DISCUSSION PAPER SERIES

DP14805
(v. 2)

THE DRIVERS OF CYBER RISK

Inaki Aldasoro, Leonardo Gambacorta, Paolo Giudici and Thomas Leach

FINANCIAL ECONOMICS

ISSN 0265-8003

THE DRIVERS OF CYBER RISK

Inaki Aldasoro, Leonardo Gambacorta, Paolo Giudici and Thomas Leach

Discussion Paper DP14805
First Published 25 May 2020
This Revision 31 May 2020

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

This Discussion Paper is issued under the auspices of the Centre's research programmes:

- Financial Economics

THE DRIVERS OF CYBER RISK

Abstract

Cyber incidents are becoming more sophisticated and their costs difficult to quantify. Using a unique database of more than 100,000 cyber events across sectors, we document the characteristics of cyber incidents. Cyber costs are higher for larger firms and for incidents that impact several organisations simultaneously. The financial sector is exposed to a larger number of cyber attacks but suffers lower costs, on average, thanks to proportionately greater investment in information technology (IT) security. 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. Crypto-related activities, which are largely unregulated, are particularly vulnerable to cyber attacks.

JEL Classification: D5, D62, D82, G2, H41

Keywords: cyber risk, cloud services, financial institutions, Bitcoin, cryptocurrencies, cyber cost, cyber regulation

Inaki Aldasoro - inaki.aldasoro@bis.org
Bank for International Settlements

Leonardo Gambacorta - leonardo.gambacorta@bis.org
Bank for International Settlements and CEPR

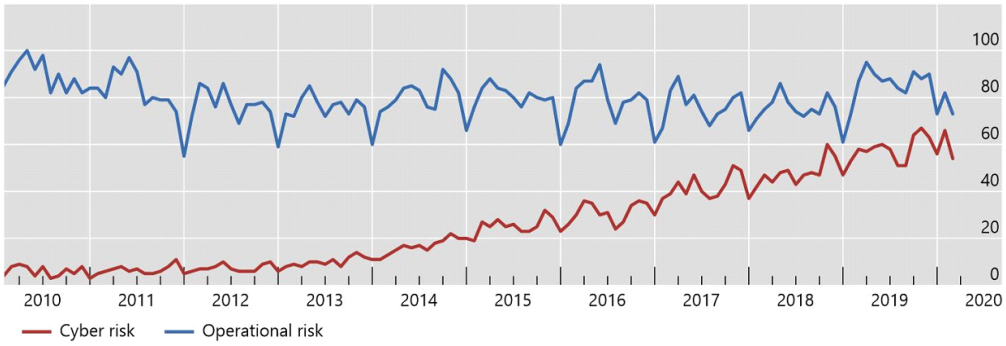
Paolo Giudici - paolo.giudici@unipv.it
Department of Economics and Management, University of Pavia, Italy

Thomas Leach - thomas.leach01@universitadipavia.it
University of Pavia

Acknowledgements
We thank Raymond Kleijmeer, Karin Reichardt, Andreas Voegtli, an anonymous referee and seminar participants at the 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. The views expressed here are those of the authors and do not necessarily represent those of the Bank for International Settlements.

I. 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. 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. 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). In the wake of recent high-profile cyber attacks such as the WannaCry incident in May 2017, public awareness of these threats is on the rise (see Figure 1).



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.

Figure 1
Interest on cyber risk is on par with operational risk.

Firms actively manage cyber risk and invest in cyber security. However, cyber costs are difficult to quantify. In the financial sector, cyber risks are a key "known unknown" tail risk to the system and a potential major threat to financial stability.[1] More broadly, cyber risk in sectors that play a critical role in the economic infrastructure could have systemic implications and should be viewed as a matter of national security (Brenner, 2017). Despite the acknowledgement of such consequences, information concerning the costs, drivers and potential mitigating factors of cyber incidents is relatively scarce.

