Journal of Financial Stability 60 (2022) 100989

Contents lists available at ScienceDirect

Journal of Financial Stability

journal homepage: www.elsevier.com/locate/jfstabil

The drivers of cyber risk☆

Iñaki Aldasoro [a], Leonardo Gambacorta [a,*], Paolo Giudici [b], Thomas Leach [b]

[a] Bank for International Settlements, Centralbanhnplatz 2, 4002, Basel, Switzlerland
[b] University of Pavia, Via San Felice 5, 27100, Pavia, Italy

ARTICLE INFO

JEL classification:
D5
D62
D82
G2
H41

Keywords:
Cyber risk
Cloud services
Financial institutions
Cyber cost
Cyber regulation

ABSTRACT

Cyber incidents are becoming more sophisticated and their costs difficult to quantify. Using a unique database of cyber events across sectors in the US, we document the characteristics and drivers of cyber incidents. Cyber costs are higher for larger firms and for incidents that impact several organisations simultaneously. Events with malicious intent (i.e. cyber attacks) tend to be less costly, unless they are on the upper tail of the loss distribution. The financial sector is exposed to a larger number of cyber attacks but suffers lower costs, on average. The use of cloud services is associated with lower costs, especially when cyber incidents are relatively small. As cloud providers become systemically important, cloud dependence is likely to increase tail risks. Finally, we document that higher expenditure on IT is associated with future reduced costs from cyber incidents.

1. Introduction

Information technology (IT) has become a critical component of well-functioning economies, underpinning economic growth over the past decades. Organisations of all sizes in both the public and private sector are becoming ever more interconnected and reliant on IT products and services, such as cloud-based systems and artificial intelligence. Accordingly, there is a growing exposure to cyber risks, and public awareness of these threat has been on the rise (see Fig. 1). Cyber risk commonly refers to the risk of financial loss, disruption or reputational damage to an organisation resulting from the failure of its IT systems.[1] The increasing reliance on cloud technologies exacerbates these risks, as it increases interdependence across firms that have shared exposures to similar (or even the same) cloud service providers.

Firms actively manage cyber risk and invest in cyber security. However, cyber costs are difficult to quantify.[2] In the financial sector, cyber risks are a key "known unknown" tail risk to the system and a potential major threat to financial stability.[3] More broadly, cyber risk in sectors that play a critical role in the economic infrastructure could have systemic implications and can be viewed as a matter of national security (Brenner, 2017). Despite such considerations, information concerning the costs, drivers and potential mitigating factors of cyber incidents is relatively scarce.

☆ We would like to thank Iftekhar Hasan (the editor) and two anonymous referees for the comments during the review process. We also thank Raymond Kleijmeer, Karin Reichardt, Andreas Voegtli and conference/seminar participants at Annual conference on Bank Regulation and Supervision (Columbia University), Bank of Italy, Bank for International Settlements, Central Bank of Malta, IFABS 2019, IFI European Chapter Forum, IFI Insurance Management Forum and OECD Blockchain Policy Forum 2019 for helpful comments and suggestions. We are grateful to Duane Kennedy and Theophanis Stratopoulos for sharing their data on IT spending. Declarations of interest: none. The views expressed here are those of the authors and do not necessarily represent those of the Bank for International Settlements. Paolo Giudici and Thomas Leach acknowledge funding from the European H2020 project FIN-TECH: A financial supervision and technological compliance programme, Grant agreement ID: 825215.

* Corresponding author.
E-mail addresses: inaki.aldasoro@bis.org (I. Aldasoro), leonardo.gambacorta@bis.org (L. Gambacorta), paolo.giudici@unipv.it (P. Giudici), thomas.leach01@universitadipavia.it (T. Leach).

[1] These episodes include malicious cyber incidents (cyber attacks) where the threat actor intends to do harm (e.g. ransomware attacks, hacking incidents or data theft by employees). High-profile attacks such as the WannaCry incident in May 2017 contributed to the growing concern around cyber risk.

[2] The high degree of uncertainty and variability surrounding cost estimates for cyber security incidents has consequences for policy-makers. For example, it is difficult to foster robust insurance markets, as well as to make decisions about the appropriate level of investment in security controls and defensive interventions (Biener et al., 2015; Wolff and Lehr, 2017).

[3] In March 2017, the G20 Finance Ministers and Central Bank Governors noted that "the malicious use of information and communication technologies could disrupt financial services crucial to both national and international financial systems, undermine security and confidence, and endanger financial stability".

https://doi.org/10.1016/j.jfs.2022.100989
Received 3 May 2021; Received in revised form 16 February 2022; Accepted 22 February 2022
Available online 4 March 2022

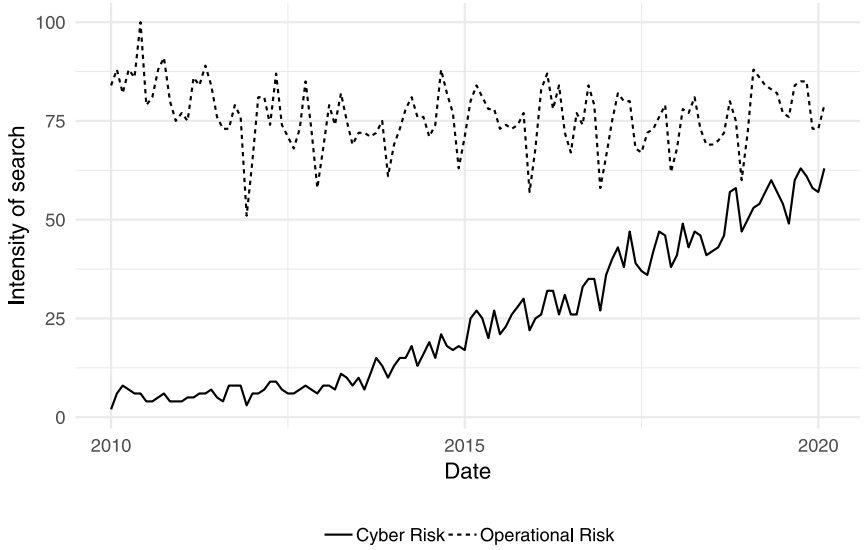

Fig. 1. Interest on cyber risk is on a par with operational risk. **Notes:** Number of online searches for "cyber risk" and "operational risk" over the last decade. Worldwide search interest is relative to the highest point (=100). Data accessed on 7 Feb 2020.
Source: Google Trends.

This paper seeks to help fill this gap by using a sample of 3705 cyber events across all economic sectors in the US, sourced from the Advisen cyber loss database. First, we document a series of stylised facts. The frequency of cyber incidents rose strongly in the decade leading up to 2016, but has since moderated somewhat. This reduction could reflect increased investment in cyber security, but also delays in discovery or reporting.[4] We find that certain economic sectors display a greater resilience to cyber incidents: for example, the financial sector has experienced a higher frequency of cyber incidents but these have been, on average, relatively less costly. Regarding the type of incident, data breaches, phishing or skimming and security incidents, appear to be most costly. Of particular concern is that data breaches are not only costly, but also relatively frequent.

The paper then documents cyber risk drivers. We first identify the key drivers contributing to the costs of cyber-related events. Firm size – measured in terms of total revenues – is positively correlated with the average cost of an event, implying that larger firms tend to incur larger costs. However, the elasticity is quite low: a 1% increase in total revenues is associated with a 0.2% increase in cyber costs. Cyber events impacting multiple firms at the same time (i.e. "connected" events) are also associated with higher costs. Cyber-related incidents can occur unintentionally – e.g. a bug in some internally developed software – or can also be caused by an actor with malicious intent.[5] Malicious cyber attacks have, *on average*, lower costs. However, a quantile analysis reveals that at the tail of the sample distribution this relationship is reversed and in fact malicious incidents are associated with higher costs. This finding indicates that, while most attackers are stopped before they can do considerable harm, a successful attacker can go on to cause extensive damage.

We then study the effects of reliance on cloud services and digital technologies more broadly. In principle, reliance on the same service provider by multiple organisations can yield positive externalities by fostering economies of scale and information sharing (Rowe, 2007). Cloud technology can thus reduce IT costs, improve resilience and enable firms to scale better (Financial Stability Board, 2019). However, it also strengthens interdependence, not least given the high concentration of the market for cloud service providers. By analysing the cost–benefit trade-off, we find that the use of cloud services is associated with lower costs of cyber events. While this speaks to the resilience of cloud technology, it should be interpreted with caution. As firms' exposure to cloud services continues to increase and cloud providers become systemically important, cloud dependence is likely to increase tail risks (Danielsson and Macrae, 2019).

Finally, we use data on the level of IT spending across sectors to assess the relationship between investment in IT and the cost of cyber incidents. This analysis can act as a helpful indicator to policymakers as to which sectors may be exposed due to underinvestment in their IT systems. We find that higher expenditure in IT is associated with lower costs at the mean and at the tail of the distribution. Sectors that appear to benefit from this higher level of spending include the manufacturing and the finance and insurance sector. For the latter this could be due to the effects of regulation and a years of experience being a prime target for cyber criminals. The dividend of such investments is evident through our additional analyses, whereby an annual increase in IT investment is associated with a reduction of costs in the subsequent year.

The rest of the paper is organised as follows. Section 2 discusses related literature. Section 3 contains a description of the data. Section 4 discusses our baseline results. 5 explores whether exposure to cloud services affects the cost of cyber events. Section 6 analyses the optimal amount of IT spending across sectors. Finally, Section 7 concludes.

2. Related literature

Most of the few empirical studies on cyber risk rely on collected publicly available data sources. Goldstein et al. (2011) study how the exposure to IT operational risk could translate into significant losses

[4] This phenomena is widely recognised in the operational risk literature (see Aldasoro et al. (2020) and Carrivick and Cope (2013)). The dataset used here does not allow us to accurately estimate such "end-of-sample" bias.

