ a standard event-time setting where we estimate the model:

$$CyberAttacks_i = \alpha_{i,e} + \alpha_{t,e} + \beta\, Vulnerability_i + \gamma X_i + \epsilon_i \tag{3}$$

To construct the event-time sample we proceed as follows. For each event in our sample, we construct an event-specific cohort of firms that includes affected firms and a comparison group of firms. We define the group of affected firms as those hit by a newly discovered vulnerability in the event quarter. We construct the comparison group of firms by selecting all the firms covered by the Factset Revere dataset that are never affected by *any* vulnerability. The main independent variable, $Vulnerability_i$, is an indicator with value 1 if the firm is affected by a vulnerability in the event quarter 0. For each event we observe firms outcomes from 4 quarters before the event to 4 quarters after. Following the stacked DiD methodology described in Baker, Larcker, and Wang (2022), we include $\alpha_{i,e}$ and $\alpha_{t,e}$, which indicate firm-cohort and time-cohort fixed effects, respectively.

To perform the second falsification test, we adopt the block permutation procedure following Chetty, Looney, and Kroft (2009). In each iteration, the date of the *Vulnerability* variable is randomly re-assigned by firm with replacement as a placebo through the sample period. Equation (3) is then estimated on the falsified data. The top plot in Figure IA1 reports the empirical cumulative distribution function (cdf) generated from running the regression model of Column (2) of Table IA5, Panel A in 1,000 random iterations of this procedure and capturing the placebo coefficient estimate. The vertical dotted line indicates the position of the actual coefficient estimate for the impact that a vulnerability discovery has on the outcome variables and implied p-value when placed in the context of cdf. Our coefficient has a p-value of 0.000, indicating that the timing of the vulnerability is essential to determine the effect on the number of cyberattacks that we document. A similar conclusion can be drawn from the bottom plot of Figure IA1, which is

generated from the regression model of Column (2) of Table IA5, Panel B. We conclude from the falsification tests that unobserved factors are unlikely to drive the link between SVs and future cyberattacks.

4.2 Customer Risk

In this section we study the impact of software vulnerabilities on four popular measures of firm risk based on stock returns. To be consistent with the tests on the predictability of cyber attacks, we calculate all measures at the quarterly horizon. The first two are standard measures of stock return volatility. Specifically, we use return volatility, defined as the standard deviation of daily stock returns; and idiosyncratic volatility, defined as the standard deviation of residuals from a regression of daily stock returns on the CAPM model.[11] The next two measures are instead meant to capture the downside risk of a firm, i.e. capture the probability of extreme negative events. To that end, we use: Value-at-risk (VaR), defined as (the absolute value of) the worst 5% daily return; and lower partial moment of the second order (LPM), defined as the standard deviation of negative daily stock returns.

We re-estimate equation (1) by substituting $CyberAttacks_{c,t+1}$ with each of the four proxies described. The results are displayed in Table 4, Panel A. For all measures, the discovery of an additional vulnerability in the software used by customer c in quarter t leads to an increase in firm risk, regardless of the variable we use. The economic magnitude of the effect is meaningful: we observe an increase in return volatility of 0.046 standard deviations; an increase of 0.035 standard deviations in idiosyncratic volatility; an drop of 0.062 standard deviations in value-at-risk; and an increase of 0.053 standard deviations in lower partial moment. The size of the effects is economically meaningful and comparable to effects documented in prior research, for example the impact of political risk on firm stock volatility documented in Hassan et al. (2019), Table IV.

In Panel B of Table 4 we re-estimate equation (2), replacing the dependent variable $CyberAttacks_{s,t+1}$ with the four risk proxies aggregated at the supplier level. To aggregate the risk metrics we take the equal-weighted average across all customers of software supplier s in quarter t. The results

[11] More details on the construction of the variables can be found in the data appendix.

indicate that when a product from software supplier s is found to be affected by a vulnerability, all customers of the software company, on average, experience a significant increase in both firm risk and downside risk. The magnitude of the effect is quite consistent across measures and an ranges between 0.011 and 0.013 standard deviations.

Overall, the results in this section suggest that the increased probability of cyberattacks following the discovery of SVs translates into a rise in stock market-based measures of firm risk, observable at the quarterly horizon.

4.3 The Real Effects of SVs

In this section, we examine the impact of SVs on firm-level investments and sales, as well as how SVs correlate with actions taken by the firm. Both theoretical and empirical literature suggest that an increase in any kind of risk tends to decrease firm-level investments (see, e.g., Bloom, Bond, and Van Reenen (2007), Bloom (2009), Hassan et al. (2019)). In turn, reduced investments may lead to lower firm revenues. Additionally, to understand how firms react to the surge in cyber risk due to SVs, a logical starting point is to investigate whether they engage the services of specialized cybersecurity companies.

Table 5 tests these predictions. In Panel A the setting is the same as the one described in equation 1. In the first column we use as dependent variable the capital investment rate I_{t+1}/K_t, measured quarterly and calculated recursively as detailed in Stein and Stone (2013). The coefficient indicates that exposure to an additional SV decreases firm investments by 0.044 standard deviations. In the second column we study the R&D investment rate, again computed recursively as in Stein and Stone (2013). Results show that the discovery of a vulnerability also depresses intangible investments, with a magnitude that is very similar to the effect on the tangible ones. The third column focuses on sales growth. We observe a significant reduction in the growth rate of sales in response to the emergence of a SV. The economic magnitude indicates that an additional vulnerability leads to a drop equal to 0.037 standard deviations.

Finally, in the last column of Panel A of Table 5, we use as dependent variable an indicator that takes a value of one if the firm hires for the first time a cybersecurity firm in the quarter

following the discovery of a vulnerability. As we discuss in Section 2 the intricate nature of firms' digital supply chains and the specialized expertise required for patch application are critical factors in preventing firms from safeguarding against SVs. Consequently, firms may opt to employ the services of cybersecurity specialists. The results support this prediction, showing that the probability of hiring a cybersecurity company increases by 0.111 standard deviations in response to the emergence of a SV. Aside from providing an insight on how firms reach to SV exposure, this test also strenghten the idea that our measure for SV indeed captures cyber risk.

As in the previous two sections, in Panel B of Table 5 we aggregate our variables at the software level by taking the equal-weighted averages of the capital investment rate, R&D investment rate, and sales growth. Additionally, we define a new indicator variable that takes a value of one if any of supplier s' customers hires a cybersecurity firm for the first time in quarter $t+1$. The sign of the results in all columns is consistent with Panel A. Moreover, with the exception of sales growth, all columns display highly statistically significant coefficients.

Taken together, our findings indicate that SVs exposure not only is a first order driver of firms' cyber risk, but also has meaningful economic implications, affecting firms' real activities.

4.4 How are SVs Discoveries Incorporated in Market Prices?

Given the results in 4.2, which indicate that firm risk increases in the quarter following the discovery of a software vulnerability, an interesting question that arises is when the market incorporates the negative news about the vulnerability discovery into stock prices. Answering this question is the aim of this section. This issue is especially relevant because cyber risk is a recent phenomenon, making its accurate measurement inherently challenging. Underscoring the importance of our question, Gomez Cram and Lawrence (2024) provide evidence that market participants have consistently undervalued software companies over the past 20 years. This suggests that the financial implications of SV-related cyber risks may not be fully appreciated by the market.

We begin by studying customer firms' stock price reaction in a short window around the discovery of a SV. We employ a standard event-study setting, and the sample construction is similar to the one we employ to estimate Equation (3). Specifically, to construct the event-time

sample, we proceed as follows. For each event in our sample, we create an event-specific cohort of firms that includes both affected firms and a comparison group. The affected firms are defined as those impacted by a newly discovered SV in the event quarter. The comparison group consists of firms covered by the Factset Revere dataset that have never been affected by any SV. To measure the short term reaction of stock returns around the announcement of a SV discovery, we first calculate the 3-day cumulative abnormal returns (CAR) in a window starting one day before and ending one day after the event. The estimation period of the expected return includes 220 trading days from day 31 to day 251 before the event, and we require a minimum of 63 days to include the event in the sample. Next, we estimate the model:

$$CAR_c = \alpha_t + \beta\, Vulnerability_c + \gamma X_c + \epsilon_c \tag{4}$$

the dependent variables are the market-adjusted, CAPM, and FF5-factor risk-adjusted cumulative abnormal returns (CAR[-1,1]) surrounding an event affecting customer firm c. The set of controls X_c includes market capitalization, book-to-market, ROA, previous 12-month return, leverage. We also include the number of cybersecurity incidents that the software customer c experienced over the year leading up to each event. α_t indicates time fixed effects. Standard errors are clustered at the customer level.

We present the results in Table 6, Panel A. Regardless of how we define CAR, we do not observe a statistically significant effect on abnormal returns around the discovery of a SV in the supplier's software. This result suggests that the market fails to promptly react to the event. There are three different interpretations of this evidence. First, the market may not react because the vulnerability is inconsequential to the firm. This explanation is unlikely given the results in the previous two sections, which show that the probability of cyberattacks, firm risk as well as firm investments and sales growth are affected following the emergence of a SV. Second, the market may be unaware of the consequences of SVs for the customers of software companies. This explanation is plausible, as cyber risks have only recently become relevant and consequential for corporations. However, this explanation contrasts with the effect we observe on stock market-based measures of firm risk at a quarterly frequency. The final potential explanation is that the market only slowly

incorporates the information related to the SV into prices. To test this hypothesis, we study the effect of software vulnerabilities on stock returns over a longer horizon. The underlying intuition is that if the lack of short-term market reaction is due to slow information incorporation, we should nonetheless observe a market reaction at some point after the discovery of a SV.

Following this intuition, we employ again the setting described by Equation (1), replacing $CyberAttacks_{c,t+1}$ with measures of customer c risk-adjusted stock returns in quarter $t+1$. We estimate the model:

$$R_{c,t+1} = \alpha_t + \alpha_c + \beta_1 Vulnerability_{c,t} + \gamma X_{c,t} + \epsilon_{c,t+1} \tag{5}$$

we use three different measures of risk-adjusted stock return: the stock's market-adjusted return, calculated as the difference between the stock's return and the market return in quarter $t+1$; CAPM alpha, calculated as the intercept in a CAPM regression of daily excess stock returns on daily excess market returns in quarter $t+1$; five-factor alpha, calculated as the intercept in a five-factor regression of daily excess stock returns on the five Fama and French (2016) factors in quarter $t+1$. $X_{c,t}$ is a set of controls that includes market capitalization, book-to-market, ROA, previous 12-month return, leverage, and the number of cybersecurity incidents that the software customer c experienced over the previous year. α_t and α_c denote time and customer fixed effects, respectively. Standard errors are clustered at the customer level.

Results are reported in Table 6, Panel B. The coefficients indicate a negative and significant stock market reaction to the discovery of software vulnerabilities over the following quarter. In terms of economic significance, the magnitude of the effect ranges between -0.6% and -0.8%, depending on the risk adjustment method used. We conclude that the evidence supports the notion that market participants are slow to incorporate information about cyber risks originating from the discovery of SVs.