This paper seeks to help fill this gap. The analysis uses a detailed dataset of over 100,000 cyber events across all sectors of economic activity. We first document some stylised facts. The frequency of cyber incidents rose strongly in the decade to 2016, but has since receded somewhat. This reduction could reflect

[1] 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".

increased investment in cyber security, but also delays in discovery or reporting.[2] The average cost of cyber events has been increasing constantly over the last decade. 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, privacy violations and phishing/skimming scams fraud in short are the most frequent but least costly. Data breaches, in turn, are both relatively frequent and costly, while business disruptions are quite infrequent but can have high costs.

The richness of the database also allows us to examine the relationship between firm, sector and event-specific characteristics and the relative cost of cyber events. The main empirical results can be summarised as follows.

First, we 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. We also find that that cyber events which impact multiple firms at the same time (i.e. "connected" events) are also associated with higher costs. Cyber-related incidents can occur unintentionally by human error, e.g. a bug in some internally developed software; or can also be caused by an actor with malicious intent.[3] Malicious cyber attacks have, on average, lower costs, because most incidents simply reflect general discontent. However, some actors seek a profit or to inflict the largest possible losses and damage. Indeed, 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. Incidents related to crypto exchanges, which are largely unregulated, produce higher losses.

We then look at the role of developing technological capabilities to reduce cyber costs. Many firms, especially if they are smaller, could lack the specific knowledge needed to make rational decisions about which software or cyber security provider to choose. Information asymmetries between firms can further exacerbate problems when investments in new technologies do not pan out as expected (Zhu and Weyant, 2003). How can firms mitigate these risks? We find that investment in technological skills pays off in terms

[2]This phenomena is widely recognised in the operational risk literature (see Aldasoro et al. (2020); Carrivick and Cope (2013). The dataset used in this analysis did not contain sufficient information concerning the dates of events, thus an "end-of-sample" bias could not be accurately estimated.

[3]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. The discovery of a zero-day exploit can result in a situation where both the customers and vendors of the IT asset are now subject to cyber attacks for which no pre-defined detection signatures or remedial patches are available. Exacerbating this situation are commercial firms that conduct research to sell zero-day exploits on the open market. Some of these firms, such as Zerodium, pay large cash rewards (up to $2.5 million) for high-risk vulnerabilities.

of reducing the cost of cyber events. In particular, firms in sectors that employ more IT specialists, use more computers and provide more IT training to staff, are better equipped to mitigate the costs stemming from a cyber event. From a policy perspective, these findings can inform governments and cross-sector regulators regarding the mitigating steps that can be taken to reduce the cost of cyber incidents and which sectors are lacking in such areas.

Cybersecurity activities provided by third-party service providers are an alternative to risk transfer mechanisms. Rowe (2007) argues that, if multiple organisations share the same service provider, economies of scale and information-sharing can create positive externalities. Cloud technology can reduce IT costs, improve resilience and enable firms to scale better (Financial Stability Board, 2019). However, the technology strengthens interdependence across firms that have shared exposures to similar (or even the same) cloud service providers. For example, several cloud suppliers may use a common operating system so that, if the operating system has a vulnerability, it could create a correlated risk across all cloud suppliers. 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 increases and cloud providers become systemically important, cloud dependence is likely to increase tail risks (Danielsson and Macrae, 2019).

Of particular concern is the exposure to cyber risk of financial institutions and infrastructures, given the critical services they provide (Kopp et al., 2017; Committee on Payments and Market Infrastructures, 2014). Following the financial crisis, banks in particular became a target for activists.[4] We interact a finance sector dummy variable with our baseline regressors to assess average costs of losses relative to other sectors. While the frequency of attacks in the financial sector is high relative to others, the sector is better at mitigating the cost of attacks. This could be the outcome of more proactive policy, regulation and investment in risk management and governance practices with respect to information technology.