[5] The best known types of cyber attack are: man-in-the-middle attacks, cross-site scripting, denial-of-service attacks, password attacks, phishing, malware and zero-day exploits. Man-in-the-middle attacks occur when attackers insert themselves into a two-party transaction. Cross-site scripting is a web security vulnerability that allows attackers to compromise the interactions a victim has with a vulnerable application. Denial-of-service attacks flood servers with traffic to exhaust bandwidth or consume finite resources. Phishing is the practice of stealing sensitive data by sending fraudulent emails that appear to be from a trustworthy source. Malware (i.e. "malicious software") is a software designed to cause damage to IT devices and/or steal data (examples include so-called Trojans, spyware and ransomware). A zero-day exploit is an attack against a software or hardware vulnerability that has been discovered but not publicly disclosed.

in firms' market values. Biener et al. (2015) emphasise the distinct characteristics of cyber risks compared to other operational risks. The presence of highly interrelated cyber losses, lack of data, and severe information asymmetries, hinder the development of a sustainable cyber insurance market, an essential element to encourage improvements in cyber resilience. Romanosky (2016) and Chande and Yanchus (2019) use the Advisen dataset to study losses from cyber events across sectors and provide an initial estimate of firm risk by sector. Our paper builds on their work by looking at how characteristics of sectors' management of IT resources can mitigate costs.

An important part of the cost of cyber events is arguably given by the reputational component, which is notably hard to assess. Makridis (2020) find that for the subset of the largest data breaches, brand power (a survey-based measure of reputation) and familiarity decrease by 5%–9% after the event (whereas they increase by 26%–29% for a typical data breach). Kamiya et al. (2021) find that a successful data breach can decrease shareholder wealth by 1.09% in the three-day window around the attack. These findings suggest economically large reputation costs.

The literature highlights that the observed heterogeneity in cyber costs across sectors heavily depends on the environment in which each firm operates, as well as IT security investments. Kamiya et al. (2018) find that cyber attacks are more likely in industries that face less intense product market competition and with higher growth opportunities. Moreover, controlling for firm characteristics, they find that, among the major industries, cyber attacks are more likely in service industries, wholesale/retail trade, and transportation and communications. Makridis and Dean (2018) find heterogeneity in cyber attack episodes amongst sectors when it comes to data breaches. In particular, companies in the finance, insurance, retail and merchant sectors are the biggest targets. Makridis and Liu (2021) also suggest that higher productivity firms have fewer cyber security vulnerabilities and are able to gain access to more and better human capital to help mitigate cyber security vulnerabilities.

Regulation can also play a key role in firms' motives for security investments. Based on a survey of more than 700 firms, Rowe and Gallaher (2006) find that the vast majority believe that regulation has increased the overall level of security. However, some firms reject this view, because excessive cyber security costs imposed by regulation could stifle firms' ability to innovate (Etzioni, 2011). While our paper does not enter into the debate on who should bear the cost of cyber security, we find that sectors with a more robust policy framework toward cyber risk tend to reap benefits in terms of lower costs of cyber incidents.

Some sectors make better targets than others, as they provide critical infrastructure for the functioning of the economy. The financial sector is frequently targeted due to its high exposure to IT and its credit intermediation role (Kopp et al., 2017). Cyber attacks on this sector could create cascade failures that are not completely understood nor adequately quantified by sector-specific simulations (Brenner, 2017). Kashyap and Wetherilt (2019) outline some principles for regulators to consider when regulating cyber risk in the financial sector. The Basel Committee has also published guidelines for banks regarding best practice regarding cyber risk (Basel Committee on Banking Supervision, 2018). Given that financial institutions tend to maintain better data collection practices due to regulatory reporting, empirical studies focusing on this sector are more developed.

Using a large cross-country panel, Aldasoro et al. (2020) find that cyber losses represent a relatively small share of operational losses for banks. In recent years, however, losses from cyber events saw a spike, with a corresponding increase in risk. The value-at-risk (VaR) associated with cyber events can range from 0.2% to 4.2% of banks' income.[6] This amounts to around a third of operational VaR. The extent of operational and cyber losses depends on the supervisory environment. A higher quality of supervision – as measured by a financial and supervisory quality index – is associated with lower losses, in terms of both frequency and amount.

Duffie and Younger (2019) find that a sample of 12 systemically important U.S. financial institutions have sufficient stocks of high-quality liquid assets to cover wholesale funding runoffs in an extreme cyber event.[7] From the literature on operational risk, the size of financial institutions is positively linked with the that of operational losses (Shih et al., 2000; Curti et al., 2019). A large share of banks' operational losses can be traced to a breakdown of internal controls (Chernobai et al., 2011). We devote particular attention to the drivers of cyber risks in the financial sector and how these could differ from other economic sectors.

3. Data

The data are obtained from Advisen, a for-profit organisation which collects information from reliable and publicly verifiable sources such as websites, newsfeeds, specialised legal information services, multiple online data breach clearinghouses and federal and state governments in the United States.[8] The entire Advisen database contains a total of 137,164 cyber incidents. Each cyber incident is linked to an ultimate parent company and includes, amongst others, the following characteristics: (i) case type (e.g. data breach, phishing); (ii) affected count (e.g. in the event of a data breach, how many details were stolen); (iii) accident date; (iv) source of the loss; (v) type of loss; (vi) actor (e.g. state-sponsored, terrorist, etc.); (vii) loss amount; (viii) company size (proxied by total revenues); (ix) company type (e.g. government, private); (x) number of employees; (xi) North American Industry Classification System (NAICS) code identifying the sector of the firm that suffered the cyber incident; and (xii) geography (i.e. the area where the incident occurred).

The majority of events reported in the database occur in the Americas region (North, Central and South America). In particular, 86 per cent of the episodes took place in the United States. This is largely due to the fact that information regarding cyber losses is easier to collect there as a result of a higher degree of freedom of information. To remove unobserved country heterogeneity from our analysis we focus on the incidents in the database that occurred in the United States, which leaves a sample of 116,387 incidents. However, due to the nature of how the data are collected, it is not possible to obtain all information desirable for each event.

Data on actual loss amounts per event represent only a subset of the larger database. The cost of cyber events can be categorised into three components (Anderson et al., 2019). The *direct cost* is the value of loss, damage and other suffering incurred by the victim of the cyber incident. The *indirect costs* are the losses and opportunity costs borne by society as a consequence of a cyber incident.[9] Firms

[6] Estimates by Bouveret (2018) – based on data collected from media and newspaper articles across countries – point to sizeable potential losses in the financial sector. His estimate of value-at-risk ranges between 14% to 19% of net income.

[7] Using a broader network of US banks, Eisenbach et al. (2020) find that the impairment of any of the five most active banks can result in significant spillovers to other banks, with 38% of the network affected on average.

[8] Losses are not estimated by Advisen, but collected from various third party sources. For example, losses for publicly traded companies are typically sourced from financial statements. In the case of litigated amounts, court records are the most likely source. There are as well other situations where information is gathered from news sources. Most cyber incidents go unreported. Typically, only the larger and the more relevant ones become public and are included in the Advisen database.

[9] Examples of direct costs are those related to the time and effort of repairing IT systems damaged as a result of an incident, the ransom paid to attackers in a successful cyber attack or regulatory fines and penalties. Indirect costs could in turn include reduced uptake by citizens of electronic services whether from companies or governments due to the perceived threat of a cyber incident or the losses incurred by an individual after having their personal data stolen.

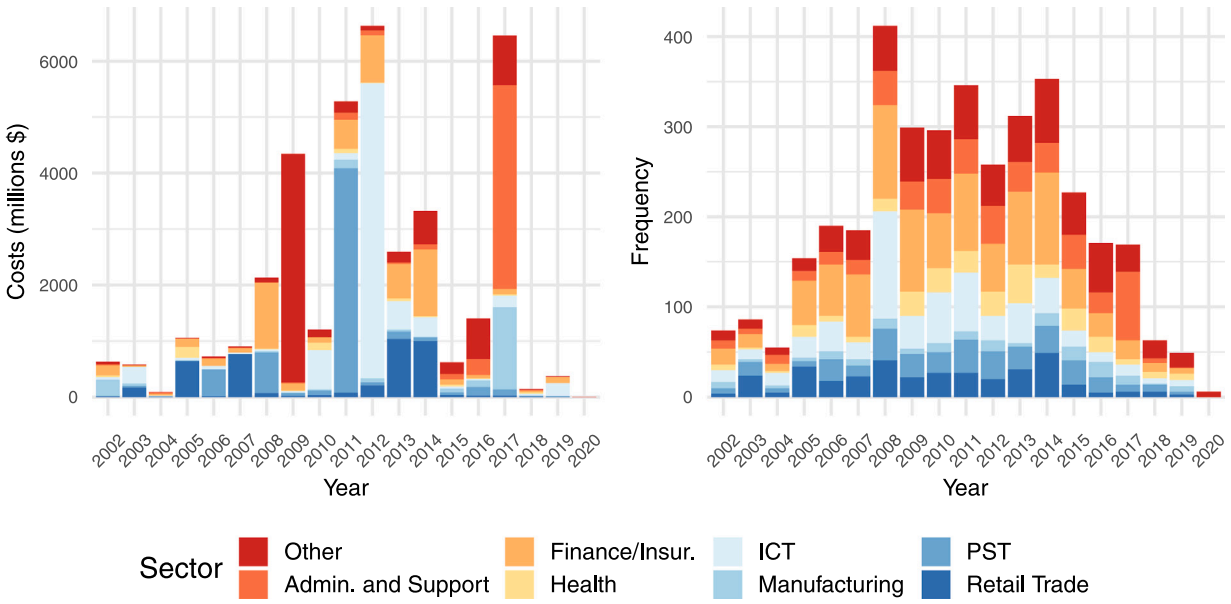

Fig. 2. Frequency and cost of cyber incidents across sectors. **Notes:** ICT stands for "Information and Communication Technology", while PST for "Professional Scientific and Technical". The plot shows costs and frequency over time with partitioning by the sectors that suffered the most incidents based on Table A.2 in the online annex, remaining sectors are subsumed into the "Other" category. The graphs are based on the 3705 observations used in our regressions and further described in Table A.2 in the online annex.

also bear *mitigation costs*, which include inter alia investment in IT personnel or in security products such as antivirus software or cyber threat awareness training for staff. The data from Advisen can best be interpreted as a measure of *direct costs* to a firm as a result of a cyber incident. Individual components of each loss (e.g., fines or penalties from regulators, payments made to a plaintiff in the event of a claim and financial damages) are provided in the data, but are rarely populated in sufficient detail to allow for a meaningful analysis. For our regression analysis, we remove observations missing such critical data, which leaves us with a sample of 3705 observations for our baseline empirical analysis.[10]

The frequency and costs of cyber events differ across sectors (see Table A.2 in the online annex for summary statistics).[11] By frequency, "Financial and insurance activities" (FI) is the most affected sector. However, it shows some resilience, as despite being subject to many attacks, the average cost of a cyber incident is not as high as for other sectors. The sector with the highest average costs is "Wholesale Trade", followed by "Transportation and storage" and "Professional, Scientific and Technical" (PST). The standard deviation of costs across sectors is quite large, implying that most likely the distribution of losses has a heavy tail.