4.4.1 Slow Reaction to Supply Chain Linkages

In this section, we explore the frictions that may drive the slow incorporation of information about software vulnerabilities. To this end, we draw from the prior work of Cohen and Frazzini

(2008), who document that investors do not promptly incorporate news about economically related firms, leading to return predictability. In our setting, we conjecture that market participants fail to immediately incorporate the consequences of software vulnerabilities for the customer firms because they overlook the economic links between the supplier of the vulnerable software product and its customers.

To assess this conjecture, we perform two tests where we measure the price reaction to events relevant for cyber risk where market participants do not need to consider customer–supplier links. First, we study the software company's stock price reaction to the discovery of a SV on one of its products. The emergence of a SV should prompt a negative stock market reaction for the software company for several reasons. To start, it damages the company's reputation, eroding customer trust and potentially leading to customer loss. Also, financially, the firm incurs significant costs for patch development, customer support, and potential legal liabilities. Finally, regulatory scrutiny and potential fines may follow, and competitors might exploit the situation to gain market share. At the same time, the link between the event and the affected firm is direct, as the vulnerability is directly associated with its producer.

We construct an event-time sample that consists of the sample of software companies affected by vulnerabilities in their products, as well as a comparison group made of the other firms in our software companies sample. In this setting, we estimate the model:

$$CAR_s = \alpha_t + \beta\, Vulnerability_s + \gamma X_s + \epsilon_s \tag{6}$$

the dependent variables are the market-adjusted, CAPM, and FF5-factor risk-adjusted cumulative abnormal returns (CAR[-1,1]) surrounding an event affecting a product of supplier firm s. The set of controls X_s includes market capitalization, book-to-market, ROA, previous 12-month return, leverage. We also include the number of cybersecurity incidents that customers of the software supplier s experienced over the year leading up to each event. α_t indicates time fixed effects. Standard errors are clustered at the supplier level.

We present the results in Table 7, Panel A. No matter the way we define CAR, we always see that the announcement of a vulnerability in the supplier's software decreases the company's

cumulative abnormal returns. The effect is sizable and ranges between -0.22% and -0.27% over the 3-day window, depending on the risk adjustment. The market appears to be able to promptly process the negative consequences of a SV discovery for the software company.

The second test focuses on measuring the market reaction to the occurrence of a cyberattack caused by a SV that impacts one of the customer firms in our sample. This attack represents the realization of the cybersecurity risk that market participants initially fail to consider at the moment of the vulnerability discovery. However, upon the occurrence of an attack, investors do not need to consider customer-supplier links, as the affected firm is directly impacted. The event-time sample we use to test market reaction in this setting includes, as affected firms, the customer firms victim of a cyberattack caused by a software vulnerability. The comparison group consists of firms in our sample that have never been victims of a cyberattack. We estimate the model:

$$CAR_c = \alpha_t + \beta Attack_c + \gamma X_c + \epsilon_c \tag{7}$$

the dependent variables are the market-adjusted, CAPM, and FF5-factor risk-adjusted cumulative abnormal returns (CAR[-1,1]) surrounding a cyberattack affecting customer firm c. The set of controls X_c includes market capitalization, book-to-market, ROA, previous 12-month return, leverage. We also include the number of cybersecurity incidents that the customer c experienced over the year leading up to each event. α_t indicates time fixed effects. Standard errors are clustered at the customer level.

We present the results in Table 7, Panel B. Across all columns we observe a strongly significant negative market reaction to the realization of a cyberattack. The economic magnitude is also significant, ranging between -1.1% and -1.3% over the 3-day window, depending on the risk adjustment. Hence, the negative consequences of a cyberattack for a company are incorporated into the price in a short 3-day window.

Overall, our results strongly indicate that market participants fail to immediately incorporate the effect of SVs on companies using the flawed software because they struggle to account for supplier-customer links in a timely manner. Consequently, they only manage to do so with a delay, resulting in the slow incorporation of information about cyber risk stemming from SVs.

5 Software Companies as a Systemic Source of Cyber Risk

The goal of this section to explore the idea that software companies can be a systemic source of cybersecurity risk for other firms in the economy. Recent events, such as the CrowdStrike global IT crash of July 2024, have stimulated the discussion about the systemic nature of software companies (Welburn (2024)). Further, in April 2024 the U.S. government has instructed the Cybersecurity and Infrastructure Security Agency (CISA) to identify and categorize certain critical infrastructure entities as Systemically Important Entities, with software companies being among those.[12] It is therefore important to quantitatively assess the relevance of software companies as systemic source of cyber risk in the economy through the spread of SVs to customer firms. To do so, we perform two separate set of tests. First, we test whether the size of the software company amplifies the effect of SVs to customer firms. Second, we explore the aggregate implications of SVs and repeat our main tests at the industry level.

5.1 The role of Software Companies' Market Share

We start by investigating whether the growth of software companies is a major driver of the recent increase in cybersecurity risk, explaining the pattern displayed in Figure 1. The motivating intuition for the tests in this section is straightforward. If the growing market share of software companies is a determinant of the surge in cybersecurity risk, our results should be more pronounced when a SV affects the product of a larger supplier. We bring this conjecture to the data using the same supplier-level panel used in equation (2). The key addition is an interaction term between $Vulnerability_{s,t}$ and a measure of supplier market share. For a given supplier s, $Supplier\ Mktshare_{s,t}$ is defined as the ratio of the cumulative market capitalization of supplier s's customers to the total market capitalization of customers in our data in quarter t.[13]

In Panel A of Table 8 we analyze the probability of cyberattacks. Column (1), which re-

[12] See here for details.

[13] In Appendix Table IA6 we present results using the customer-level panel data employed in equation (1). Here, the measure of supplier market share is defined as the average market share across all of customer c's software suppliers. While these results lead to similar conclusions, we prefer to report the supplier-level analysis in the main body of the paper. This approach allows us to interact $Vulnerability_{s,t}$ with a measure of supplier market share that is specific to each supplier s.

ports results for all types of cyberattacks, shows that the coefficient on the interaction term $Supplier\ Mktshare_{s,t} \times Vulnerability_{s,t}$ is positive and significant. The magnitude indicates that a one standard deviation increase in $Supplier\ Mktshare_{s,t}$ amplifies the effect of a vulnerability discovery on the probability that customers of supplier s are victimes of cyberattacks by 0.015 standard deviations. This effect is entirely driven by cyberattacks originating from software vulnerabilities, as displayed in column (2). In contrast, when we examine attacks unrelated to vulnerabilities, the interaction term is not statistically significant.

Panel B of Table 8 analyzes the effect on firm risk. Across all four proxies for risk, we observe a consistent pattern: the larger the market share of the software company, the greater the impact that the emergence of a SV has on stock market-based measures of firm risk. The economic magnitude is significant, as the coefficients on the interaction terms represent between 30% and 40% of the baseline effects documented in Panel B of Table 4.

In Panel C, we examine whether the real effects of SVs are influenced by the market share of the supplier. The results in this panel indicate that firm tangible and intangible investment rates, as well as sales growth, suffer more when the discovered vulnerability pertains to the product of a supplier with a large market share. Since larger software suppliers have a more pronounced impact on firm risk, this evidence supports the view that the surge in firm risk depresses investment and sales. However, software supplier market share does not play a role in increasing the likelihood that firms hire a cybersecurity specialist. It is plausible that the increase in firm risk due to a SV originating from a smaller supplier is already sufficient to lead firms to outsource their cybersecurity needs.

Finally, the role of $Supplier\ Mktshare_{s,t}$ is clearly demonstrated also in the last panel of the table, where we focus on quarterly risk-adjusted returns. In all columns, the coefficient on the interaction term $Supplier\ Mktshare_{s,t} \times Vulnerability_{s,t}$ is negative and significant. This indicates that the average negative market reaction at the quarterly frequency of customers of supplier s worsens when the vulnerability affects the product of a larger software supplier.

To sum up, the findings in this section are consistent with our conjecture that the recent growth of software companies plays an important role in the larger incidence of cyberattacks over

the same period.

5.2 Aggregate Implications at the Industry Level

If software companies were a systemic source of cyber risk, the effect of SVs should extend beyond the firm level and have aggregate implications. To empirically assess this, we examine whether the consequences of SVs are detectable at the industry level. Although we observe significant effects at the customer firm level, it is not immediately clear the impact survives when aggregated at the industry level. For instance, while some firms in an industry may be negatively impacted due to their use of vulnerable software, other firms using competitor products might benefit by gaining market share. Alternatively, the occurrence of a cyberattack on one firm using the vulnerable software could prompt their peers to enhance their cybersecurity measures, thereby reducing overall industry risk.

We move the analysis at the Fama-French 49-industry-level, and begin by studying the determinants of the number of cyberattacks that hit an industry. We do so by estimating the model:

$$CyberAttacks_{i,t+1} = \alpha_i + \beta_1 Vulnerability_{i,t} + \epsilon_{i,t+1} \tag{8}$$

where i indexes Fama-French 49 industries and t quarters. We use as dependent variables the number of cybersecurity incidents of any type, related to vulnerabilities, or not related to vulnerabilities that hit the industry i in quarter $t+1$. In all columns, the main independent variable is the number of vulnerabilities that affect industry i in quarter t. α_i indicates industry fixed effects. Standard errors are clustered at the industry level. For ease of interpretation we standardize the dependent variables.

Results reported in Table 9 Panel A align with those in Table 3 at the customer firm level. We find that the emergence of an additional vulnerability increases the number of cyberattacks of any type occurring in an industry by 0.09 standard deviations. This impact is more pronounced at 0.104 standard deviations for attacks specifically caused by SVs, while it is negligible for attacks due to other causes. Overall, these findings suggest that SVs have the potential to impact the

cyber risk exposure of entire industries.

In Panel B of Table 9 we re-estimate Equation (8) using stock market-based measures of industry risk as dependent variables. We employ return volatility, defined as the standard deviation of daily industry returns over a one-quarter horizon, and idiosyncratic volatility, defined as the standard deviation of residuals from a regression of daily industry returns on the CAPM model, also computed over a one-quarter horizon. To capture the downside risk of an industry, we use Value-at-Risk (VaR), defined as the absolute value of the worst 5% daily industry returns over a one-quarter horizon, and the lower partial moment of the second order (LPM), defined as the standard deviation of negative daily industry returns over a one-quarter horizon.[14] The results indicate that SVs increase industry-level risk across all measures, with a magnitude ranging between 0.016 and 0.037 standard deviations.

A similar pattern is displayed in Panel C and D, where we analyze industry-level real effects and risk-adjusted quarterly returns, respectively. Panel C shows that aggregated investment rates drop in response to the emergence of software vulnerabilities affecting customers operating in the industry. The number of firms hiring the services of cybersecurity company specialists also increases. The last panel of the table shows that industries experience a decline in quarterly risk-adjusted performance between 0.1% and 0.4% in response to an additional software vulnerability, depending on the risk adjustment.