Cryptocurrencies have emerged as a challenge to established financial institutions and currencies. Despite initial claims of superior security, the cryptocurrency space has suffered numerous cyber attacks. This notoriety stems both from attacks on crypto-exchanges due to poor security standards and due diligence on internal controls,[5] as well as from the use of cryptocurrencies as ransomware that is difficult to trace, e.g. WannaCry (Kshetri and Voas, 2017). We find that the average cost of crypto-related events is significantly higher. These costs are not independent of the soaring price of cryptocurrencies in recent years. We document the existence of a strong positive correlation between the price of bitcoin and the intensity of attacks on crypto-exchanges. To quantify this relationship, we use a Probit model to show that an increase in the price of bitcoin increases the likelihood of future attacks on crypto-exchanges. However, the inverse relationship is not found to be significant, i.e. there is no price decrease following cyber incidents related to cryptocurrencies.

Finally, we use data on the level of IT spending across sectors to assess which sectors may be over- or

[4]For a list of attacks on banks see: https://carnegieendowment.org/specialprojects/protectingfinancialstability/timeline.

[5]Analytics firm Chainalysis reported that approximately $1 billion worth of cryptocurrency was stolen from digital currency exchanges in 2018 (https://news.bitcoin.com/report-two-hacker-groups-stole-1-billion-from-crypto-exchanges/).

underspending on their IT security. This analysis can act as a helpful indicator to policymakers as to which sectors may be exposed due to underinvestment in IT systems. We find that, across all sectors, there is a deficit in IT-spending. This is concerning, as the threat of cyber-related incidents is likely to increase in the future. Some sectors, however, exhibit an adequate level of spending. Such is the case of the finance and insurance sector. The dividend of such investments is evident through our other analyses, whereby the observed cost of attacks for that sector is lower. This result does not provide a reason for firms in sectors that are in a spending surplus to cut back on their investments, as the threat landscape continues to evolve.

The rest of the paper is organised as follows. Section II discusses related literature. Section III contains a description of the data. Section IV discusses our baseline results. Section V looks at the impact of technological skills on cyber losses. Section VI explores whether exposure to cloud services affects the cost of cyber events. Section VII examines the relationship between the price of bitcoin and the hacking of crypto exchanges. Section VIII zooms in on the financial sector. Section IX analyses the optimal amount of IT spending across sectors. Finally, Section X summarises the main conclusions.

II. Related literature

Cyber- and IT-related risks can be seen as a relevant component of operational risks. Costs are difficult to quantity, however, given the absence of high-quality data. Establishing policies to encourage long-term data collection about incidents and security breaches is crucial in informing effective policies for security actions and outcomes (Wolff, 2014).[6] Recent policy initiatives represent an encouraging step forward,[7] but information on cyber incidents and cost is still limited.

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, or the risk of failures of operational IT systems, could translate into significant losses in firms' market value. 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.

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)

[6]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).

[7]The European Union recently passed the Network and Information Security Directive (NISD). Meanwhile, the U.S. Congress passed the Cybersecurity Information Sharing Act (CISA). These new initiatives aim to encourage, or require, that more information about cyber security incidents be shared with, or reported to, entities other than the ones who detect those incidents. Due to the new class of threats and corresponding insurance needs, frequent updating of the reporting requirements is needed to maintain an accurate picture of the cyber threat landscape (Wolff and Lehr, 2018; Rowe and Gallaher, 2006).

find that cyber attacks are more likely in industries that face less intense product market competition and in industries 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) also 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.

Regulation can also play a key role in firms' motives for security investments. Using the results of a survey conducted on more than 700 firms, Rowe and Gallaher (2006) find that the vast majority of reporting firms 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 by reducing the costs of cyber incidents.

Some sectors provide critical infrastructure for the functioning of the economy. Cyber attacks on the financial sector could create possible cascade failures that are not completely understood or adequately quantified by sector-specific simulations (Brenner, 2017). Kopp et al. (2017)) note that the financial sector is a frequently targeted sector due to its high exposure to IT and its credit intermediation role. 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.[8] 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.[9] This amounts to around a third of operational VaR, despite representing a minor share of the latter in terms of frequencies and loss amounts. 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. Credit booms and periods of accommodative monetary policy are associated with higher operational losses in the future, but have no effect on cyber losses.