Fig. 2 shows how events are distributed by sector over time. The overall distribution across sectors in terms of frequency remains relatively stable. Much of the growing frequency of events can be attributed to the FI sector as well as the "Administrative and Support Service" sector. Increases in the frequency of cyber incidents in the FI sector following the great financial crisis may be partly driven by targeted attacks on banks. The peak in costs in 2012 was shared largely amongst the FI and "Information and communication technology" (ICT) sectors.

Fig. 3 depicts the distribution by case type through time.[12] Privacy violations are the frequent type (44 per cent of the cases). This is likely due to the fact that reporting requirements have been in place for a longer period for such incidents, as well as the relative ease of assigning conclusive responsibility when they occur (Chande and Yanchus, 2019). Data breaches have been responsible for a significant portion of costs over time. The total cost of a data breach grows with the amount of records stolen. Therefore, if hackers are able to obtain large volumes of records, the costs can soar as millions of individuals can be affected.

In terms of frequency, the overall trend has been positive, in line with the growing concern over cyber risks.[13] This is likely driven by a few factors. First, several frameworks and legislation have come into place that encourage the reporting of cyber incidents. Second, the barrier to carrying out cyber attacks has become lower as competent computing skills are no longer required to carry out attacks. The reduction in more recent years could represent the effects of increased investment in cyber security, but should be taken with caution due to the above mentioned reporting bias. Regarding the distribution by frequency over time, the increase has largely been attributed to privacy violations. Costs, on the other hand, peaked in 2011, largely due to spikes in privacy violations and data breaches.

4. Identifying the drivers of cyber costs

4.1. Empirical approach

Our analysis aims to explain the costs of cyber events by a series of event and firm/sector characteristics.[14] We model the direct costs of a

[10] The sample used for the regression analysis (i.e., which includes loss data) presents some differences with respect to the full sample of US firms. This can be seen by comparing the mean and the standard deviation of the main variables of interest in the regression sample and in the full sample (see Table A.1 in the annex). Firms used in the regression sample are bigger and experience a higher number of connected cyber incident events with respect to the full sample if US firms.

[11] The sectors are based on NAICS. For details, see https://www.census.gov/eos/www/naics/.

[12] Case types are based on the definitions of Romanosky (2016) and are used as fixed effects in our regression (see Section 4 for details). Further categorisations are possible, though they do not provide enough variation for econometric analysis. We provide additional details on case types in the online annex (see Figures A.1–A.3).

[13] As noted earlier, the most recent end of the data is probably subject to an under-reporting bias, as it takes time for incidents to be discovered and acknowledged. Therefore, we expect the numbers in the most recent years of our sample to increase as more information becomes available in the future.

[14] These include observable direct costs from cyber events. Costs are likely to be a lower bound for a number of reasons. For one, there are many costs

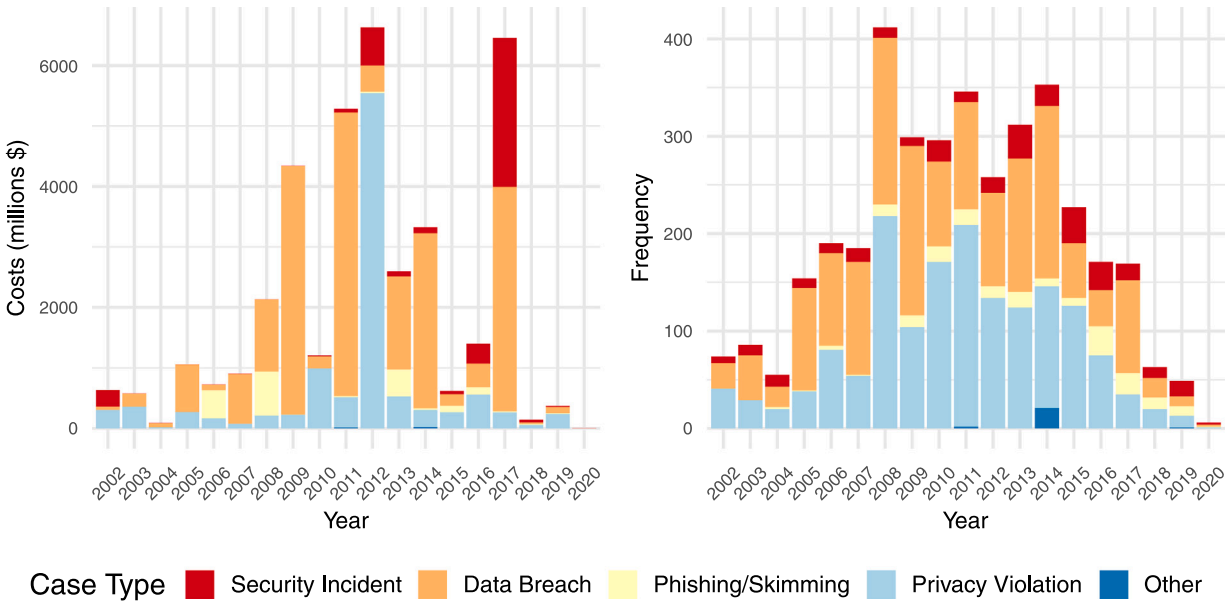

Fig. 3. Frequency and cost of cyber incidents by case type. **Notes:** The graphs show costs and frequency over time by case type. The graphs are based on the 3705 observations used in our regressions and further described in Table A.2 in the online annex.

cyber incident through the following regression:

$$C_{i,f,g} = \beta Z_{i,f,g} + \lambda W_{f,g} + \theta X_g + \eta_k + \alpha_t + u_{i,f,g} \quad (1)$$

where, i denotes the individual incident ($N_i = 3705$), f denotes the firm at which the incident took place ($N_f = 2445$) and g the sector of the firm ($N_g = 19$), based on the NAICS sector categorisation. $C_{i,f,g}$ denotes the cost of the incident; X_g denotes sector-level controls; $W_{f,g}$ denotes firm-level variables and $Z_{i,f,g}$ stands for variables that vary at the individual incident level. We control for year fixed effects (α_t) and for fixed effect for incident types (η_k). Finally, $u_{i,f,g}$ denotes the random error term. For clustering of standard errors we take a conservative approach in our baseline estimation by clustering at the sector level.[15]

Firm size is proxied by the revenues of the firm that suffered a cyber incident.[16] Shih et al. (2000) hypothesise a relationship between firm size and costs stemming from operational risks of the form: $C = R^{\alpha} F(\theta)$, where C denotes costs, R stands for the revenues of the firm, θ for a vector of unobserved risk factors that explain the variation in costs not attributed to revenues, and α for the degree of returns to scale in terms of costs.[17] The authors estimate (in log form) an $\hat{\alpha}$ of 0.15, with a low R^2 (0.05). Replicating this equation, we estimate $\hat{\alpha}$ to be 0.23 and similarly a low R^2 of 0.09. The $\alpha < 1$ indicates a decreasing marginal cost with respect to increases in revenues. Inspection of the residuals of this equation does not indicate obvious signs of heterogeneity across firm size or non-normality of the residuals (see the plot of the residuals in Figure A.4 in the online annex). Contrary to some of the literature (Biener et al., 2015), we do not find evidence in favour of the existence of a U-shape relationship between firm size and the average costs from cyber incidents (details can be found in the annex).

Cyber incidents are likely to exhibit features of contagion: a failure in a firms' IT systems could have spillover effects on other firms (Baldwin et al., 2017; Eisenbach et al., 2020; Crosignani et al., 2020). Incidents that impact multiple firms could contribute to greater costs in the aggregate through other means as well. Affected firms could for instance seek damages and respond by pursuing litigation against the firm at which the incident originated, increasing the costs for the firm that originally suffered the incident. On the other hand, costs could be distributed across firms, thus lowering the average cost across affected firms. We include a variable, *connected events*, that captures how many firms were linked to one specific cyber incident to investigate this effect. To illustrate, if a hacker infiltrated one firm and subsequently managed to penetrate the system of another firm, and both firms recognise they have been affected by the same hacking incident, the *connected events* variable would be 2.[18]

We collect sector level data to estimate the impact of differences in the adoption of information technologies across firms from different sectors. We obtain two variables from the Digital Module of the 2018 Annual Business Survey undertaken by the Bureau of Economic Analysis (BEA). The first variable proxies the *digital share of business activity*. The survey asks: *In 2017, how much of each type of information was kept in digital format at this business?* We collect the percentage of firms that responded that more than 50% of their information is kept in digital format. Therefore, a higher value in this variable indicates

associated to any event, which may either be not easily quantifiable nor publicly reported, even if the event has some cost reported. Furthermore, as discussed earlier the costs of cyber events can also be indirect: these refer to the losses and opportunity costs borne by society as a consequence of a cyber incident, and may include hard-to-quantify damage to the reputation of a firm. Finally, companies may also incur mitigating costs.

[15] The inclusion of $W_{f,g}$ and X_g implies perfect correlation within firm and sector level. Consequently, the error term will be perfectly correlated within clusters, which could lead to bias. This type of clustering has a nested structure, i.e., firm within a sector. The conventional wisdom suggests clustering at the highest level of aggregation, in this case the sector (Cameron and Miller, 2015). We present robustness to alternative clustering choices in the online annex.

[16] We test the robustness of our results by performing regressions using number of employees as an alternative proxy for firm size. Results are unaffected by this choice.

[17] With Eq. (1) we aim to capture some of the unexplained variance (θ) with the inclusion of control variables discussed below. Shih et al. (2000) posit that the unexplained part of this regression could be attributed to variation in firms' attributes regarding risk management, e.g. nature of the business, quality of internal controls, etc. Firm size may implicitly capture the difference in corporate structures and variation in management.