Overall, the findings in this section allow us to conclude that SVs have significant aggregate implications and can substantially alter the exposure to cyber risk across entire industries.

6 Conclusion

This paper shows that the expansion of the software market is, trough SVs, a significant factor contributing to the increase in cybersecurity risk for firms in the economy. The underlying mechanism operates as follows: the proliferation of software products and the interconnectedness of the software industry lead to a greater number of potential targets for cyberattacks, and each software vulnerability represents a potential entry point for malicious actors. Consequently,

[14] Industry daily returns are sourced from Kenneth French's website.

the expansion of the software industry inadvertently amplifies the overall exposure to cyber risk landscape by introducing more points of vulnerability into the corporate sector. This interconnectedness means that as more firms integrate software into their operations, the broader economy becomes increasingly vulnerable to cyber threats.

We see our paper as first step in understanding how cybersecurity risk originates and propagates, even before an actual cyberattack occurs. Our work extends and complements recent studies on the effect of cyberattacks on firms and financial institutions. Importantly, we show that the expansion and concentration of software companies in the economy can be a significant driver of such risk, thus contributing on the recent debate on the role of software companies as systemic source of cyber risk.

References

Acemoglu, Daron, and Pascual Restrepo, 2020, Robots and jobs: Evidence from us labor markets, *Journal of Political Economy* 128, 2188–2244.

Adelino, Manuel, Miguel A Ferreira, Mariassunta Giannetti, and Pedro Pires, 2023, Trade credit and the transmission of unconventional monetary policy, *The Review of Financial Studies* 36, 775–813.

Akerman, Anders, Ingvil Gaarder, and Magne Mogstad, 2015, The Skill Complementarity of Broadband Internet *, *The Quarterly Journal of Economics* 130, 1781–1824.

August, Terrence, Marius Florin Niculescu, and Hyoduk Shin, 2014, Cloud implications on software network structure and security risks, *Information Systems Research* 25, 489–510.

August, Terrence, Daehoon Noh, Noam Shamir, and Hyoduk Shin, 2022, Cyberattacks, operational disruption and investment in resilience measures, *Available at SSRN 4032257* .

Autor, David H, Frank Levy, and Richard J Murnane, 2003, Skill demand, inequality, and computerization: Connecting the dots, *Technology, growth, and the labor market* 107–129.

Baker, Andrew C, David F Larcker, and Charles CY Wang, 2022, How much should we trust staggered difference-in-differences estimates?, *Journal of Financial Economics* 144, 370–395.

Barrot, Jean-Noël, and Julien Sauvagnat, 2016, Input Specificity and the Propagation of Idiosyncratic Shocks in Production Networks *, *The Quarterly Journal of Economics* 131, 1543–1592.

Bloom, Nicholas, 2009, The impact of uncertainty shocks, *Econometrica* 77, 623–685.

Bloom, Nicholas, Tarek Alexander Hassan, Aakash Kalyani, Josh Lerner, and Ahmed Tahoun, 2021, The diffusion of new technologies, Working Paper 28999, National Bureau of Economic Research.

Bloom, Nick, Stephen Bond, and John Van Reenen, 2007, Uncertainty and investment dynamics, *Review of Economic Studies* 74, 391–415.

Carvalho, Vasco M, Makoto Nirei, Yukiko U Saito, and Alireza Tahbaz-Salehi, 2021, Supply chain disruptions: Evidence from the great east japan earthquake, *The Quarterly Journal of Economics* 136, 1255–1321.

Chetty, Raj, Adam Looney, and Kory Kroft, 2009, Salience and taxation: Theory and evidence, *American economic review* 99, 1145–1177.

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

Crosignani, Matteo, Marco Macchiavelli, and André F. Silva, 2023, Pirates without borders: The propagation of cyberattacks through firms' supply chains, *Journal of Financial Economics* 147, 432–448.

Curti, Filippo, Ivan Ivanov, Marco Macchiavelli, and Tom Zimmemrmann, 2023, City hall has been hacked! the financial costs of lax cybersecurity.

Cybersecurity Ventures, 2022, 2022 Official Cybercrime Report https://s3.ca-central-1.amazonaws.com/esentire-dot-com-assets/assets/resourcefiles/2022-Official-Cybercrime-Report.pdf.

DellaVigna, Stefano, and Joshua M. Pollet, 2007, Demographics and industry returns, *American Economic Review* 97, 1667–1702.

Duffie, Darrell, and Joshua Younger, 2019, Cyber runs.

Eisenbach, Thomas M., Anna Kovner, and Michael Junho Lee, 2022, Cyber risk and the u.s. financial system: A pre-mortem analysis, *Journal of Financial Economics* 145, 802–826.

Eisenbach, Thomas M., Anna Kovner, and Michael Junho Lee, 2023, When it rains, it pours: Cyber risk and financial conditions.

Fama, Eugene F, and Kenneth R French, 2016, Dissecting anomalies with a five-factor model, *Review of Financial Studies* 29, 69–103.

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

Gomez Cram, Roberto, and Alastair Lawrence, 2024, The value of software, *Working paper, London Business School* .

Gormley, Todd A., and David A. Matsa, 2011, Growing out of trouble? corporate responses to liability risk, *Review of Financial Studies* 24, 2781–2821.

Hassan, Tarek A, Stephan Hollander, Laurence van Lent, and Ahmed Tahoun, 2019, Firm-Level Political Risk: Measurement and Effects*, *The Quarterly Journal of Economics* 134, 2135–2202.

Hertzel, Michael G, Zhi Li, Micah S Officer, and Kimberly J Rodgers, 2008, Inter-firm linkages and the wealth effects of financial distress along the supply chain, *Journal of Financial Economics* 87, 374–387.

Hong, Harrison, and Jeremy C. Stein, 1999, A unified theory of underreaction, momentum trading, and overreaction in asset markets, *The Journal of Finance* 54, 2143–2184.

Hong, Harrison, Walter Torous, and Rossen Valkanov, 2007, Do industries lead stock markets?, *Journal of Financial Economics* 83, 367–396.

Jamilov, Rustam, Hélène Rey, and Ahmed Tahoun, 2023, The anatomy of cyber risk, Working Paper 28906, National Bureau of Economic Research.

Kogan, Leonid, Dimitris Papanikolaou, Amit Seru, and Noah Stoffman, 2017, Technological Innovation, Resource Allocation, and Growth*, *The Quarterly Journal of Economics* 132, 665–712.

Kotidis, Antonis, and Stacey L. Schreft, 2022, Cyberattacks and financial stability: Evidence from a natural experiment.

Merton, Robert C., 1987, A simple model of capital market equilibrium with incomplete information, *The Journal of Finance* 42, 483–510.

Peng, Lin, and Wei Xiong, 2006, Investor attention, overconfidence and category learning, *Journal of Financial Economics* 80, 563–602.

Pástor, Ľuboš, and Pietro Veronesi, 2009, Technological revolutions and stock prices, *American Economic Review* 99, 1451–83.

Software.org, 2021, Software: Supporting US Through COVID https://software.org/reports/software-supporting-us-through-covid-2021/.

Stein, Luke CD, and Elizabeth Stone, 2013, The effect of uncertainty on investment, hiring, and r&d: Causal evidence from equity options, *Working Paper (SSRN 1649108)* .

Ward, Colin, 2020, Is the it revolution over? an asset pricing view, *Journal of Monetary Economics* 114, 283–316.

Welburn, Jonathan, 2024, Crowdstrike is too big to fail, *Wall Street Journal* https://www.wsj.com/articles/crowdstrike-is-too-big-to-fail-77d20f1d.

Figure 1: Trends in Software Companies and Cyber Risk

This figure shows the yearly market share of U.S. software companies and number of cyberattacks affecting U.S. publicly listed companies. Market share is calculated as the fraction of the total stock market accounted for by software companies' market capitalization.

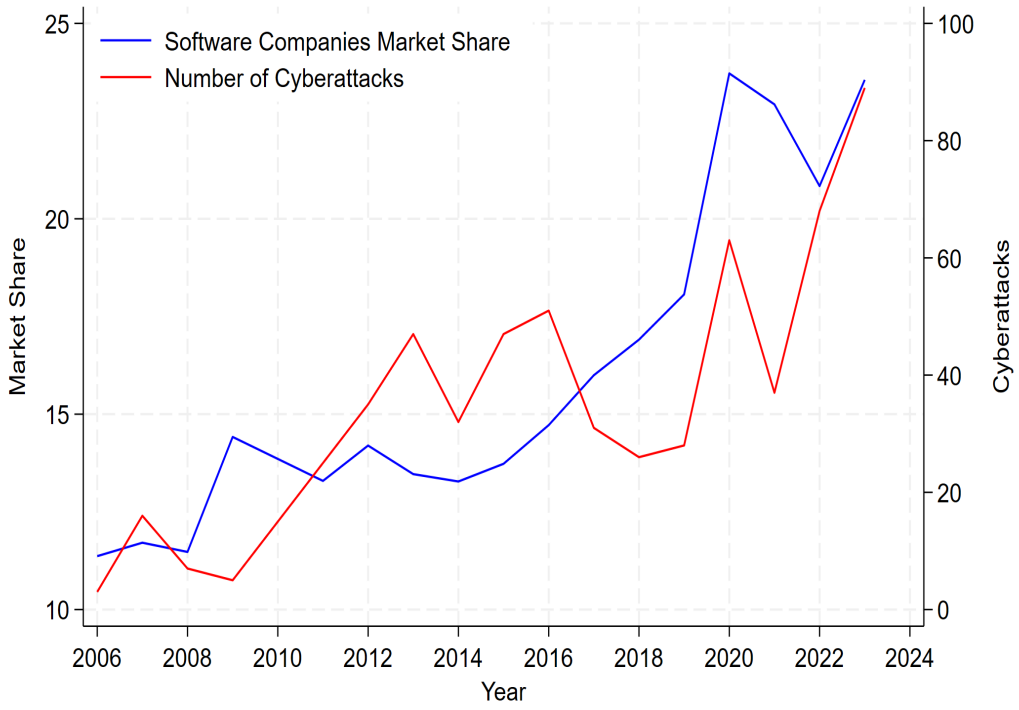

Table 1: Summary Statistics

This table presents mean, standard deviation, 10th percentile, median, 90th percentile and number of observations for the main variables used in the paper. A complete list of definitions for these variables is provided in the Data appendix. Panel A presents summary statistics for the software supplier sample, while in Panel B we present statistics for the software customers' sample. The sample period runs from 2006 to 2023.