Duffie and Younger (2019) analyse a sample of twelve systemically important U.S. financial institutions and suggest these firms have sufficient stocks of high-quality liquid assets to cover wholesale funding runoffs in a relatively extreme cyber event. From the literature on operational risk, the size of financial institutions is positively linked with the size 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). Due to

[8] See Basel Committee on Banking Supervision (2018)

[9] 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.

its special nature, in Section VIII, we devote particular attention to the drivers of cyber risks in the financial sector and how these could differ from other economic sectors.

Another element that could alter the relative cost of a cyber attack is the reliance on cloud services. Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction (Mell and Grance, 2011). Firms are benefiting from cheap computing resources, although as more firms become dependent on clouds, this increases risk correlation among them. Through shared software, hardware and vendors, incidents could spread more quickly, leading to higher overall costs (Welburn and Strong, 2019).

Lloyds (2017) suggests that economic losses due to cloud service disruption can vary. The average loss of such disruptions can be between US$15.6 billion and US$53 billion, depending on factors such as the different organisations involved and for how long the cloud service disruption lasts. Lloyds (2018) reports that, given the state of the cyber insurance industry today, a cyber incident that takes a top-three cloud provider offline in the US for 3 to 6 days would result in losses to policy holders between US$6.9 billion and US$14.7 billion.

A few studies also try to evaluate the potential systemic costs of cyber incidents by making a number of assumptions on interlinkages among sectors and cascade effects. Dreyer et al. (2018) estimate the direct gross domestic product costs of cyber crime to be between US$275 billion and US$6.6 trillion, using a sample of 60 countries. Romanosky (2016) estimates the annual cost of cyber incidents to be approximately US$8.5 billion.

A novel part of the literature looks at the occurrence of cyber attacks on cryptocurrency exchanges due to their relatively poor security standards. Moore and Christin (2013) find that the average transaction volume on bitcoin exchanges is negatively correlated with the probability they will close prematurely and positively correlated with the probability of experiencing a breach. We further develop this aspect by analysing the correlation between the price of bitcoin and the intensity of attacks on crypto-exchanges.

III. Data

A. Cyber events

Our dataset includes 115,415 cyber incidents. 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.[10] Due to the nature of how the data are collected, it is not possible to obtain all information desirable for each event. For example, the total loss amount could be disclosed by the public source from which it was collected, but in many cases firms do not to release details of their losses due to

[10]Most cyber incidents go unreported. Typically, only the larger and the more relevant ones become public and are included in the Advisen database.

reputational concerns (Biener et al., 2015; Pretty, 2018). We return to this particular issue later (Section IV), but it is important to recognise upfront that whilst there are many observations in the database, one drawback is that several variables are sparsely populated.

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); iv) 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 losses reported in the database occur in the Americas region (North, Central and South America; see Table I). 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 in the US as a result of a higher degree of freedom of information. The region with the next largest occurrence of cyber events is Europe and, lastly, Africa. Data limitations imply these statistics are not to be taken as a representation of the population of cyber incidents that occur globally.

	Asia	Africa	Americas	Europe	Oceania	Total
Frequency	3739	345	102459	7627	1024	115415
Total losses	4079.95	1793.37	28988.08	4299.89	362.41	39523.82
Mean loss	19.43	99.63	2.06	6.29	4.53	2.62
Standard deviation of loss	89.34	399.53	47.41	44.36	19.64	49.97

Notes: All numbers referring to losses are reported in USD millions. The mean and standard deviations are calculated only over the observations that have data for the loss amount available. 221 incidents could not be assigned to a country / region.

Table I
Summary of geographic location of incidents

Table II contains summary statistics of the frequency and losses by sector.[11] By frequency, "Financial and insurance activities" (FI) is the most affected. This sector is typically targeted due to lucrative rewards from a successful attack. However, the FI sector shows some resilience, as despite being subject to many attacks, the mean cost of a cyber incident is not as high as in other sectors. The sector with the highest average costs is "Transportation and storage", followed by "Wholesale Trade" and "Professional, Scientific and Technical" (PST). The standard deviation of costs in these sectors is quite large, implying that most likely the distribution of losses has a heavy tail. In terms of overall costs, the "Information and Communication Technology" (ICT) sector ranks first.