[18] The variable does not provide information on the relationship between the root cause and the affected parties. We note that this variable likely acts as a lower bound on the number of related incidents, as some are unable to be traced to a root cause or may have gone unnoticed or unreported by some firms.

a sector with a stronger dependence on digital technologies for its storage of information. Firms with a higher dependence on IT and digital technologies may expose themselves to more cyber risk (Florakis et al., 2020).[19]

Cloud technologies have become synonymous with cyber risk as policy institutions grapple with the consequences of having centralised IT storage infrastructures. Incidents that involve cloud technology could lead to significant "spillover costs". The survey from BEA also includes data gathered on the penetration of cloud services across sectors. The survey asks: *Considering the amount spent on each of these [IT functions] how much was spent on cloud services [provided by a third party on-demand via the internet]?* We take an average across all IT functions and collect data on firms that indicated 50%–100% was spent on cloud services. The variable proxies an indicator of sectors with a higher exposure to cloud technologies.

Cyber incidents include a broad set of *malicious* and non-malicious events. We test whether cyber attacks (malicious) cause more damage or whether inadvertent incidents are equally damaging. We divide the categorical variable of case types (e.g. DDoS attack, accidental data leak, IT processing error) into two broad categories, malicious and non-malicious, based on whether the incident was done with intent to cause damage or occurred as a result of an accident. Based on this categorisation, we construct a dummy variable labelled *Malicious*, which is equal to one if the event resulted from malicious intent. Around 44% of the incidents recorded fall within this category.

We include in Eq. (1) a set of dummy variables, η_k, for different *types of incidents*, based on the classifications in Romanosky (2016). *Security incident* relates to an incident that compromises or disrupts corporate IT systems (computers or networks) or their intellectual property — examples include hacking and extorting corporate information or a denial of service (DoS) attack. *Data breach* includes unintended disclosure of information (e.g. accidental public disclosure of customer data, improper disposal of information) and/or theft of computers containing personal information of employees or customers of a firm. *Phishing/skimming* are the sending of emails purporting to be from reputable sources in order to convince individuals to reveal personal information to subsequently commit identity theft and the illegal copying of information from the magnetic strips found on credit and debit cards (usually via hardware devices on ATM machines). *Privacy violation* refers to unauthorised collection, use or sharing of personal information — examples include unauthorised collection from cell phones, GPS devices, cookies, web tracking or physical surveillance. This is distinguished from data breaches as an act committed *by* the firm as opposed to *against* the firm. *Other* denote cyber-related losses that were not attributed to one of the above categories.

Table 1 contains summary statistics of the variables used in our regressions. The mean cost incurred by a firm is $10.4 million, with a median of $117,000 and standard deviation of $122 million. This implies a high coefficient of variation and is indicative of the heavy-tailed nature of cyber risks. Cyber risk can be considered a subset of firms' operational risk (Aldasoro et al., 2020). The severity of operational losses is typically characterised by a set of long-tailed distributions, including the log-normal, such that $\ln(C_{i,f,g}) \sim \mathcal{N}(\mu, \sigma^2)$. Figure A.5 in the online annex shows the density of the costs after the log transformation has been applied. There appears to be bi-modality around the mean, but the data are approximately normally distributed.

The mean of the digital share implies that 14.8% of firms across all sectors' maintain more than 50% of their information in digital format. The value ranges from 9%–24%, with sectors at the lower end of the spectrum including Construction and Transportation and Warehousing, and at the top end Manufacturing and Management of Companies and Enterprises. In turn, the average of the cloud variable is 18.6%: roughly a fifth of firms across all sectors spend upwards of 50% of their IT budgets on cloud services. The value ranges from 6 to 26%. Sectors with a lower exposure include Agriculture, Forestry, Fishing and Hunting and Mining, Quarrying, and Oil and Gas Extraction. Those at the other end of the spectrum include Health Care and Social Assistance; Finance and Insurance; Professional, Scientific, and Technical Services; and Information.

Table 1
Summary of variables used in the regression.

	Mean	Median	Std. dev.	Minimum	Maximum
Variables varying at individual event level					
Costs ($ mil)	10.4	0.117	122	0[a]	5000
Connected events	4.90	2.00	9.75	0	79.0
Variables varying at firm level					
Firm size (Revenues $ mil)	12,800	27.0	41,700	0[a]	521,000
Variables varying at sector level					
Digital share of business activity	14.8	15.2	2.28	9.23	24.3
Cloud service purchases	18.6	20.1	5.66	5.60	26.2
Binary variables at event level					
Malicious indicator	0.437	0	0.496	0	1.00
Security incidents	0.0815	0	0.274	0	1.00
Data breaches	0.427	0	0.495	0	1.00
Phishing/Skimming	0.0494	0	0.217	0	1.00
Privacy violations	0.436	0	0.496	0	1.00
Other incidents	0.00648	0	0.0802	0	1.00

Notes:
[a]Zeros are a consequence of rounding accuracy. The top panel reports the variables from Eq. (1) that vary with each individual event in the sample. The second panel contains variables from Eq. (1) that vary by each firm contained in the sample. The third panel are the variables from Eq. (1) that vary at the sector level and obtained from the US census Bureau 2018 Annual Business Survey. The bottom panel are dummy variables that indicate the type of the incident.

4.2. Baseline results

The results of the baseline regressions are presented in Table 2. The cost of a cyber attack is positively correlated with both firm size and the number of connected events. Columns I–III report the baseline regression with and without sector and year effects. We favour the regressions with their inclusion as the coefficients remain robust to unobserved heterogeneity across sectors and variation common to all firms (e.g., the macroeconomic environment).[20] The point estimate of firm size – the logarithm of firm revenues – in Column III is 0.231. A coefficient smaller than 1 suggests the marginal cost is decreasing with respect to revenues, i.e. costs do not increase linearly with the size of the firm.[21] The partial elasticity between firm size and costs implies that for a 1% increase in size there is an increase in the expected cost of 0.23%. Incidents that affect multiple firms – i.e. *connected events* – are similarly associated with higher expected costs: a unit increase in the number of affected firms translates to approximately a 2.6% increase in expected costs.[22] Finally, in column IV we show that these results are robust to the inclusion of more granular sector fixed effects.

[19] The effect of this variable likely manifests in two ways. First, a higher "digital presence" widens the surface of attack to cyber-criminals, which may increase the likelihood of being attacked. Moreover, it suggests which sectors maintain more of their assets in digital format and thus stand to lose more given a cyber incident.

[20] As discussed above the standard errors are clustered at the sector level. The results with Ecker–White errors and firm-level clustering are reported in the online annex. The magnitude of standard errors varies, although this has little impact on the precision of the estimates.

[21] An alternative way to see this is to correct for firm size on the costs variable, i.e. using a ratio of costs to firm size. We present the results of this regression in Table A.4 in the online annex. They confirm that the losses are not proportionate to firm size, and are decreasing relative to firm size.

[22] Revenues, like various measures of firm size, could be heterogeneous across sectors. Sector fixed effects should go some way into controlling for this.

Table 2
The drivers of cyber risk — baseline results.

Dependent variable: Log(Cost)				
Regressor	I	II	III	IV
log(Firm size)	0.241***	0.220***	0.231***	0.220***
	(0.0300)	(0.0228)	(0.0234)	(0.0222)
Connected events	0.0176**	0.0257***	0.0257***	0.0238***
	(0.00740)	(0.00555)	(0.00548)	(0.00661)
Malicious	−1.31***	−1.33***	−1.20***	−1.09***
	(0.230)	(0.179)	(0.207)	(0.324)
Security incident	11.0***	13.0***	13.6***	13.8***
	(0.338)	(0.631)	(0.676)	(0.632)
Data breach	11.6***	14.1***	14.6***	14.8***
	(0.186)	(0.469)	(0.477)	(0.603)
Phishing/Skimming	12.7***	14.7***	15.1***	15.4***
	(0.486)	(0.573)	(0.554)	(0.683)
Privacy violation	10.8***	13.2***	14.0***	14.3***
	(0.387)	(0.708)	(0.755)	(0.701)
Other	12.6***	14.4***	15.3***	15.2***
	(0.405)	(0.636)	(0.666)	(0.895)
Year fixed effects	N	Y	Y	Y
Sector fixed effects	N	N	Y	N
Sub sector fixed effects	N	N	N	Y
R^2	0.11	0.19	0.21	0.25
N	3705	3705	3705	3705

Notes: Results from estimating Eq. (1). *, ** and *** denote significance at the 10, 5 and 1 percent level respectively. Standard errors (reported in parentheses) are clustered by sector. Column I is an OLS regression without controls for Year or Sector fixed effects. Column II is an OLS regression without Sector fixed effects. Column III is an OLS regression including both fixed effects. Column IV replaces the sector fixed effects with the finer categorisation of sub-sectors.

Events with malicious intent are associated with a lower expected cost. Taking the estimate from the third column suggests that on average malicious events are associated with costs 66% lower than other event types.[23] This is perhaps surprising, given the significant press coverage that cyber attacks get and the concern expressed by multiple organisations.[24] Looking more closely at the distribution of costs within each category can provide evidence as to what may be the key driver. In Figure A.6 of the online annex, we show the distribution of costs per case type, malicious versus non-malicious. Security incidents and data breaches are the only case types with variation across both malicious and non-malicious events. Within security incidents in particular there is a stark contrast between the distribution of malicious and non-malicious events, with the latter being significantly more costly.[25] Further, regarding incident types we note that the Other and Phishing/Skimming incidents are, on average, more costly.[26] Finally, while on average the cost of malicious events may be lower, it may still be the case that when focusing at the worst type of events in terms of losses – i.e. when looking at the right tail of losses – malicious events regain prominence. Our finding should thus not be taken as a reason to gloss over the threat that is posed by malicious cyber attacks, as we show in the next section.

4.3. Beware of the tails

Losses stemming from operational and cyber incidents are typically characterised by a set of "heavy-tailed" distributions (Cohen et al., 2019). Therefore, it is reasonable to assume that the conditional distribution is not homogeneous across cost quantiles. Of particular interest in this context is the tail of this distribution, which characterises events of low frequency but high severity. Identifying the features of such events is important to policy-makers and supervisors as they carry the potential to generate substantial economic losses and systemic disruption.