Panel A: Customer-Level Statistics

	Mean	SD	P10	P50	P90	Observations
Vulnerability	1.20	4.53	0.00	0.00	3.00	81,125
CyberAttacks	0.00	0.07	0.00	0.00	0.00	81,125
Size ($ B)	18.62	40.05	0.13	3.42	50.17	81,125
Book-to-Market	0.59	0.55	0.11	0.44	1.17	81,125
Profitability	0.00	0.04	-0.03	0.01	0.03	81,125
Leverage	0.23	0.20	0.00	0.21	0.51	81,125
Past 12-month Return	0.10	0.48	-0.43	0.06	0.62	81,125
Supplier MktShare	10.74	7.89	1.60	9.13	23.16	81,125

Panel B: Supplier-Level Statistics

	Mean	SD	P10	P50	P90	Observations
Vulnerability	0.04	0.83	0.00	0.00	0.00	26,627
CyberAttacks	20.86	61.20	0.00	0.00	100.00	26,627
Size ($ B)	15.31	94.70	0.04	0.92	19.61	26,627
Book-to-Market	0.51	0.82	0.08	0.36	1.07	26,627
Profitability	-0.01	0.09	-0.06	0.00	0.04	26,627
Leverage	0.15	0.18	0.00	0.08	0.41	26,627
Past 12-month Return	0.14	0.76	-0.51	0.04	0.77	26,627
Supplier MktShare	3.56	4.21	0.06	2.13	8.77	26,627

Table 2: Firm-Level Determinants of SVs

This table studies firm-level determinants of SVs. The dependent variable in Panel A is an indicator that takes value 1 if customer c is exposed to a vulnerability in one of the software products it uses in quarter t. The dependent variable in Panel B is an indicator with value 1 if one of the software products of supplier s has a vulnerability in quarter t. In both panels, all right-hand-side variables are measured with one quarter lag. A complete list of definitions for these variables is provided in the Data appendix. t-statistics based on standard errors clustered by customer (or supplier) are shown in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. The sample period runs from 2006 to 2023.

Panel A: Customer Vulnerability Determinants

	Vulnerability$_{c,t}$				
	(1)	(2)	(3)	(4)	(5)
Size$_{c,t-1}$	0.004	-0.035***	0.021	0.009	0.009
	(0.30)	(-2.91)	(0.54)	(0.24)	(0.23)
Book-to-Market$_{c,t-1}$	-0.053***	-0.033***	0.006	0.005	0.005
	(-3.86)	(-2.95)	(0.85)	(0.76)	(0.76)
Profitability$_{c,t-1}$	-0.049***	-0.018*	-0.002	-0.003	-0.003
	(-3.18)	(-1.82)	(-0.44)	(-0.53)	(-0.52)
Leverage$_{c,t-1}$	0.010	-0.035***	0.044**	0.038**	0.038**
	(0.89)	(-3.29)	(2.33)	(2.19)	(2.19)
Past 12-month Return$_{c,t-1}$	-0.000	-0.006	-0.006	-0.005	-0.005
	(-0.01)	(-0.50)	(-0.63)	(-0.46)	(-0.47)
Supplier MktShare$_{c,t-1}$				2.150***	2.147***
				(5.26)	(5.26)
Past CyberAttacks$_{c,t-1}$					0.010
					(1.37)
Customer FE		Yes	Yes	Yes	Yes
Time FE			Yes	Yes	Yes
Observations	81,125	81,125	81,125	81,125	81,125
Adjusted r^2	0.005	0.134	0.402	0.412	0.412

Panel B: Supplier Vulnerability Determinants

	Vulnerability$_{s,t}$				
	(1)	(2)	(3)	(4)	(5)
Size$_{s,t-1}$	0.107*	0.101*	0.057	0.052	0.050
	(1.88)	(1.82)	(1.12)	(1.06)	(1.04)
Book-to-Market$_{s,t-1}$	-0.005	-0.007	-0.040*	-0.042*	-0.041*
	(-0.43)	(-0.63)	(-1.72)	(-1.76)	(-1.75)
Profitability$_{s,t-1}$	-0.006	-0.003	0.002	0.003	0.003
	(-1.14)	(-0.60)	(1.14)	(1.28)	(1.39)
Leverage$_{s,t-1}$	-0.006	-0.012	-0.005	-0.006	-0.007
	(-0.79)	(-1.16)	(-0.40)	(-0.43)	(-0.51)
Past 12-month Return$_{s,t-1}$	-0.018	-0.013*	-0.011	-0.011	-0.011
	(-1.47)	(-1.77)	(-1.43)	(-1.45)	(-1.48)
Supplier MktShare$_{s,t-1}$				0.032	0.010
				(0.99)	(0.56)
Past CyberAttacks$_{s,t-1}$					0.050
					(1.34)
Supplier FE		Yes	Yes	Yes	Yes
Time FE			Yes	Yes	Yes
Observations	26,627	26,627	26,627	26,627	26,627
Adjusted r^2	0.011	0.017	0.147	0.147	0.149

Table 3: Vulnerabilities and Cybersecurity Incidents

This table studies the relation between the probability of cybersecurity attacks and exposure to SVs. In both panels, we report estimates of the model:

$$CyberAttacks_{i,t+1} = \alpha + \beta_1 Vulnerability_{i,t} + \gamma X_{i,t} + \epsilon_{i,t+1}$$

the dependent variables in Panel A are the number of cybersecurity incidents of any type (columns (1)), related to vulnerabilities (columns (2)), or due to causes different from vulnerabilities (columns (3)) that hit the software customer c in quarter $t+1$. In all columns, the main independent variable is an indicator with value 1 if customer c is exposed to a vulnerability of one of its software suppliers in quarter t. The dependent variables in Panel B are the number of cybersecurity incidents of any type (columns (1)), related to vulnerabilities (columns (2)), or due to causes different from vulnerabilities (columns (3)) that hit the software supplier s customers in quarter $t+1$. In all columns, the main independent variable is an indicator with value 1 if one of the software products of software company s has a vulnerability in quarter t. Controls include market capitalization, book-to-market, ROA, previous 12-month return, leverage. In Panel A we also include the number of cybersecurity incidents that the software customer c experienced over the previous year. In Panel B we also include the number of cybersecurity incidents that the software supplier s customers experienced over the previous year. α indicates different sets of fixed effects, including year-quarter and software customer (or supplier) fixed effects. t-statistics based on standard errors clustered by customer (or supplier) are shown in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. A complete list of definitions for these variables is provided in the Data appendix.

Panel A: Customer Attacks

	All Attacks$_{c,t+1}$ (1)	Vulnerability Attacks$_{c,t+1}$ (2)	Other Attacks$_{c,t+1}$ (3)
Vulnerability$_{c,t}$	0.106***	0.187***	0.029
	(4.41)	(5.75)	(1.43)
Customer Controls$_{c,t}$	Yes	Yes	Yes
Customer FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Observations	81,125	81,125	81,125
Adjusted r^2	0.016	0.004	0.013

Panel B: Customer Attacks Aggregated at the Supplier-Level

	All Attacks$_{s,t+1}$ (1)	Vulnerability Attacks$_{s,t+1}$ (2)	Other Attacks$_{s,t+1}$ (3)
Vulnerability$_{s,t}$	0.036**	0.038**	0.012
	(2.24)	(2.14)	(0.95)
Supplier Controls$_{s,t}$	Yes	Yes	Yes
Supplier FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Observations	26,627	26,627	26,627
Adjusted r^2	0.326	0.173	0.254

Table 4: Vulnerabilities and Customer Risk

This table studies the relation between the firm risk and exposure to SVs. In both panels, we report estimates of the model:

$$Risk_{i,t+1} = \alpha + \beta_1 Vulnerability_{i,t} + \gamma X_{i,t} + \epsilon_{i,t+1}$$

the dependent variables in Panel A are the return volatility, idiosyncratic volatility, value-at-risk (VaR), and the second-order LPM of software customer c in quarter $t+1$. In Panel A, the main independent variable is an indicator with value 1 if customer c is exposed to a vulnerability of one of its software suppliers in quarter t. The dependent variables in Panel B are the equal-weighted averages of the return volatility, idiosyncratic volatility, value-at-risk (VaR), and the second-order LPM of software supplier s customers in quarter $t+1$. In Panel B, the main independent variable is an indicator with value 1 if one of the software products of software company s has a vulnerability in quarter t. Controls include market capitalization, book-to-market, ROA, previous 12-month return, leverage. In Panel A we also include the number of cybersecurity incidents that the software customer c experienced over the previous year. In Panel B we also include the number of cybersecurity incidents that the software supplier s customers experienced over the previous year. α indicates different sets of fixed effects, including year-quarter and software customer (or supplier) fixed effects. t-statistics based on standard errors clustered by customer (or supplier) are shown in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. A complete list of definitions for these variables is provided in the Data appendix.

Panel A: Customer Risk

	$\text{RetVol}_{c,t+1}$ (1)	$\text{Ivol}_{c,t+1}$ (2)	$\text{Var 95\%}_{c,t+1}$ (3)	$\text{LPSD}_{c,t+1}$ (4)
$\text{Vulnerability}_{c,t}$	0.046***	0.035***	-0.062***	0.053***
	(3.39)	(2.62)	(-4.66)	(3.87)
$\text{Customer Controls}_{c,t}$	Yes	Yes	Yes	Yes
Customer FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Observations	81,125	81,125	81,125	81,125
Adjusted r^2	0.560	0.577	0.541	0.438

Panel B: Customer Risk Aggregated at the Supplier-Level

	$\text{RetVol}_{s,t+1}$ (1)	$\text{Ivol}_{s,t+1}$ (2)	$\text{Var 95\%}_{s,t+1}$ (3)	$\text{LPSD}_{s,t+1}$ (4)
$\text{Vulnerability}_{s,t}$	0.011***	0.011***	-0.011***	0.013***
	(3.47)	(3.74)	(-2.87)	(3.89)
$\text{Supplier Controls}_{s,t}$	Yes	Yes	Yes	Yes
Supplier FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Observations	26,627	26,627	26,627	26,627
Adjusted r^2	0.528	0.535	0.505	0.459

Table 5: Real Effects of Software Vulnerabilities

This table studies the real effects of customers' exposure to SVs. In both panels, we report estimates of the model:

$$Y_{i,t+1} = \alpha + \beta_1 \, Vulnerability_{i,t} + \gamma X_{i,t} + \epsilon_{i,t+1}$$

the dependent variables in Panel A are firm c's capital investment rate $\mathrm{I}_{c,t+1}/\mathrm{K}_{c,t}$, R&D rate $\mathrm{R}_{c,t+1}/\mathrm{G}_{c,t}$, sales growth, and an indicator variable that takes value of 1 if customer c has a relation with a cybersecurity company in quarter $t+1$. In Panel A, the main independent variable is an indicator with value 1 if customer c is exposed to a vulnerability from one of its software suppliers in quarter t. The dependent variables in the first three columns of Panel B are the equal-weighted average capital investment rate, R&D rate, and sale growth across all customers of supplier s. In the last column the dependent variable is an indicator variable that takes value of 1 if any customer of software company s has a relation with a cybersecurity company in quarter $t+1$. In Panel B, the main independent variable is an indicator with value 1 if one if the software of supplier s has a vulnerability in quarter t. Controls include market capitalization, book-to-market, ROA, previous 12-month return, leverage. In Panel A we also include the number of cybersecurity incidents that the software customer c experienced over the previous year. In Panel B we also include the number of cybersecurity incidents that the software supplier s customers experienced over the previous year. α indicates different sets of fixed effects, including year-quarter and software customer (or supplier) fixed effects. t-statistics based on standard errors clustered by customer (or supplier) are shown in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. A complete list of definitions for these variables is provided in the Data appendix.