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

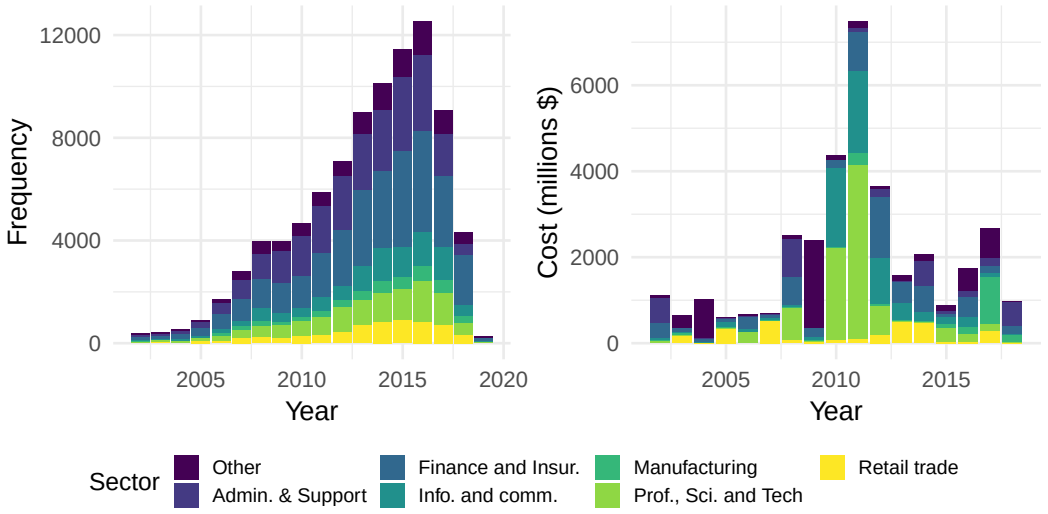
	Frequency	Total loss	Mean loss	Std. dev. of loss
Accommodation, food and beverage	2829	496.91	2.22	7.31
Administrative and support service	22062	3477.81	0.78	17.83
Agriculture, forestry, fishing and hunting	80	0.01	0.01	0.01
Arts, entertainment, and recreation	969	147.22	1.94	5.27
Construction	1020	358.66	2.49	23.42
Educational services	4631	147.26	0.57	2.34
Financial and insurance activities	27792	6321.09	1.69	15.45
Healthcare and social assistance	8688	597.80	1.06	9.55
Information and communication	9524	6402.59	5.51	76.22
Management of companies and enterprises	437	259.62	4.12	13.95
Manufacturing	3926	2006.67	5.61	32.24
Mining, quarrying, and oil and gas	126	8.92	0.52	1.69
Other services	2655	300.42	1.28	7.24
Professional, scientific and technical	10629	8777.63	5.54	114.79
Public administration	6704	2145.67	4.25	50.08
Real estate activities	1370	110.97	0.78	3.09
Retail trade	6452	2889.54	3.44	23.78
Transportation and storage	1154	2201.43	16.80	90.68
Utilities	530	36.13	0.38	1.39
Wholesale trade	2422	2202.03	8.01	120.60
Total	115415	39523.82	2.62	49.97

Notes: All numbers referring to losses are reported in USD millions. The mean and standard deviations are calculated only over the observations that have data for the loss amount available. In addition, 1415 incidents were not assigned a sector.

Table II
Summary of Advisen data by sector

Figure 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 due to their role in the crisis. The peak in costs in 2011 was shared largely amongst the FI, PST and ICT sectors. This coincides with a year of excessive losses due to data breaches (Figure 3), which suggests that multiple firms across sectors may have been affected by cyber incident that resulted in a loss of data either for the firm or personal records.