Fig. 4 displays the estimates of the coefficients of *firm size*, *connected* events, and *malicious* events at quantiles ranging between the 0 and 100th percentile. Estimates do vary at different quantiles. The estimate of *firm size* has a lower magnitude at the both ends of the distribution, i.e. it shows an inverted U pattern. *Connected* events have larger estimates towards the upper end of the distribution. Most interestingly, we observe that as the *malicious* variable approaches the 90th percentile the coefficient trends upwards, towards zero, and eventually into positive territory. Malicious events thus do exhibit a significantly different behaviour at the tail-end of the distribution.

Turning the attention to the tail of the distribution, we next present the results of cost quantile regressions between the 95th and 99.5th percentiles. The specification of the regression is analogous to the baseline regression in Table 2 (Column III). *Firm size* and *connected* events are both lower than their mean estimates. The *malicious* indicator has a positive coefficient across all the upper quantiles. A significant effect is observed at the 99.5% level. This result suggests that, ceteris paribus, the tail of the loss distribution is more sensitive to shocks from malicious events. Well coordinated malicious attacks – that happen less frequently – are likely to exceed the costs of non-malicious cyber events. These estimates should be taken with some caution: with limited observations, estimates of what occurs in quantiles can be subject to bias (Chernozhukov and Umantsev, 2001). Nonetheless, uncovering this relationship reveals an important caveat of only studying the central measures of the distribution. While our benchmark regression may show that malicious events are less damaging, sophisticated hacks can actually exacerbate costs at the tail end of the distribution. Understanding the potential damage of high-frequency, low-probability events is paramount from a policy perspective (see Table 3).

5. Digitalisation and cloud-based technologies

Until not so long ago, firms looking to adopt digital technology had to invest in their own data infrastructure and hardware. With the advent of cloud technologies, this has dramatically changed. Cloud technology enables firms to rent computing power and storage from service providers, turning some fixed costs into marginal costs and giving firms more flexibility in handling their operations in a potentially

In untabulated results available upon request, we construct a dummy variable that equals 1 if firm revenues are above the median within that sector, such that we have a within-sector measure of small versus large firms. Including this in the regression confirms the robustness of our original result. When interacting dummies of firm size with specific event types we do not find evidence that firm size plays a role in specific event types.

[23] The percentage change is calculated using the bias correction of Kennedy (1981), $g = \exp(\hat{\beta} - \frac{1}{2}V(\hat{\beta})) - 1$.

[24] A number of other factors may help explain this finding. For one, cyber security actions adopted by many firms protect them from the effects of malicious cyber incidents. There are various well-developed tools that are built to predict and manage cyber attacks, which may be less effective against events that occur as a result of human error inside firms. Moreover, well coordinated cyber attacks can go undiscovered for a long time, in which case the cost of the attack can be difficult to estimate or even identify. Finally, some cyber attacks potentially carry large reputational costs that are hard to quantify and are hence not adequately reflected in loss data.

[25] Non-malicious security incidents include events such as network failures or software bugs that can be very costly. Network outages could be caused by operator errors, surge or usage spike, hardware infrastructure failure, or loss of electrical power. Firms could expect to face 1.6 h of downtime every week, which has been estimated to cost them, on average, $5600 per minute (Knobbe, 2020).

[26] To check if these are significantly different from other variables we test whether the difference between Phishing/Skimming and Data Breaches is statistically significantly larger than zero. Using the results from Table 2 (Column III), we find evidence of statistical significance at the 10% level. Based on this we assume Other and Phishing/Skimming incidents are more costly.

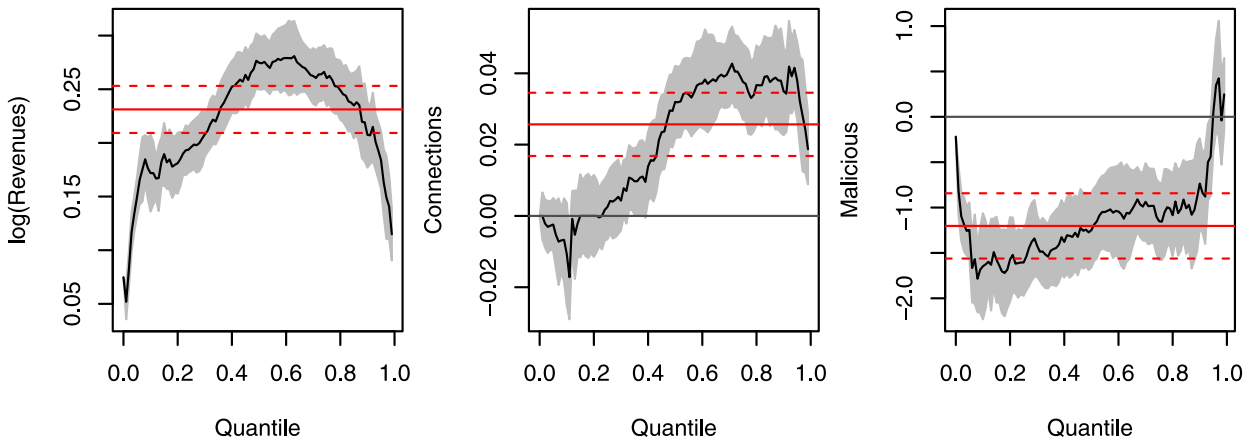

Fig. 4. Selected coefficients from cost quantile regressions. **Notes:** The plot shows the change in selected coefficients as cost quantiles vary. The specification of the regression is analogous to the baseline regression of Table 2 (Column III). The grey shading highlights the 90% confidence interval of the coefficient and red lines denote the estimate of the conditional mean by OLS.

Table 3
Quantile regressions.

Dependent variable: Log(Cost)				
Regressor	95%	97.5%	99%	99.5%
Panel A: Wild Bootstrap				
log(Firm size)	0.184***	0.149***	0.115***	0.153***
	(0.0170)	(0.0167)	(0.0225)	(0.0416)
Connected events	0.0381***	0.0241***	0.0187**	0.00301
	(0.00763)	(0.00756)	(0.00768)	(0.0179)
Malicious	0.0508	0.525*	0.247	0.342
	(0.309)	(0.316)	(0.347)	(1.11)
Panel B: Clustered Bootstrap				
log(Firm size)	0.184***	0.149***	0.115***	0.153***
	(0.0177)	(0.0140)	(0.0202)	(0.00303)
Connected events	0.0381***	0.0241***	0.0187***	0.00301***
	(0.00592)	(0.00370)	(0.00473)	(0.000870)
Malicious	0.0508	0.525	0.247	0.342***
	(0.239)	(0.335)	(0.454)	(0.102)
Year fixed effects	Y	Y	Y	Y
Sector fixed effects	Y	Y	Y	Y
Case type fixed effects	Y	Y	Y	Y

Notes: Results from estimating Eq. (1) at different quantiles. *, ** and *** denote significance at the 10, 5 and 1 percent level, respectively. In Panel A, the standard errors (reported in parentheses) are calculated using the wild bootstrap method; in Panel B standard errors are calculated using the clustered bootstrap method. Both methods are computed using the R Package `quantreg`. For the definition of the regressors, see Table 1.

Table 4
Regressions including the sector level cloud and digital storage variables.

Dependent variable: Log(Cost)					
Regressor	I	II	III	IV	V
log(Firm size)	0.222***	0.221***	0.220***	0.454***	0.450***
	(0.0224)	(0.0224)	(0.0228)	(0.128)	(0.126)
Connected events	0.0256***	0.0256***	0.0258***	0.0253***	0.0254***
	(0.00550)	(0.00549)	(0.00549)	(0.00523)	(0.00525)
Share of digital		−0.0142	0.0554	0.0461	0.114*
		(0.0445)	(0.0484)	(0.0631)	(0.0671)
log(Firm size) × Share of digital				−0.0156*	−0.0154*
				(0.00858)	(0.00846)
Share of cloud	−0.0211		−0.0378**		−0.0371*
	(0.0154)		(0.0188)		(0.0198)
Malicious	−1.33***	−1.33***	−1.33***	−1.34***	−1.34***
	(0.172)	(0.173)	(0.171)	(0.171)	(0.169)
Year fixed effects	Y	Y	Y	Y	Y
Sector fixed effects	N	N	N	N	N
Case type fixed effects	Y	Y	Y	Y	Y
R^2	0.19	0.19	0.19	0.19	0.20
N	3705	3705	3705	3705	3705

Notes: Results from estimating Eq. (1). *, ** and *** denote significance at the 10, 5 and 1 percent level, respectively. All standard errors (reported in parentheses) are clustered by sector to account for the correlation in the sector-level variables. Firm size refers to the firm revenues; Share of digital is the percentage of firms per sector that keep more than 50% of information stored in digital format; Share of cloud is the percentage of firms per sector that keep more than 50% of their data in a dedicated cloud storage. Sector controls are dropped in all regressions due to multicollinearity with the sector-level variables.

more protected environment. This can be particularly advantageous for smaller firms with fewer resources to spend on IT.[27] Cloud computing also exhibits positive externalities such as the reduction of energy consumption and carbon emissions (Etro, 2015). Evidence suggests that firms increasingly take advantage of these benefits, as adoption of digital technology continues to trend upwards (Chen and Srinivasan, 2019).

Digital technologies also pose risks and challenges. Networked production facilities, vehicles, transport infrastructure, and a host of other devices connected to the internet present new opportunities to cyber criminals. The growing complexity of digital infrastructures could increase the likelihood of failures and interruptions, as well as the attendant costs. Cloud service providers have recently drawn the attention of regulators due to the risks associated to their operations, not least given the high degree of concentration in the sector that increases the risk of single points of failure. Tail-risks associated with an outage of a cloud service provider could lead to substantial losses and potentially bring the economy to a halt (Danielsson and Macrae, 2019).[28]

The jury is still out on whether the benefits outweigh the risks, or vice versa. To date, there is little empirical evidence to support either claim. We contribute to this discussion by extending our regression framework with variables that proxy for firms' exposure to digital and cloud services. In particular, we consider *share of digital* and *share of cloud*, which capture sector-level exposure to digital technologies and cloud technology, respectively. To avoid multicollinearity

[27] However, firms are still responsible for the configuration of machines and safe storage of sensitive data while interfacing with external applications.