Panel A: Customer Real Effects

	$\mathrm{I}_{c,t+1}/\mathrm{K}_{c,t}$ (1)	$\mathrm{R}_{c,t+1}/\mathrm{G}_{c,t}$ (2)	Sale Growth$_{c,t+1}$ (3)	Cybersecurity Relation$_{c,t+1}$ (4)
Vulnerability$_{c,t}$	-0.044**	-0.043***	-0.037***	0.111***
	(-2.38)	(-2.68)	(-2.89)	(3.36)
Customer Controls$_{c,t}$	Yes	Yes	Yes	Yes
Customer FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Observations	70,069	70,069	70,069	70,069
Adjusted r^2	0.202	0.756	0.025	0.652

Panel B: Customer Real Effects Aggregated at the Supplier-Level

	$\mathrm{I}_{s,t+1}/\mathrm{K}_{s,t}$ (1)	$\mathrm{R}_{s,t+1}/\mathrm{G}_{s,t}$ (2)	Sales Growth$_{s,t+1}$ (3)	Cybersecurity Relation$_{s,t+1}$ (4)
Vulnerability$_{s,t}$	-0.011***	-0.007***	-0.002	0.011***
	(-2.90)	(-3.64)	(-1.00)	(3.71)
Supplier Controls$_{s,t}$	Yes	Yes	Yes	Yes
Supplier FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Observations	24,448	24,448	24,448	24,448
Adjusted r^2	0.134	0.397	0.003	0.492

Table 6: Vulnerability and Customer Stock Market Reaction

This table studies short-term stock market performance of firms around the discovery of a vulnerability. In Panel A we report estimates of the model:

$$CAR_c = \alpha_t + \beta\, Vulnerability_c + \gamma X_c + \epsilon_c$$

the dependent variables are the market-adjusted, CAPM, and FF5-factor risk-adjusted cumulative abnormal returns (CAR[-1,1]) surrounding an event. To calculate CAR[-1,1], we utilize an estimation window of 220 trading days (-251, -31), with day 0 being the date of the vulnerability discovery. A minimum of 63 non-missing returns is required within this estimation window. In Panel B we report estimates of the model:

$$R_{c,t+1} = \alpha_t + \alpha_c + \beta_1\, Vulnerability_{c,t} + \gamma X_{c,t} + \epsilon_{c,t+1}$$

the dependent variables are the the market-adjusted, CAPM, and FF5-factor risk-adjusted quarterly returns of software customer c in quarter $t+1$. The main independent variable is an indicator with value 1 if customer c is exposed to a vulnerability of one of its software suppliers in quarter t. Controls include market capitalization, book-to-market, ROA, previous 12-month return, leverage. We also include the number of cybersecurity incidents that the software customer c experienced over the year leading up to each event. α_t indicates time fixed effects, while α_c indicates customer fixed effects. t-statistics based on standard errors clustered by customer are shown in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. A complete list of definitions for these variables is provided in the Data appendix.

Panel A: Customer CAR[-1,1]

	Market-Adjusted CAR_c (1)	CAPM CAR_c (2)	5-factor CAR_c (3)
$Vulnerability_c$	-0.008	0.001	-0.011
	(-0.23)	(0.03)	(-0.43)
$Customer\ Controls_c$	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Observations	645,846	645,846	645,846
Adjusted r^2	0.039	0.042	0.015

Panel B: Customer Quarterly Returns

	Market-Adjusted$_{c,t+1}$ (1)	CAPM$_{c,t+1}$ (2)	5-factor Model$_{c,t+1}$ (3)
$Vulnerability_{c,t}$	-0.008**	-0.007***	-0.006**
	(-2.45)	(-2.86)	(-2.08)
$Customer\ Controls_{c,t}$	Yes	Yes	Yes
Customer FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Observations	77,357	77,357	77,357
Adjusted r^2	0.062	0.062	0.040

Table 7: Stock Market Reaction Without Supply-Chain Linkages

In Panel A we study the short-term stock market performance of supplier firms around the discovery of a vulnerability. In Panel A we report estimates of the model:

$$CAR_s = \alpha_t + \beta\, Vulnerability_s + \gamma X_s + \epsilon_s$$

the dependent variables are the market-adjusted, CAPM, and FF5-factor risk-adjusted cumulative abnormal returns (CAR[-1,1]) surrounding the discovery of a vulnerability in one of the products of software supplier s. In Panel B, we report estimates of the model:

$$CAR_c = \alpha_t + \beta Attack_c + \gamma X_c + \epsilon_c$$

the dependent variables are the market-adjusted, CAPM, and FF5-factor risk-adjusted cumulative abnormal returns (CAR[-1,1]) measured around a successful cyberattack against customer c that occurs at event time 0. To calculate CAR[-1,1], we utilize an estimation window of 220 trading days (-251, -31), with day 0 being the date of the vulnerability discovery. A minimum of 63 non-missing returns is required within this estimation window. α_t indicates time fixed effects. t-statistics based on standard errors clustered by supplier (or customer) are shown in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. A complete list of definitions for these variables is provided in the Data appendix.

Panel A: Supplier CAR[-1,1] Around Vulnerability Discovery

	Market-Adjusted CAR_s (1)	CAPM CAR_s (2)	5-factor CAR_s (3)
$Vulnerability_s$	-0.221***	-0.247***	-0.277***
	(-2.80)	(-2.73)	(-2.97)
$Supplier\ Controls_s$	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Observations	213,024	213,024	213,024
Adjusted r^2	0.012	0.012	0.012

Panel B: Customer CAR[-1,1] Around Cyber Attacks

	Market-Adjusted CAR_c (1)	CAPM CAR_c (2)	5-factor CAR_c (3)
$Attack_c$	-1.280***	-1.081***	-1.125***
	(-3.59)	(-2.97)	(-2.91)
$Customer\ Controls_c$	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Observations	49,407	49,407	49,407
Adjusted r^2	0.016	0.025	0.010

Table 8: The Role of Supplier Market Share

This table studies how supplier market share affects the relation between the probability of cybersecurity attacks, risk variables, quarterly stock returns, and exposure to software vulnerabilities. In all panels, we report estimates of the model:

$$Y_{s,t+1} = \alpha + \beta_1 \, Vulnerability_{s,t} + \beta_2 Supplier\ Mktshare_{s,t} + \beta_3 \, Vulnerability_{s,t} \times Supplier\ Mktshare_{s,t} + \gamma X_{s,t} + \epsilon_{s,t+1}$$

the dependent variables in Panel A are are the number of cybersecurity incidents of any type (columns (1)), related to vulnerabilities (columns (2)), or due to causes different from vulnerabilities (columns (3)) that hit the software supplier s customers in quarter $t+1$. The dependent variables in Panel B are the equal-weighted averages of the return volatility, stock market crash risk, value-at-risk (VaR), and the second-order LPM of software supplier s customers in quarter $t+1$. The dependent variables in the first three columns of Panel C are the equal-weighted average capital investment rate, R&D rate, and sale growth across all customers of supplier s. In the last column the dependent variable is an indicator variable that takes value of 1 if any customer of supplier s has a relation with a cybersecurity company in quarter $t+1$. The dependent variables in Panel D are the equal-weighted averages of the market-adjusted, CAPM, and FF5-factor risk-adjusted quarterly returns of software supplier s customers in quarter $t+1$. In all panels, the main independent variable is an indicator with value 1 if one of the software of supplier s has a vulnerability in quarter t. *Supplier Mktshare* is computed as the ratio of the sum of market capitalization of software supplier s customers to the total market capitalization of customer firms in our sample in quarter t. Controls include market capitalization, book-to-market, ROA, previous 12-month return, leverage. We also include the number of cybersecurity incidents that the software supplier s customers experienced over the previous year. α indicates different sets of fixed effects, including year-quarter and software supplier fixed effects. t-statistics based on standard errors clustered by supplier are shown in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. A complete list of definitions for these variables is provided in the Data appendix.

Panel A: Customer Attacks Aggregated at the Supplier-Level

	All Attacks$_{s,t+1}$ (1)	Vulnerability Attacks$_{s,t+1}$ (2)	Other Attacks$_{s,t+1}$ (3)
Supplier MktShare$_{s,t}$	0.165***	0.071**	0.157***
	(4.29)	(2.11)	(3.94)
Vulnerability$_{s,t}$	0.028**	0.026	0.011
	(2.36)	(1.60)	(1.04)
Supplier MktShare$_{s,t}$ × Vulnerability$_{s,t}$	0.015***	0.015***	0.003
	(2.67)	(2.93)	(0.47)
Supplier Controls$_{s,t}$	Yes	Yes	Yes
Supplier FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Observations	26,627	26,627	26,627
Adjusted r^2	0.336	0.175	0.262

Panel B: Customer Risk Aggregated at the Supplier-Level

	RetVol$_{s,t+1}$ (1)	Ivol$_{s,t+1}$ (2)	Var 95%$_{s,t+1}$ (3)	LPSD$_{s,t+1}$ (4)
Supplier MktShare$_{s,t}$	-0.075***	-0.091***	0.061***	-0.063***
	(-5.50)	(-6.09)	(4.86)	(-4.63)
Vulnerability$_{s,t}$	0.010***	0.009***	-0.009***	0.011***
	(3.83)	(3.76)	(-3.75)	(4.53)
Supplier MktShare$_{s,t}$ × Vulnerability$_{s,t}$	0.003**	0.003**	-0.004**	0.003**
	(2.19)	(1.99)	(-2.41)	(2.05)
Supplier Controls$_{s,t}$	Yes	Yes	Yes	Yes
Supplier FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Observations	26,627	26,627	26,627	26,627
Adjusted r^2	0.530	0.538	0.507	0.460

Panel C: Customer Real Effects Aggregated at the Supplier-Level

	I$_{s,t+1}$/K$_{s,t}$ (1)	R$_{s,t+1}$/G$_{s,t}$ (2)	Sale Growth$_{s,t+1}$ (3)	Cybersecurity Relation$_{s,t+1}$ (4)
Supplier MktShare$_{s,t}$	0.568**	1.304***	-0.066	-0.080
	(2.38)	(3.65)	(-0.43)	(-1.33)
Vulnerability$_{s,t}$	-0.002	0.002	-0.001	0.014*
	(-0.74)	(0.43)	(-0.28)	(1.96)
Supplier MktShare$_{s,t}$ × Vulnerability$_{s,t}$	-0.155***	-0.115**	-0.091**	-0.042
	(-2.84)	(-2.01)	(-2.30)	(-0.38)
Supplier Controls$_{s,t}$	Yes	Yes	Yes	Yes
Supplier FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Observations	23,337	23,337	23,337	23,337
Adjusted r^2	0.141	0.395	-0.001	0.492