Following the categorisation of Bouveret (2018), we present in Table III a summary of losses and the frequency by categories of event types. The bulk of cyber incidents are related to data breaches and fraud. Together, these account for over 100,000 of the events in the sample. However, for observations that contain information on loss amounts, business disruptions seem to cause the most damage on average, where the mean loss is twice as large as for data breaches. While a business disruption is likely to have a large "direct" costs for the firm, the reputational costs of a data breach may be much higher and difficult to calculate, as several people may be affected as a result of the breach.



Notes: The graph contains a breakdown that reflects the sectors that are highlighted in our regression. See Section IV for more details. The Other category contains Construction; Wholesale trade; Real estate activities; Accommodation, food and beverage; and Transportation and storage.

Figure 2

Distribution of frequency and costs of cyber incidents by sector over time

For comparison, we also present in Table IV summary statistics using an alternative definition adopted by Romanosky (2016). Here there is a nuanced difference related to privacy events. Privacy violations refer to when the firm is actually responsible for the incident, i.e. it has committed a breach of individuals' privacy and not incurred the incident itself. In the categorisation presented in Table III, privacy violations were grouped together with fraud. For the other categories, security incidents are more or less analogous to business disruptions, phishing and skimming are a fraction of fraud as the larger component was made up by privacy violations under the previous definition. Data breaches are defined similarly to the previous definition. In terms of frequency, privacy violations are the most commonly occurring incident (44 per cent of the cases). This is likely due to the fact that reporting requirements having been in place for a longer period for privacy incidents than for other types of incident. Moreover, it is easier to assign conclusive responsibility of the incident in privacy cases than in others (Chande and Yanchus, 2019). 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.

	Data breach	Business disruptions	Fraud	Other	Total
Frequency	53500	4915	56308	692	115415
Total losses	19155.30	8657.35	11679.12	32.04	39523.82
Mean loss	9.97	26.48	0.91	1.69	2.62
Standard deviation of loss	119.23	123.77	19.71	2.57	49.97

Notes: All numbers referring to losses are reported in USD millions. The mean and standard deviations are calculated only over the observations that have data for the loss amount available. The Data breach category contains incidents that resulted in a loss of data either for the firm or personal records; The Business disruptions category contains items such as hacking or denial of service (DoS) attack; The Fraud category contains privacy violations and phishing/skimming incidents.

Table III
Summary by type of cyber incident

Figure 3 depicts the distribution by case type through time.[12] The most recent end of the data is probably subject to an underreporting bias, as it takes time for incidents to be discovered and acknowledged. Therefore, we expect the numbers from 2016 onwards to increase as more information becomes available in the future. There has been an increasing trend in terms of frequency, in line with the growing concern over cyber risks. This is likely driven by a few factors. First, several frameworks and legislation have come into place that encourages the reporting of cyber incidents. Second, the barrier to accessibility 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. Concerning the distribution by frequency over time, the increase has largely been attributed to privacy violations, which may point to the improvements in reporting. Costs, on the other hand, peaked in 2011, largely due to spikes in privacy violations and data breaches.

	Security incident	Data breach	Phishing/skimming	Privacy	Other	Total
Frequency	5908	54066	3629	51120	692	115415
Total losses	8699.98	20090.55	2732.95	7968.30	32.04	39523.82
Mean loss	10.05	9.02	11.06	0.68	1.69	2.62
Std dev. of loss	77.07	110.85	73.35	17.50	2.57	49.97

Notes: All numbers referring to losses are reported in USD millions. The mean and standard deviations are calculated only over the observations that have data for the loss amount available. Security incidents include hacking, DDoS-type attacks or internal IT failures. The Data breach category contains incidents that resulted in a loss of data either for the firm or personal records; The Privacy category refers to incidents when the firms have breached the privacy rights of individuals; Phishing and skimming largely refer to the sending of emails that are intended to solicit personal information from victims, as well as the copying of credit card details.