[28] For a wider discussion of the benefits and risks of cloud computing, see for example Catteddu (2009) and Carr et al. (2019).

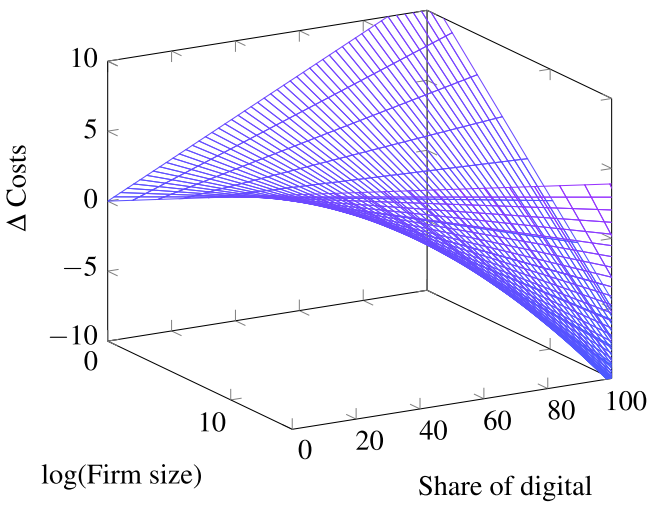

Fig. 5. The expected joint effect of firm size and dependence on digital technology on firms costs.

between sector-level variables and sector fixed effects, we drop the sector controls.

We present the results of the regressions in Table 4.[29] Our results suggest that firms in sectors with a higher exposure to cloud technologies benefit from a mitigating effect on expected costs stemming from cyber incidents. To interpret the magnitude of coefficients, recall how the digital and cloud variables are constructed: the proportion of firms that stated over 50% of their information was stored digitally and of their IT spending was on cloud technology, respectively. The measures loosely reflect the probability that any given firm within a sector has more of its data stored digitally and the probability that the firm has a higher spending on cloud services. The variables are recorded on a 0–100 percentage scale. Consider Column V in Table 4: a one percent increase in the probability of firms spending more than 50% of their IT budgets on cloud services is associated with a reduction of around 4% in expected costs.[30] A stronger dependence on digital services appears to have no statistically significant effect, if considered in isolation (Column II). However, when we add an interaction term between firm size and digital dependence in Column V, we find a mitigating effect.[31] Fig. 5 plots the surface of this mitigating effect on costs as a function of firm size and digital dependence. As firm size increases, more exposure to digital technologies appears to have a mitigating effect on costs. A possible explanation for this relationship could be that as firms expand so do their resources and investment in personnel that ensure that digital technologies are safely maintained.

Overall, a higher exposure to digital and cloud infrastructures is associated with a mitigating effect on costs, which is also firm-size dependent. Given data limitations, it is challenging to derive an exact identification strategy that enables us to define the causal relationship between firms' use of digital technologies and cyber risk. That said, our findings represent a first-pass analysis for a better understanding of the role that digital technologies and in particular cloud technology have to play in shaping cyber risk.

5.1. Revisiting the tails

In this section, we briefly revisit the behaviour of cyber costs at the tails of the loss distribution in relationship to our *digital* and *cloud* variables. Reliance on cloud services in particular has the potential to increase tail-risks (Danielsson and Macrae, 2019). If such risk would be present in our dataset, we would expect to observe a similar relationship to that seen for the *malicious* variable. That is, the coefficients on *cloud* and *digital* become trend upwards as we move right in the distribution, eventually becoming positive.

Fig. 6 shows the estimates for the *cloud* (left) and *digital* (right) variables across quantiles. The estimate for the *cloud* variable shrinks from negative toward zero at the upper quantiles, but does not reach positive territory. The mitigating effect found in the baseline regression is reduced at the tails, but does not turn into a factor that could exacerbate tail risks. The *digital* variable shows a positive relationship with cyber costs that tend to decline towards the upper quantiles (this confirms the sign of the interaction term in Column V of Table 4).

6. Dealing with cyber risk: is current IT investment enough?

In this section we analyse if investment in IT can help to mitigate costs from cyber incidents.[32] Investment in IT security arguably has obvious benefits, yet to date there appears to be little evidence that the continually increasing size of IT budgets and spending are correlated with the mitigation of costs stemming from cyber-related incidents, despite IT security being considered an increasingly critical part of business continuity plans. Evidence that IT spending yields a beneficial return on investment could alter the perspective of firms that are reluctant to invest in IT security as it is viewed as a "sunk" cost. Roner et al. (2021) provide some evidence that cyber security investments can reduce losses arising from a cyber breach. We provide further support to this argument with alternative data.

We use a database constructed by Kennedy and Stratopoulos (2017) based on the InformationWeek IW500 survey. The survey gathers data on IT spending from 500 firms based in the US, across various sectors. The survey focuses on firms that are the most innovative IT users, i.e. to be included in the list a firm had to demonstrate sophisticated use and deployment of IT (Lim et al., 2011).[33] The IW500 dataset provides us with an estimate of firms' IT expenditures as a percentage of revenues. Figure A.8 in the online annex displays the trend in IT spending across sectors for the period 2002–2013. The finance and insurance sector is consistently one of the largest investors in IT, whereas construction and mining are at the lower end of the spectrum. Overall, the investment in IT tends to be relatively stable over time.

Table 5 summarises costs across sectors and the implied spending on IT. To compare the typical annual cost of cyber incidents to firms across sectors, we remove outliers using the interquartile range method (we are interested in these outliers and will return to them later in this section). On average, non-malicious incidents appear to have a higher average costs. Retail Trade and Finance and Insurance are the implied largest spenders on IT. These sectors along with Information and Manufacturing are inferred to spend upwards of a billion dollars annually on IT. In Fig. 7 we show the relationship between spending and the typical annual costs for malicious and non-malicious incidents. For both, there is a weakly implied negative relationship, i.e. higher spending is correlated with lower costs when correcting for average

[29] The case type dummies displayed in previous output are subsumed into the Case Type Fixed Effects indicator.

[30] In untabulated results – available upon request – we also find that more connected events featuring firms with a larger cloud share tend to have larger costs associated to them.

[31] Results are robust to considering an alternative measure of digital dependence, namely output to the digital economy as obtained from the Bureau of Economic Analysis. Results are available upon request.

[32] According to Gartner (2021), worldwide IT spending is projected to total $4.2 trillion in 2021, with information security and risk management technology and services expected to grow 12.4% to reach $150.4 billion.

[33] Previous studies have used the IW500 data to examine the relationship between IT expenditures and various aspects of firm activity and performance. However, there is little evidence on the impact that this investment has towards reducing the risk of losses.

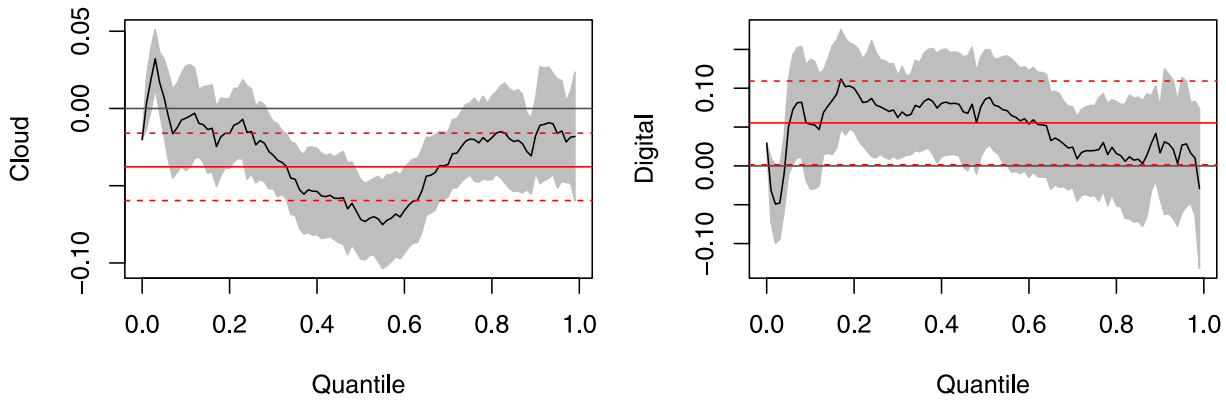

Fig. 6. Selected coefficients from quantile regressions. **Notes:** The graphs show the change in selected coefficients as quantiles vary. The specification of the regression is analogous to the baseline regression of Table 4 (Column III). The grey shading highlights the 90% confidence interval of the coefficients and red lines denote the estimate of the conditional mean by OLS.

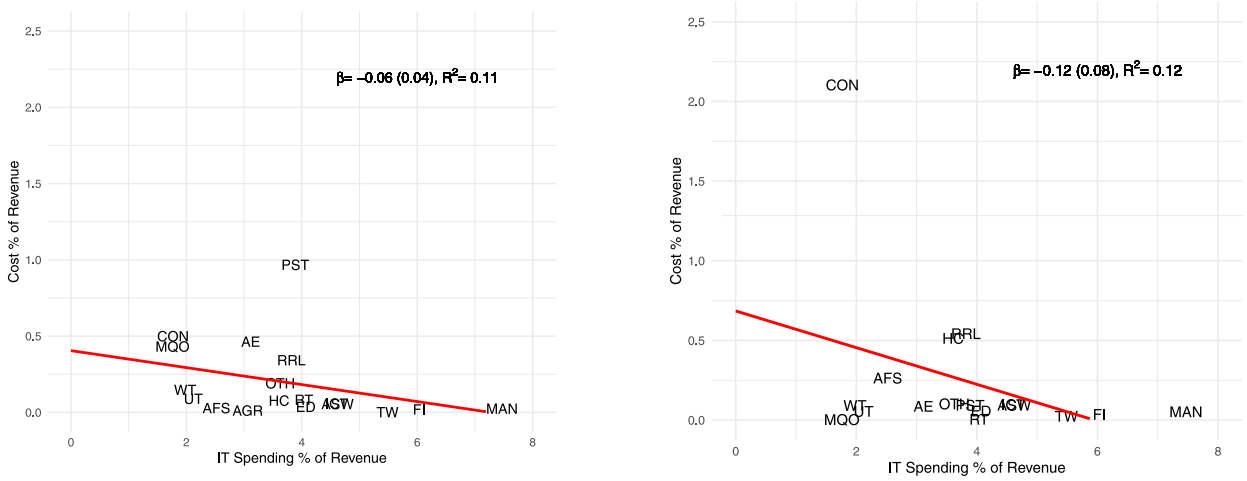

Fig. 7. Spending in IT relative to optimal and costs of cyber events per unit of revenue. **Notes:** The panel on the left hand side shows the average annual costs for malicious incidents by sector normalised by the average revenues for firms within that sector on the *y*-axis. On the *x*-axis, we plot the spending as a percentage of revenues. The panel on the right hand side shows the average annual costs for non-malicious incidents by sector normalised by the average revenues for firms within that sector on the *y*-axis. On the *x*-axis, we plot the spending as a percentage of revenues.

firm size.[34] Sectors with an implied lower spending on IT appear to incur higher costs for malicious relative to non-malicious events. Thus there appears to be evidence that higher levels of IT spending are effective at protecting firms from non-malicious incidents, even if we cannot be certain about causality. For example, investment into new hardware may not be a direct investment in security, but may lead to fewer unintended failures that could be caused by old hardware.