Panel D: Customer Quarterly Returns Aggregated at the Supplier-Level

	Market-Adjusted$_{s,t+1}$ (1)	CAPM$_{s,t+1}$ (2)	5-factor Model$_{s,t+1}$ (3)
Supplier MktShare$_{s,t}$	-0.003	0.002	0.046***
	(-0.45)	(0.26)	(4.76)
Vulnerability$_{s,t}$	-0.001	-0.004	-0.007**
	(-0.51)	(-1.26)	(-2.16)
Supplier MktShare$_{s,t}$ × Vulnerability$_{s,t}$	-0.006***	-0.006**	-0.003**
	(-3.42)	(-2.09)	(-2.15)
Supplier Controls$_{s,t}$	Yes	Yes	Yes
Supplier FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Observations	25,112	25,112	25,112
Adjusted r^2	0.621	0.167	0.069

Table 9: The Industry-Level Consequences of Software Vulnerabilities

This table studies the effect of software vulnerabilities on Fama-French 49-industry-level outcomes. We report estimates of the model:

$$Industry\ Outcome_{i,t+1} = \alpha + \beta_1 Vulnerabilities_{i,t} + \epsilon_{i,t+1}$$

the dependent variables in Panel A are the number of cybersecurity incidents of any type (columns (1)), related to vulnerabilities (columns (2)), or related to supplier-specific vulnerabilities (columns (3)) that hit the industry i in quarter $t+1$. The dependent variables in Panel B are the return volatility, stock market crash risk, value-at-risk (VaR), and the second-order LPM of industry i in quarter $t+1$. The dependent variables in the first three columns of Panel C are the equal-weigthed average capital investment rate, R&D rate, and sale growth across all customers operating in industry i. In the last column the dependent variable is the number of customers operating in industry i with a relation with a cybersecurity company in quarter $t+1$. The dependent variables in Panel D are the market-adjusted, CAPM, and FF5-factor equal-weighted risk-adjusted quarterly returns of industry i in quarter $t+1$. In all panels, the main independent variable is the number of vulnerabilities affecting customers operating in industry i in quarter t. α indicates Fama-French 49-industry fixed effects. t-statistics based on standard errors clustered by industry are shown in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. A complete list of definitions for these variables is provided in the Data appendix.

Panel A: CyberSecurity Attacks

	All Attacks$_{c,t+1}$ (1)	Vulnerability Attacks$_{c,t+1}$ (2)	Other Attacks$_{c,t+1}$ (3)
Vulnerabilities$_{i,t}$	0.091***	0.104***	0.004
	(5.56)	(11.09)	(0.23)
Industry FE	Yes	Yes	Yes
Observations	3,381	3,381	3,381
Adjusted r^2	0.130	0.209	0.074

Panel B: Risk Variables

	RetVol$_{i,t+1}$ (1)	Ivol$_{i,t+1}$ (2)	Var 95%$_{i,t+1}$ (3)	LPSD$_{i,t+1}$ (4)
Vulnerabilities$_{i,t}$	0.026***	0.037***	-0.016**	0.016**
	(3.73)	(4.62)	(-2.38)	(2.38)
Industry FE	Yes	Yes	Yes	Yes
Observations	3,381	3,381	3,381	3,381
Adjusted r^2	0.180	0.293	0.143	0.139

Panel C: Real Effects

	I$_{i,t+1}$/K$_{i,t}$ (1)	R$_{i,t+1}$/G$_{c,t}$ (2)	Sales Growth$_{i,t+1}$ (3)	Cybersecurity Relation$_{i,t+1}$ (4)
Vulnerabilities$_{i,t}$	-0.104*	-0.098*	0.013	0.148***
	(-1.78)	(-1.78)	(0.38)	(6.63)
Industry FE	Yes	Yes	Yes	Yes
Observations	3,228	3,228	3,228	3,228
Adjusted r^2	0.049	0.050	0.024	0.800

Panel D: Quarterly Returns

	Market-Adjusted$_{i,t+1}$ (1)	CAPM$_{i,t+1}$ (2)	5-factor Model$_{i,t+1}$ (3)
Vulnerabilities$_{i,t}$	-0.003**	-0.004**	-0.001*
	(-2.39)	(-2.57)	(-1.71)
Industry FE	Yes	Yes	Yes
Observations	3,381	3,381	3,381
Adjusted r^2	-0.003	0.012	0.006

Internet Appendix for "Do Software Companies Spread Cyber Risk?"

by Giorgio Ottonello and Emanuele Rizzo

Data appendix: Variable definitions

Variable	Definition
Vulnerability Variables	
$Vulnerability_{c,t}$	Defined as an indicator with a value of 1 if customer c is affected by a software vulnerability in quarter t. In constructing our main independent variable we consider only vulnerabilities with a Common Vulnerability Scoring System (CVSS) above or equal 7. A score above or equal 7 identifies a vulnerability as having high or critical risk. Source: CISA/ZDI databases.
$Vulnerability_{s,t}$	Defined as an indicator with a value of 1 if a software of software supplier s is affected by a software vulnerability in quarter t. In constructing our main independent variable we consider only vulnerabilities with a Common Vulnerability Scoring System (CVSS) above or equal 7. A score above or equal 7 identifies a vulnerability as having high or critical risk. Source: CISA/ZDI databases.
$MinorVulnerability$	Defined as indicator with a value of 1 if a firm is affected by a minor vulnerability. We define a vulnerability minor when the Common Vulnerability Scoring System (CVSS) is below 7. A score below 7 identifies a vulnerability as having medium or low risk. Source: CISA/ZDI databases.
Dependent Variables	
$AllAttacks_{c,t+1}$	Defined as the number of cyberattacks of any type that customer c experiences in quarter $t+1$. Source: Our database.
$AllAttacks_{s,t+1}$	Defined as the number of cyberattacks of any type that customers of software company s experience in quarter $t+1$. Source: Our database.
$VulnerabilityAttacks_{c,t+1}$	Defined as the number of cyberattacks directly caused by a vulnerability in our sample that customer c experiences in quarter $t+1$. Source: Our database.
$VulnerabilityAttacks_{s,t+1}$	Defined as the number of cyberattacks directly caused by a vulnerability in our sample customers of software company s experience in quarter $t+1$. Source: Our database.
$OtherAttacks_{c,t+1}$	Defined as the number of cyberattacks that we are unable to link to a vulnerability in our sample that customer c experiences in quarter $t+1$. Source: Our database.
$OtherAttacks_{s,t+1}$	Defined as the number of cyberattacks that we are unable to link to a vulnerability in our sample customers of software company s experience in quarter $t+1$. Source: Our database.
$RetVol_{c,t+1}$	Computed as the standard deviation of daily stock returns over a quarter. A minimum number of 21 daily returns is required for the calculation. Source: CRSP.
$RetVol_{s,t+1}$	The equal-weighted average of $RetVol_{c,t+1}$ across all customers of software supplier s in quarter t. Source: CRSP.
$Ivol_{c,t+1}$	Computed as the standard deviation of residuals from a regression of daily stock returns on the CAPM, and over quarter. A minimum number of 21 daily returns is required for the calculation. Source: CRSP.
$Ivol_{s,t+1}$	The equal-weighted average of $Ivol_{c,t+1}$ across all customers of software supplier s in quarter t. Source: CRSP.
$Var95\%_{c,t+1}$	Computed as the absolute value of the worst 5% daily return over a quarter. A minimum number of 21 daily returns is required for the calculation. Source: CRSP.

Continued on next page

– continued from previous page

Variable	Definition
$Var95\%_{s,t+1}$	The equal-weighted average of $Var95\%_{c,t+1}$ across all customers of software supplier s in quarter t. Source: CRSP.
$LPSD_{c,t+1}$	Computed as the standard deviation of negative daily stock returns over a 12-month period. A minimum number of 63 daily returns is required for the calculation. Source: CRSP.
$LPSD_{s,t+1}$	The equal-weighted average of $LPSD_{c,t+1}$ across all customers of software supplier s in quarter t. Source: CRSP.
$\mathrm{I}_{c,t+1}/\mathrm{K}_{c,t}$	The capital stock $\mathrm{K}_{c,t}$ is computed recursively using a perpetual-inventory method, as described in Stein and Stone (2013). $\mathrm{I}_{c,t+1}$ represents capital expenditures (CAPX) Source: CRSP/Compustat Merged.
$\mathrm{I}_{s,t+1}/\mathrm{K}_{s,t}$	The equal-weighted average of $\mathrm{I}_{c,t+1}/\mathrm{K}_{c,t}$ across all customers of software supplier s in quarter t. Source: CRSP.
$\mathrm{R}_{c,t+1}/\mathrm{G}_{c,t}$	The R&D stock $\mathrm{G}_{c,t}$ is computed recursively using a perpetual-inventory method, as described in Stein and Stone (2013). $\mathrm{R}_{c,t+1}$ represents R&D expenditures (XRDQ) Source: CRSP/Compustat Merged.
$\mathrm{R}_{s,t+1}/\mathrm{G}_{s,t}$	The equal-weighted average of $\mathrm{R}_{c,t+1}/\mathrm{G}_{c,t}$ across all customers of software supplier s in quarter t. Source: CRSP/Compustat Merged.
Sale $\text{Growth}_{c,t+1}$	The difference in firm sales in $t-1$ minus sales in t, divided by sales in t. Source: CRSP/Compustat Merged.
Sale $\text{Growth}_{s,t+1}$	The equal-weighted average of Sale $\text{Growth}_{c,t+1}$ across all customers of software supplier s in quarter t. Source: CRSP/Compustat Merged.
Cybersecurity $\text{Relation}_{c,t+1}$	Indicator variable that takes value of 1 if customer c hires a cybersecurity company in quarter $t+1$. Source: Factset Revere.
Cybersecurity $\text{Relation}_{s,t+1}$	Indicator variable that takes value of 1 if any of the customers of software supplier s hires a cybersecurity company in quarter $t+1$. Source: Factset Revere.
CAR_c	Computed as the cumulative abnormal return over a 3-day symmetric window around an event. To compute the expected return we use an estimation window of 220 trading days (-251, -31) where day 0 is the date of the event. We use three different risk-adjustments: market-adjusted, calculated as the difference between the stock's return and the market return; CAPM; and Fama-French five-factor. We require at least 63 nonmissing returns in the estimation window. Source: CRSP.
$R_{c,t+1}$	Quarterly cumulative risk-adjusted stock return. We use three different measures of risk-adjusted stock return: the stock's market-adjusted return, calculated as the difference between the stock's return and the market return in quarter $t+1$; CAPM alpha, calculated as the intercept in a CAPM regression of daily excess stock returns on daily excess market returns in quarter $t+1$; five-factor alpha, calculated as the intercept in a five-factor regression of daily excess stock returns on the five Fama and French (2016) factors in quarter $t+1$. Source: CRSP.
Other Variables	
$SupplierMktShare_{s,t}$	Computed as the ratio of the cumulative market capitalization of a software company's customers to the total market capitalization of customers in our data. Source: CRSP and Factset Revere.