Table IV
Summary by type of cyber incident – Alternative classification

[12]The categorisation is based on the fixed effects used in the regression which is closely aligned to Romanosky (2016). See Section IV for details.

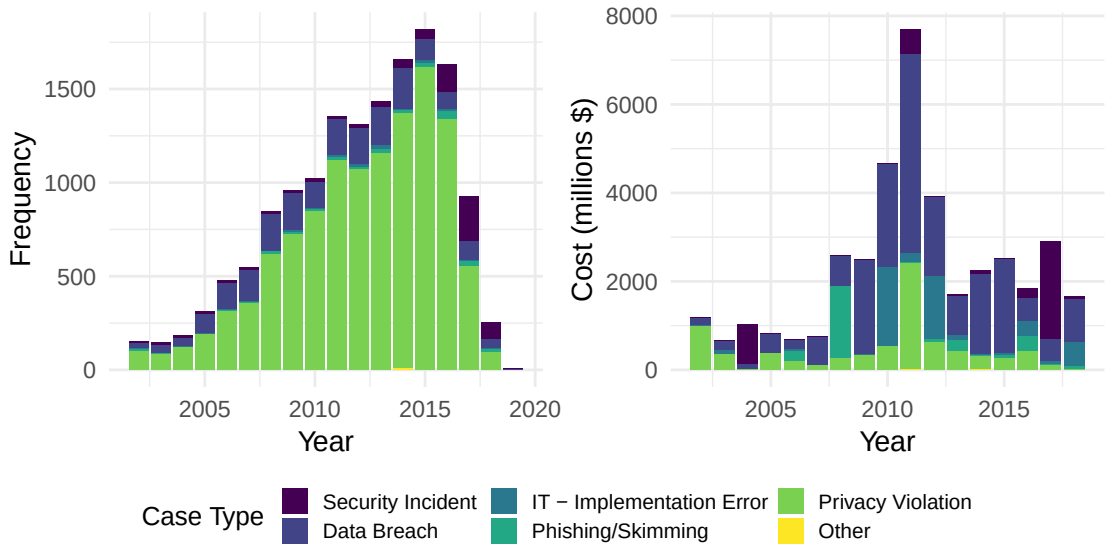


Figure 3
Distribution of frequency and costs of cyber incidents by case type over time

Figure 4 reports the frequency of cyber incidents that stem from external actors, including criminal organisations, state-backed entities or terrorist groups. Attacks deriving from these entities tend to be highly sophisticated, with the potential to cause significant monetary costs. Criminal groups' and terrorists' cyber crime skills are continuously improving. Governments, firms and individuals thus need to be increasingly proactive in guarding their cyber-assets. States have also turned to cyber attacks. A spike is observed in 2016, when hackers associated with North Korea breached Bangladesh Bank's systems and used the SWIFT network to send fraudulent money transfer orders.[13]

B. Cloud dependency and technological skills

We use data from the OECD to proxy heterogeneity in the development of IT resources per sector and also to study the effects of a different adoption of cloud services on the cost of cyber incidents. Data are only available for a subset of sectors and also for a selected number of countries. When merging the two datasets, we therefore lose information on some sectors. Moreover, a lack of consistent data across countries and time in the database limits to some extent the analysis of cross-country heterogeneity and the differential effects through time. We assume that there is heterogeneity amongst sectors in these variables, but that the levels by sectors persists across the countries included in the database and also through time in our sample data, which range from 2002-2018. While these assumptions may be restrictive, the exercise provides a first attempt at giving empirical evidence of the link between IT investment /cloud dependency and the costs of

[13] See https://reut.rs/34Qhh4a.