One-off events that lead to significant damage and disruption are of particular interest. In Fig. 8, we show a similar plot to the previous but using the 90th percentile of the distribution of costs across each sector for malicious and non malicious events. First we note a similar relationship, higher spending is associated with lower costs per unit of revenue at the 90th percentile. The estimated regression lines (in red) are similar. Finance and Insurance, Transportation and Warehousing and Manufacturing are some of the highest spenders as a percentage of revenue and appear to benefit from a reduction in relative costs at the mean and at the 90th percentile.

We next look at the difference between spending and costs at firm level. Figure A.9 in the online annex shows the histogram of the difference in percentage of revenues spent on IT and the annual costs of cyber incidents as a percentage of revenues. The bulk of observations are centred around 0, indicating that spending and costs are approximately similar, as also noted by Romanosky (2016). However, we observe a longer left tail that point to the annual cost of cyber incidents sometimes exceeding investment in IT, and occasionally quite significantly. Of course, firms would not expect cyber incidents of such nature to occur with certainty every year and thus the optimal investment requires some balance over time. This leaves open the question of the optimal amount that firms should invest into IT security. The seminal work of Gordon and Loeb (2002) suggest that this should be where the marginal benefit of investment (reduction in costs) is equal to the marginal cost (dollars invested). Such work requires consideration of the distribution of cyber losses for firms and how to incorporate external information from databases similar to the one used within this analysis. We leave this issue for future research.

We attempt to identify if there is a significant effect of IT investment on reducing the costs of a cyber incident.[35] We match our IW500 spending data with the sector and years of observations available in our sample. Data on spending ranges from 2002–2013 therefore we drop observations beyond 2013. We then include the level of spending as

[34] Without normalising the costs by firm size (revenues) we would simply observe the fact that larger sectors have higher costs.

[35] Using UK based data (Roner et al., 2021) show that investments in IT security lead to a reduction in the amount of a loss from a cyber incident.

Table 5
Summary of costs and spending by sector.

Sector	Revenues	Costs (all)	Costs (non-mal)	Costs (mal)	IT spending
Accommodation and Food (AFS)	1512.75	1.58	3.99	0.43	38.12
Admin., Support, WM (ASW)	659.86	0.65	0.62	0.36	30.48
Agriculture (AGR)	82.74	0.01	0.00	0.01	2.53
Arts and Entertainment (AE)	112.11	0.59	0.10	0.52	3.49
Construction (CON)	39.18	0.65	0.82	0.20	0.69
Educational Services (ED)	725.13	0.24	0.41	0.27	29.50
Finance and Insurance (FI)	21,362.75	5.52	7.69	4.18	1288.88
Health Care (HC)	920.09	2.68	4.72	0.71	33.18
Information (ICT)	21,735.79	20.49	21.06	12.49	1000.99
Manufacturing (MAN)	14,483.91	8.56	7.56	3.55	1081.79
Mining (MQO)	1621.28	2.45	0.05	7.00	28.53
Other Services (OTH)	31.73	0.04	0.03	0.06	1.15
Professional, Sci. and Tech. (PST)	388.15	0.97	0.36	3.76	15.09
Real Estate (RRL)	953.15	4.16	5.17	3.26	36.40
Retail Trade (RT)	33,325.43	12.70	1.74	27.63	1344.87
Transportation and Warehousing (TW)	13,208.20	1.69	3.16	0.21	724.57
Utilities (UT)	1892.43	1.37	1.04	1.69	40.14
Wholesale Trade (WT)	1726.54	1.84	1.58	2.55	34.16
Total	6376.73	3.88	3.34	3.83	318.59

Notes: The table summarises revenues, costs and IT spending across sectors. Revenues in the first column denote the average revenue of a firm within each sector. The three cost columns report the average annual cost of cyber incidents incurred by firms in the Advisen database. We remove outliers using the interquartile range method. We report the average for all incidents in the second column and then distinguish between non-malicious and malicious incidents in the third and fourth columns. The final column reports the implied total IT spending by sector (Revenues × IW500 measure). All figures are expressed in millions of US dollars. Sector abbreviations are denoted in parenthesis next to the sector.

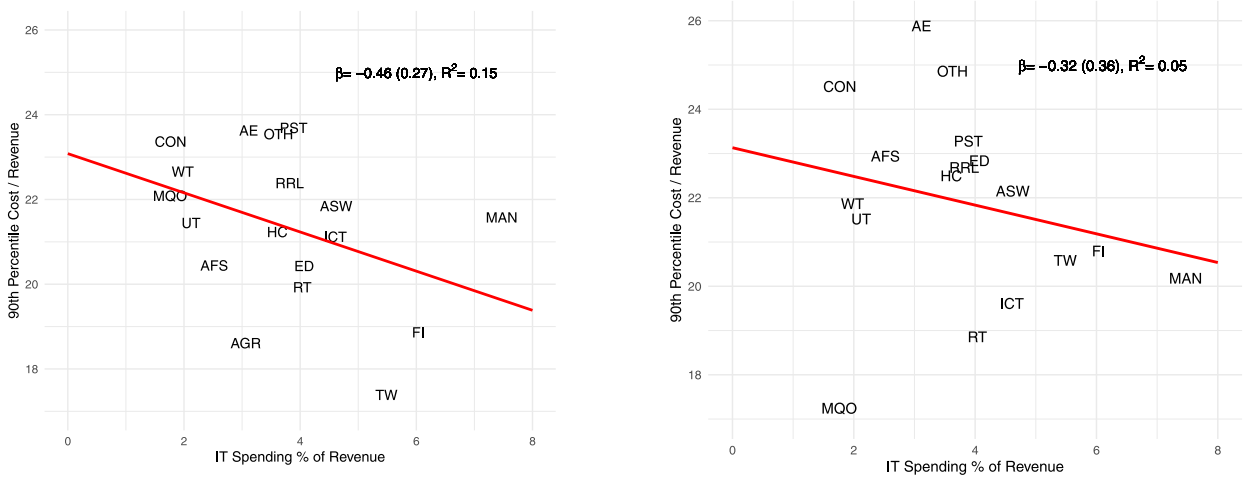

Fig. 8. Spending in IT relative to optimal and costs of cyber events per unit of revenue (90th percentile). **Notes:** The panel on the left hand side shows the logarithm of the 90th percentile annual costs for malicious incidents by sector normalised by the average revenues for firms within that sector on the *y*-axis. On the *x*-axis, we plot the spending as a percentage of revenues. The panel on the right hand side shows the logarithm of the 90th percentile for non-malicious incidents by sector normalised by the average revenues for firms within that sector on the *y*-axis. On the *x*-axis, we plot the spending as a percentage of revenues.

a variable in our regression. Table 6 displays the results. In columns I and II we match the spending year with the year in which the incident occurred. In III, IV, and V we match the year in which the incident with the lag of spending.[36]

The regressions using the contemporaneous spending measure indicate no effect of spending on the costs of cyber incidents. However, when including the lag there is a negative impact (i.e., mitigating effect).[37] The magnitude of the estimate suggests that a 1% increase in IT spending is correlated with a decrease of 34% in costs the subsequent year. This is an important result and also echoes that of Roner et al. (2021). IT investments can take time to mature and the benefits may not be observed immediately, e.g. investment in staff training does not pay dividends until staff have acquired sufficient knowledge. Further, in column V of the table we note that the effect of spending is also strengthened for malicious incidents. An important caveat is in order. These results could give the false impression that simply increasing the size of IT budgets can be a silver bullet to deal with cyber-related losses. This is not necessarily true for two reasons. First, not only the level but also the composition of IT matters. Second, our results cannot properly account for the role of human capital. Expanding investment in IT hardware and related security expenses do not necessarily help if staff are not properly trained and capabilities are not fit for purpose.

7. Conclusions

The digital revolution has increased the interconnectivity and complexity of the economic system. The use of technology and internet have improved firms' productivity, but also exposes them to cyber attacks. Moreover, the greater use of cloud services exposes further important economic sectors to common risks.

[36] Here we gain observations from 2014 and lose those from 2002, hence the changing sample size in the regressions.

[37] Additional lags of IT spending are not statistically significant, so we present results with only one lag.

Table 6
Regressions including the sector level cloud and digital storage variables.

Dependent variable: Log(Cost)					
Regressor	I	II	III	IV	V
log(Firm size)	0.228***	0.228***	0.228***	0.228***	0.230***
	(0.0232)	(0.0232)	(0.0197)	(0.0197)	(0.0190)
Connections	0.0159	0.0159	0.0245***	0.0245***	0.0251***
	(0.0101)	(0.0101)	(0.00679)	(0.00679)	(0.00708)
Malicious	−1.12***	−1.12***	−1.02***	−1.02***	
	(0.350)	(0.350)	(0.288)	(0.288)	
IT spending	−29.2	−29.2			
	(30.5)	(30.5)			
IT spending lag			−41.3*	−41.3*	−35.6*
			(22.6)	(22.6)	(21.5)
IT spending lag × Malicious					−22.9***
					(7.69)
Share of cloud		0.0217		−1.13***	−1.13***
		(0.158)		(0.135)	(0.128)
Share of digital		−0.286		1.34***	1.37***
		(0.188)		(0.152)	(0.153)
Year fixed effects	Y	Y	Y	Y	Y
Sector fixed effects	Y	Y	Y	Y	Y
Case type fixed effects	Y	Y	Y	Y	Y
R^2	0.2	0.2	0.2	0.2	0.21
N	2611	2611	2953	2953	2953

Notes: Results from estimating Eq. (1). *, ** and *** denote significance at the 10, 5 and 1 percent level, respectively. All standard errors (reported in parentheses) are clustered by sector to account for the correlation in the sector-level variables. Firm size refers to the firm revenues; Share of digital is the percentage of firms per sector that keep more than 50% of information stored in digital format; Share of cloud is the percentage of firms per sector that keep more than 50% of their data in a dedicated cloud storage. IT spending denotes the percentage of revenue spent on IT within each sector per year.