Continued on next page

– continued from previous page

Variable	Definition
$SupplierMktShare_{c,t}$	The equal-weighted average of $SupplierMktShare_{s,t}$ across all the software suppliers of customer c. Source: CRSP and Factset Revere.
$Size_{i,t}$	The market value of common equity. Market value of equity is the product of the price (PRC) at the end of quarter t times the contemporaneous number of shares outstanding (SHROUT). Source: CRSP.
$Book-to-market_{i,t}$	The ratio of the book equity divided by the market value of equity. The book value of equity is calculated as stockholder equity, plus deferred taxes and credits, minus the book value of preferred stock. Stockholders' equity is the first available from the following list: i) the Compustat item (SEQ); ii) the book value of common equity (item CEQ), plus preferred stock (item PSTK); iii) the book value of total assets (AT) minus the book value of total liabilities (LT). Deferred taxes and credits are measured through the first available of the following: i) the Compustat item (TXDITC); ii) the sum of balance-sheet deferred taxes (TXDB) and investment tax-credit, (TCB); iii) zero. The book value of preferred stock is the first available of the following: i) redemption value (PSTKRV); ii) liquidation value (PSTKL); iii) par value (PSTK); iv) zero. Source: CRSP/Compustat Merged.
$Past12-monthReturn_{i,t}$	The cumulated continuously compounded stock return from month $j-12$ to month $j-1$, where j is the first month of quarter t. Source: CRSP.
$Profitability_{i,t}$	Ratio of a firm EBITDA over the firm total assets (AT). Source: CRSP/Compustat Merged.
$Leverage_{i,t}$	The ratio of a firm's total debt (long-term debt plus short-term debt) over the firm's total asset value. Source: CRSP/Compustat Merged.
$PastCyberattacks_{i,t}$	Number of cyberattacks attacks a firm has experienced over the past 12 months. Source: Our database.

Table IA2: Vulnerabilities and Affected Customers by Supplier

This table shows the number of common vulnerabilities and exposures (CVE), severe CVEs, number of customers affected by a CVE, and the sum of market capitalization of the affected customers (in billion dollars) for each supplier hit by a vulnerability in our sample.

Supplier	Vulnerabilities	Customers	Mkt Cap. ($B)
ASUSTek	1	9	57
AVEVA Group plc	1	4	49
Adobe	85	152	477
Alphabet	49	574	1200
Apache Software Foundation	26	5	2
Apple	47	346	458
Atlassian	8	18	258
Automattic Inc	3	2	0
Aviatrix	1	1	23
Canon	1	18	30
Cisco	54	155	366
Citrix	8	24	269
Crestron Electronics	1	4	4
D-Link	14	4	0
DASAN Networks	2	2	0
Dell Technologies	1	55	19
Delta Electronics	1	9	258
Docker, Inc.	1	2	0
FatPipe	1	1	0
FatPipe Networks	1	1	0
Fortinet	8	23	37
HP	2	77	65
Hancom	1	1	44
Hangzhou Hikvision Digital Technology	1	3	9
IBM	11	181	191
Intel	1	53	536
Internet Brands	2	1	0
Ivanti	8	4	1
Juniper Networks	3	39	75
LG Electronics Inc.	1	42	167
Liferay	1	1	0
Linux Kernel	12	1	0
McAfee	1	11	8
Merit LILIN Ent Co.	1	1	0
Microsoft	332	742	1269
Mitel Networks Corporation	2	15	5
MongoDB	1	6	5
Mozilla Foundation	9	4	6
Netgear	8	18	265
Nokia	1	56	54
Oracle	33	351	740
Owl Labs	1	1	2
Palo Alto Networks	3	17	28
Progress Software Corporation	1	105	932
QNAP Systems	9	1	2
Qualcomm	2	33	400
Quest Software	1	6	4
RealNetworks	1	13	94
Realtek Semiconductor Corp	2	2	4
Samsung	2	163	1005
Schneider Electric	1	24	16
Siemens	3	126	127
Sitecore	1	3	2
SoftBank Group	2	13	108
SolarWinds	2	92	515
Sophos	3	7	3
Sophos Group plc	3	9	3
Sumavision	1	13	29
Synacor	5	42	269
TRENDnet	1	1	0
Trend Micro	9	16	21
Ubiquiti	1	1	0
Vmware	21	77	93
WatchGuard Technologies	2	3	2
Western Digital	1	11	260
Yealink Network Technology Corporation	1	3	9

Figure IA1: Non Parametric Permutation Tests: Vulnerabilities and Cybersecurity Incidents

This figure shows the results of the block permutation procedure following the method in Chetty, Looney, and Kroft (2009). In each iteration, the dates of the vulnerability discoveries are randomly re-assigned by firm with replacement as a placebo through the sample period. Our regression of column (2) of Panel A, and B of Table IA5 are then estimated on the falsified data. The plots report the empirical cumulative distribution function (cdf) generated from running each of the regression models in 1,000 random iterations of this procedure and capturing the placebo coefficient estimate. The vertical dotted line indicates the position of the actual coefficient estimate for the impact that a vulnerability discovery has on the probability of future cyberattacks and implied p-value when placed in the context of cdf. The implied p-value reported in each plot shows the proportion of the placebo coefficients that are contrasted with the actual regression coefficient.

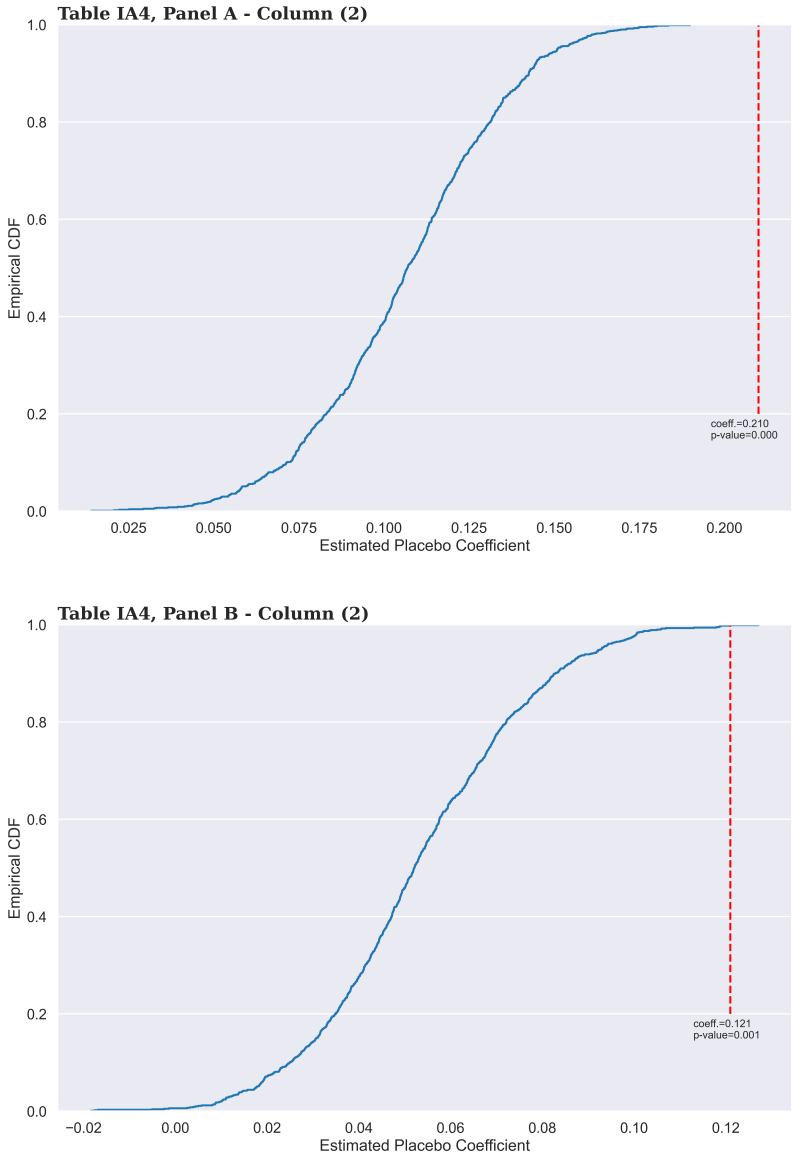

Table IA3: Vulnerabilities and Cybersecurity Incidents: Other Type of Attacks

This table studies the relation between the probability of cybersecurity attacks and exposure to software vulnerabilities. In both panels, we report estimates of the model:

$$CyberAttacks_{i,t+1} = \alpha + \beta\, Vulnerability_{i,t} + \gamma X_{i,t} + \epsilon_{i,t+1}$$

the dependent variables in Panel A are the number of cybersecurity incidents classified as phishing (columns (1) and (2)), DDoS (columns (3) and (4)), credential stuffing (columns (5) and (6)), malware (columns (7) and (8)), or miscellaneous (columns (9) and (10)) that hit the software customer c in quarter t. In all columns, the main independent variable is an indicator with value 1 if customer c is exposed to a vulnerability of one of its software suppliers in quarter t. The dependent variables in Panel B are the number of cybersecurity incidents classified as phishing (columns (1) and (2)), DDoS (columns (3) and (4)), credential stuffing (columns (5) and (6)), malware (columns (7) and (8)), or miscellaneous (columns (9) and (10)) that hit the software supplier s customers in quarter t. In all columns, the main independent variable is an indicator with value 1 if one of the software of supplier s has a vulnerability in quarter t. Controls include market capitalization, book-to-market, ROA, previous 12-month return, leverage. In Panel A we also include the number of cybersecurity incidents that the software customer c experienced over the previous year. In Panel B we also include the number of cybersecurity incidents that the software supplier s customers experienced over the previous year. α indicates different sets of fixed effects, including time (year-quarter) and firm fixed effects. t-statistics based on standard errors clustered by firm are shown in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. A complete list of definitions for these variables is provided in the Data appendix.