Despite the large and growing exposure to cyber risks, cyber costs are difficult to quantify. Using a unique database at the firm level for the US, we document the characteristics of cyber incidents and help quantify cyber risk. The average cost of cyber events has increased over the last decade. These costs are higher for larger firms and more connected events, and relatively lower for cyber events with malicious intent (cyber attacks), but only if the attack is not conducted on a large scale: malicious events can be more costly in the upper tail of the distribution.

The financial sector experiences the highest number of cyber incidents (especially of a malicious type, privacy and lost data incidents). However, banks and insurance companies incur more limited losses relative to other sectors, likely due to the effects of regulation and higher investment in cyber security.

We document that a higher exposure to digital technologies helps firms mitigate the costs of cyber incidents, as does more reliance on cloud services. This last result should be taken with caution and qualified. As cloud connectivity increases and cloud providers become systemically important, cloud dependence is also likely to increase tail risks.

Finally, we document some evidence on the effect that spending on IT has on the costs of cyber incidents. We observe a negative relationship between spending in IT and the cost of cyber-events. This result provides some evidence of the "unobserved" return on investment into IT and security and may encourage firms that are reluctant to invest into IT, as the returns on additional expenditures are hard to measure. While our analysis does not account for the systemic implications of failures in specific critical sectors, the results can inform policymakers as to where to direct their attention in order to improve the economy's overall cyber resilience.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jfs.2022.100989.

References

Aldasoro, I., Gambacorta, L., Giudici, P., Leach, T., 2020. Operational and Cyber Risks in the Financial Sector. BIS Working Papers 840, Bank for International Settlements.

Anderson, R., Barton, C., Böhme, R., Clayton, R., Hernandez Ganan, C., Grasso, T., Levi, M., Moore, T., Vasek, M., 2019. Measuring the changing cost of cybercrime. In: The 2019 Workshop on the Economics of Information Security (WEIS 2019).

Baldwin, A., Gheyas, I., Ioannidis, C., Pym, D., Williams, J., 2017. Contagion in cyber security attacks. J. Oper. Res. Soc. 68 (7), 780–791.

Basel Committee on Banking Supervision, 2018. Cyber resilience: Range of practices.

Biener, C., Eling, M., Wirfs, J.H., 2015. Insurability of cyber risk: An empirical analysis. Geneva Pap. Risk Insur.-Issues Pract. 40 (1), 131–158.

Bouveret, A., 2018. Cyber Risk for the Financial Sector: A Framework for Quantitative Assessment. IMF Working Paper WP/18/143, International Monetary Fund.

Brenner, J., 2017. Keeping America Safe: Toward More Secure Networks for Critical Sectors. Report on a Series of MIT Workshops, MIT Internet Policy Research Initiative.

Cameron, A.C., Miller, D.L., 2015. A practitioner's guide to cluster-robust inference. J. Hum. Resour. 50 (2), 317–372.

Carr, B., Pujazon, D., Vazquez, J., 2019. Cloud Service Providers and Criticality: Potential Treatments and Solutions. Technical Report, Institute of International Finance.

Carrivick, L., Cope, E.W., 2013. Effects of the financial crisis on banking operational losses. J. Oper. Risk 8 (3), 3.

Catteddu, D., 2009. Cloud computing: benefits, risks and recommendations for information security. In: Iberic Web Application Security Conference. Springer, p. 17.

Chande, N., Yanchus, D., 2019. The Cyber Incident Landscape. Working Paper 32, Bank of Canada.

Chen, W., Srinivasan, S., 2019. Going Digital: Implications for Firm Value and Performance. Working Paper 19–117, Harvard Business School.

Chernobai, A., Jorion, P., Yu, F., 2011. The determinants of operational risk in US financial institutions. J. Financ. Quant. Anal. 46 (6), 1683–1725.

Chernozhukov, V., Umantsev, L., 2001. Conditional value-at-risk: Aspects of modeling and estimation. Empir. Econ. 26 (1), 271–292.

Cohen, R.D., Humphries, J., Veau, S., Francis, R., 2019. An investigation of cyber loss data and its links to operational risk. J. Oper. Risk 14 (3), 1–25.

Crosignani, M., Macchiavelli, M., Silva, A.F., 2020. Pirates without Borders: The Propagation of Cyberattacks through Firms' Supply Chains. Staff Report 937, Federal Reserve Bank of New York.

Curti, F., Mihov, A., Frame, W.S., 2019. Are the largest banking organizations operationally more risky?. URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3210206.

Danielsson, J., Macrae, R., 2019. Systemic consequences of outsourcing to the cloud. In: VoxEU. CEPR.

Duffie, D., Younger, J., 2019. Cyber Runs. Hutchins Center Working Paper 51, Brookings Institution.

Eisenbach, T.M., Kovner, A., Lee, M.J., 2020. Cyber Risk and the US Financial System: A Pre-Mortem Analysis. Staff Report 909, Federal Reserve Bank of New York.

Etro, F., 2015. The economics of cloud computing. In: Cloud Technology: Concepts, Methodologies, Tools, and Applications. IGI Global, pp. 2135–2148.

Etzioni, A., 2011. Cybersecurity in the private sector. Issues Sci. Technol. 28 (1), 58–62.

Financial Stability Board, 2019. Third-party dependencies in cloud services: Considerations on financial stability implications.

Florakis, C., Louca, C., Michaely, R., Weber, M., 2020. Cybersecurity Risk. Working Paper 28196, National Bureau of Economic Research, http://dx.doi.org/10.3386/w28196, URL: http://www.nber.org/papers/w28196.

Gartner, 2021. Gartner forecasts worldwide security and risk management spending to exceed $150 billion in 2021. URL: https://www.gartner.com/en/newsroom/press-releases/2021-05-17-gartner-forecasts-worldwide-security-and-risk-managem.

Goldstein, J., Chernobai, A., Benaroch, M., 2011. An event study analysis of the economic impact of IT operational risk and its subcategories. J. Assoc. Inf. Syst. 12 (9), 1.

Gordon, L.A., Loeb, M.P., 2002. The economics of information security investment. ACM Trans. Inf. Syst. Secur. 5 (4), 438–457.

Kamiya, S., Kang, J.-K., Kim, J., Milidonis, A., Stulz, R.M., 2018. What is the Impact of Successful Cyberattacks on Target Firms?. NBER Working Paper 24409.

Kamiya, S., Kang, J.-K., Kim, J., Milidonis, A., Stulz, R.M., 2021. Risk management, firm reputation, and the impact of successful cyberattacks on target firms. J. Financ. Econ. 139 (3), 719–749. http://dx.doi.org/10.1016/j.jfineco.2019.05.019, URL: https://www.sciencedirect.com/science/article/pii/S0304405X20300143.

Kashyap, A.K., Wetherilt, A., 2019. Some principles for regulating cyber risk. AEA Pap. Proc. 109, 482–487.

Kennedy, P.E., 1981. Estimation with correctly interpreted dummy variables in semilogarithmic equations. Amer. Econ. Rev. 71 (4), 801, URL: http://www.jstor.org/stable/1806207.

Kennedy, D.B., Stratopoulos, T.C., 2017. Mapping IT spending across industry classifications: An open source dataset. URL: https://ssrn.com/abstract=3073236.

Knobbe, D., 2020. Network outages: Do they cost more than you think?. URL: https://info.pivitglobal.com/blog/cost-of-downtime.

Kopp, E., Kaffenberger, L., Jenkinson, N., 2017. Cyber Risk, Market Failures, and Financial Stability. IMF Working Paper WP/17/185, International Monetary Fund.

Lim, J.-H., Stratopoulos, T.C., Wirjanto, T.S., 2011. Path dependence of dynamic information technology capability: An empirical investigation. J. Manage. Inf. Syst. 28 (3), 45–84.

Makridis, C., 2020. Do data breaches damage reputation? Evidence from 43 companies between 2002 and 2018. URL: https://ssrn.com/abstract=3596933.

Makridis, C., Dean, B., 2018. Measuring the economic effects of data breaches on firm outcomes: Challenges and opportunities. J. Econ. Soc. Meas. 43 (1–2), 59–83.

Makridis, C., Liu, T., 2021. Abnormal returns and dispersion in cybersecurity exposure. URL: https://ssrn.com/abstract=3746589.

Romanosky, S., 2016. Examining the costs and causes of cyber incidents. J. Cybersecur. 2 (2), 121–135.

Roner, C., Caterina, C.D., Ferrari, D., 2021. Exponential Tilting for Zero-inflated Interval Regression with Applications to Cyber Security Survey Data. Faculty of Economics and Management at the Free University of Bozen, URL: https://ideas.repec.org/p/bzn/wpaper/bemps85.html.

Rowe, B., 2007. Will outsourcing IT security lead to a higher social level of security? In: Workshop on Economics of Information Security. Pittsburgh.

Rowe, B.R., Gallaher, M.P., 2006. Private sector cyber security investment strategies: An empirical analysis. In: The Fifth Workshop on the Economics of Information Security. The Workshop on the Economics of Information Security.

Shih, J., Samad-Khan, A., Medapa, P., 2000. Is the size of an operational loss related to firm size. Oper. Risk 2 (1), 21–22.

Wolff, J., Lehr, W., 2017. Degrees of Ignorance about the Costs of Data Breaches: What Policymakers Can and Can't Do about the Lack of Good Empirical Data. Technical Report, Available at SSRN 2943867.