Panel A: Customer Attacks

	$\text{Phishing}_{c,t+1}$ (1)	$\text{DDoS}_{c,t+1}$ (2)	$\text{Cred. Stuffing}_{c,t+1}$ (3)	$\text{Malware}_{c,t+1}$ (4)	$\text{Miscellaneous}_{c,t+1}$ (5)
$\text{Vulnerability}_{c,t}$	0.002	0.006	0.008	0.010	-0.002
	(0.35)	(1.15)	(0.95)	(1.64)	(-0.31)
$\text{Customer Controls}_{c,t}$	Yes	Yes	Yes	Yes	Yes
Customer FE	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
Observations	81,125	81,125	81,125	81,125	81,125
Adjusted r^2	-0.007	-0.007	-0.018	-0.020	-0.021

Panel B: Customer Attacks Aggregated at the Supplier-Level

	$\text{Phishing}_{s,t+1}$ (1)	$\text{DDoS}_{s,t+1}$ (2)	$\text{Cred. Stuffing}_{s,t+1}$ (3)	$\text{Malware}_{s,t+1}$ (4)	$\text{Miscellaneous}_{s,t+1}$ (5)
$\text{Vulnerability}_{s,t}$	0.010	0.003	0.005	-0.001	0.006
	(1.17)	(0.36)	(0.64)	(-0.70)	(0.43)
$\text{Supplier Controls}_{s,t}$	Yes	Yes	Yes	Yes	Yes
Supplier FE	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
Observations	25,363	25,363	25,363	25,363	25,363
Adjusted r^2	0.179	0.143	0.073	0.170	0.177

Table IA4: Vulnerabilities and Cybersecurity Incidents: Minor Vulnerabilities

This table studies the relation between the probability of cybersecurity attacks and exposure to software vulnerabilities. In both panels, we report estimates of the model:

$$CyberAttacks_{c,t+1} = \alpha + \beta Minor\ Vulnerability_{i,t} + \gamma X_{i,t} + \epsilon_{i,t+1}$$

the dependent variables in Panel A are the number of cybersecurity incidents of any type (columns (1)), related to vulnerabilities (columns (2)), or due to causes different from vulnerabilities (columns (3)) that hit the software customer c in quarter $t+1$. In all columns, the main independent variable is an indicator with value 1 if customer c is exposed to a minor vulnerability of one of its software suppliers in quarter t. The dependent variables in Panel B are the number of cybersecurity incidents of any type (columns (1)), related to vulnerabilities (columns (2)), or due to causes different from vulnerabilities (columns (3)) that hit the software supplier s customers in quarter $t+1$. In all columns, the main independent variable is an indicator with value 1 if supplier s product has a minor vulnerability in quarter t. We define a vulnerability minor when the vulnerability scoring is below seven. Controls include market capitalization, book-to-market, ROA, previous 12-month return, leverage. In Panel A we also include the number of cybersecurity incidents that the software customer c experienced over the previous year. In Panel B we also include the number of cybersecurity incidents that the software supplier s customers experienced over the previous year. α indicates different sets of fixed effects, including time (year-quarter) and firm fixed effects. t-statistics based on standard errors clustered by firm are shown in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. A complete list of definitions for these variables is provided in the Data appendix.

Panel A: Customer Attacks

	All Attacks$_{c,t+1}$ (1)	Vulnerability Attacks$_{c,t+1}$ (2)	Other Attacks$_{c,t+1}$ (3)
Minor Vulnerability$_{c,t}$	0.011	-0.003	0.014
	(1.24)	(-0.44)	(1.53)
Customer Controls$_{c,t}$	Yes	Yes	Yes
Customer FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Observations	81,125	81,125	81,125
Adjusted r^2	0.017	0.006	0.013

Panel B: Customer Attacks Aggregated at the Supplier-Level

	All Attacks$_{s,t+1}$ (1)	Vulnerability Attacks$_{s,t+1}$ (2)	Other Attacks$_{s,t+1}$ (3)
Minor Vulnerability$_{s,t}$	0.003	0.012	-0.002
	(0.96)	(1.31)	(-1.18)
Supplier Controls$_{s,t}$	Yes	Yes	Yes
Supplier FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Observations	25,363	25,363	25,363
Adjusted r^2	0.325	0.172	0.254

Table IA5: Vulnerabilities and Cybersecurity Incidents: Alternative Specifications

This table studies the relation between the probability of cybersecurity attacks and exposure to software vulnerabilities. In Panel A we report results of a "stacked regression" (see Gormley and Matsa (2011), Baker, Larcker, and Wang (2022)) with firm- and industry-time fixed effects saturated with indicators for cohort identifiers:

$$CyberAttacks_i = \alpha_{i,e} + \alpha_{t,e} + \beta\, Vulnerability_i + \gamma X_i + \epsilon_i$$

the dependent variables are the number of cybersecurity incidents of any type (columns (1) and (2)), related to vulnerabilities (columns (3) and (4)), or due to causes different from vulnerabilities (columns (5) and (6)) hit the software customer i (Panel A), or that hit the software supplier i customers (Panel B). The main independent variable is an indicator with value 1 if the firm is affected by a vulnerability in the event quarter $t = 0$. For each event we observe firms outcomes from 4 quarters before the event to 4 quarters after. Controls include market capitalization, book-to-market, ROA, previous 12-month return, leverage, institutional ownership, and the number of past cybersecurity incidents that a firm suffered. $\alpha_{i,e}$ and $\alpha_{t,e}$ indicate firm-cohort and time-cohort fixed effects, respectively. t-statistics based on standard errors clustered by firm and event date (year-quarter) are shown in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. A complete list of definitions for these variables is provided in the Data appendix.

Panel A: Customer Attacks

	All Attacks$_{i,t}$	Vulnerability Attacks$_{i,t}$	Other Attacks$_{i,t}$
Vulnerability$_i$	0.267***	0.210***	0.057
	(3.33)	(3.44)	(0.86)
Customer Controls$_{i,t-1}$	Yes	Yes	Yes
Firm x Cohort FE	Yes	Yes	Yes
Time x Cohort FE	Yes	Yes	Yes
Observations	675,603	675,603	675,603
Adjusted r^2	0.062	0.011	0.059

Panel B: Customer Attacks Aggregated at the Supplier-Level

	All Attacks$_{i,t}$	Vulnerability Attacks$_{i,t}$	Other Attacks$_{i,t}$
Vulnerability$_i$	0.180**	0.121**	0.058
	(2.54)	(2.08)	(1.23)
Supplier Controls$_{i,t-1}$	Yes	Yes	Yes
Firm x Cohort FE	Yes	Yes	Yes
Time x Cohort FE	Yes	Yes	Yes
Observations	209,618	209,618	209,618
Adjusted r^2	0.318	0.098	0.275

Table IA6: The Role of Supplier Market Share - Customer-Level Results

This table studies how supplier market share affects the relation between the probability of cybersecurity attacks, risk variables, quarterly stock returns, and exposure to software vulnerabilities. In all panels, we report estimates of the model:

$$Y_{c,t+1} = \alpha + \beta_1 \, Vulnerability_{c,t} + \beta_2 Supplier\ Mktshare_{c,t} + \beta_3 \, Vulnerability_{c,t} \times Supplier\ Mktshare_{c,t} + \gamma X_{c,t} + \epsilon_{c,t+1}$$

the dependent variables in Panel A are the number of cybersecurity incidents of any type (columns (1)), related to vulnerabilities (columns (2)), or related to supplier-specific vulnerabilities (columns (3)) that hit the software customer c in quarter $t+1$. The dependent variables in Panel B are the return volatility, stock market crash risk, value-at-risk (VaR), and the second-order LPM of software customer c in quarter $t+1$. The dependent variables in Panel C are customer c capital investment rate $I_{c,t+1}/K_{c,t}$, R&D rate $R_{c,t+1}/G_{c,t}$, sale growth, and an indicator variable that takes value of 1 if customer c has a relation with a cybersecurity company in quarter $t+1$. The dependent variables in Panel D are the market-adjusted, CAPM, and FF5-factor risk-adjusted quarterly returns of software customer c in quarter $t+1$. In all panels, the main independent variable is an indicator with value 1 if customer c is exposed to a vulnerability of one of its software suppliers in quarter t. *Supplier Mktshare* is defined as the average market share across all customer c's software suppliers. At the supplier level, market share is computed as the ratio of the cumulative market capitalization of software company s customers to the total market capitalization of customers in our data in quarter t. Controls include market capitalization, book-to-market, ROA, previous 12-month return, leverage. We also include the number of cybersecurity incidents that the software customer c experienced over the previous year. α indicates different sets of fixed effects, including year-quarter and software customer fixed effects. t-statistics based on standard errors clustered by customer are shown in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. A complete list of definitions for these variables is provided in the Data appendix.

Panel A: Customer Attacks

	All Attacks$_{c,t+1}$ (1)	Vulnerability Attacks$_{c,t+1}$ (2)	Other Attacks$_{c,t+1}$ (3)
Supplier MktShare$_{c,t}$	-0.008	0.003	-0.009
	(-1.02)	(0.43)	(-1.10)
Vulnerability$_{c,t}$	0.051**	0.094***	0.026
	(2.21)	(2.94)	(1.13)
Supplier MktShare$_{c,t}$ × Vulnerability$_{c,t}$	0.078***	0.121***	0.009
	(3.06)	(3.10)	(0.39)
Customer Controls$_{c,t}$	Yes	Yes	Yes
Customer FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Observations	81,125	81,125	81,125
Adjusted r^2	0.017	0.006	-0.008

Panel B: Customer Risk

	RetVol$_{c,t+1}$ (1)	Ivol$_{c,t+1}$ (2)	Var 95%$_{c,t+1}$ (3)	LPSD$_{c,t+1}$ (4)
Supplier MktShare$_{c,t}$	0.002	0.000	-0.004	0.003
	(0.31)	(0.07)	(-0.71)	(0.41)
Vulnerability$_{c,t}$	0.011	0.007	-0.025*	0.027*
	(0.78)	(0.47)	(-1.77)	(1.74)
Supplier MktShare$_{c,t}$ × Vulnerability$_{c,t}$	0.044***	0.037***	-0.046***	0.033***
	(3.86)	(3.29)	(-4.20)	(2.81)
Customer Controls$_{c,t}$	Yes	Yes	Yes	Yes
Customer FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Observations	81,125	81,125	81,125	81,125
Adjusted r^2	0.561	0.577	0.542	0.438

Panel C: Customer Real Effects

	I$_{c,t+1}$/K$_{c,t}$ (1)	R$_{c,t+1}$/G$_{c,t}$ (2)	Sale Growth$_{c,t+1}$ (3)	Cybersecurity Relation$_{c,t+1}$ (4)
Supplier Mktshare	0.001	-0.015**	0.001	0.019*
	(0.06)	(-2.21)	(0.14)	(1.87)
Vulnerability$_{c,t}$	-0.003	-0.031*	-0.020	-0.023
	(-0.16)	(-1.76)	(-1.41)	(-0.45)
Supplier Mktshare × Vulnerability$_{c,t}$	-0.058***	-0.009	-0.022**	0.171***
	(-4.01)	(-0.74)	(-2.17)	(3.36)
Customer Controls$_{c,t}$	Yes	Yes	Yes	Yes
Customer FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Observations	69,221	69,221	69,221	69,221
Adjusted r^2	0.203	0.757	0.025	0.656

Panel D: Customer Quarterly Returns

	Market-Adjusted$_{c,t+1}$ (1)	CAPM$_{c,t+1}$ (2)	5-factor Model$_{c,t+1}$ (3)
Supplier Mktshare$_{c,t}$	-0.004***	-0.003***	-0.002**
	(-3.05)	(-3.09)	(-2.01)
Vulnerability$_{c,t}$	0.001	-0.002	-0.002
	(0.17)	(-0.61)	(-0.68)
Supplier Mktshare$_{c,t}$ × Vulnerability$_{c,t}$	-0.009***	-0.005**	-0.003
	(-2.96)	(-2.34)	(-1.31)
Customer Controls$_{c,t}$	Yes	Yes	Yes
Customer FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Observations	77,351	77,351	77,351
Adjusted r^2	0.063	0.062	0.040