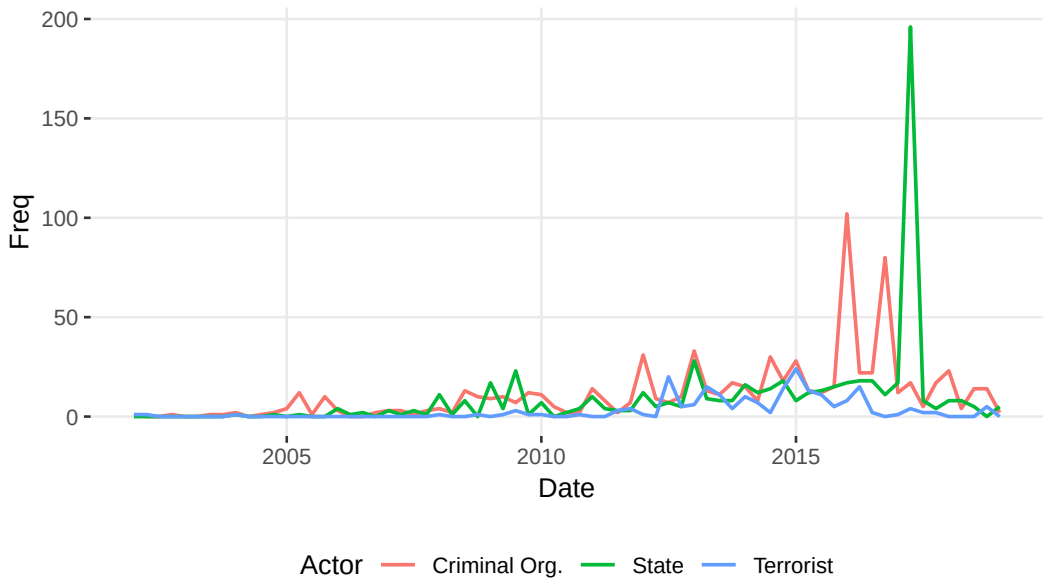


Figure 4
Frequency of cyber incidents by external actors

cyber incidents.[14] This is useful for policymakers, as it provides insights as to what factors can help reduce costs stemming from cyber risks and, in particular, which sectors are relatively more vulnerable to cyber attacks.

We construct global indicators of IT investment and cloud dependency across sectors. We take the average across regions and time to form a composite indicator for OECD countries. We take the average across regions and time to form a composite indicator for OECD countries. Formally, if $x_{i,j,t}$ is the indicator of interest in sector i, of country j, at time t, we compute the proxy variable for sector x_i as follows:

$$x_i = \frac{1}{n_i \cdot m_j} \sum_j \sum_t x_{i,t,j} \tag{1}$$

where, n_i is the number of countries that reported at time t and m_j is the number of periods for which country j reported data.

We construct three variables that explain the level of technological development of a sector:

1. *Staff training*, the percentage of firms in each sector which gives staff a specific IT training;
2. *PC users*, the percentage of staff in each sector that use a computer in their everyday work;
3. *Specialist Staff*, the percentage of firms that employ staff specialised in IT or related areas.

[14] Indeed, the US does not report this information and the database has a large skew towards the US. Thereby, much of the assumption is based on the fact that the distribution across sectors in the database acts as a good proxy for the distribution of sectors in the US.

As an indication of cloud dependency, we define the variable Cloud to be the percentage of businesses per sector that have purchased cloud services. Table V contains the computed proxies for the four variables over the sectors for which data was available, and Table VI presents summary statistics. The incidence of cloud services is highest for the ICT sector, followed by the PST sector and the FI sector. In terms of technological capabilities, the ICT and FI sectors stand out, followed by the PST sector.

	Cloud	*StaffTraining*	*PCUsers*	*SpecialistStaff*
Manufacturing	23.95	23.33	41.74	24.33
Construction	23.68	14.63	37.89	12.80
Wholesale trade	27.00	29.12	64.93	29.54
Retail trade	20.60	20.15	44.07	17.81
Transportation and storage	22.72	19.62	43.22	19.75
Accommodation, Food and beverage	18.20	12.22	31.34	10.50
Information and communication	54.41	59.11	89.80	74.61
Financial and insurance activities	32.74	56.94	84.92	65.10
Real estate activities	30.60	26.55	64.98	23.18
Professional, scientific and technical	38.19	34.99	82.43	36.64
Administrative and support service	27.84	21.09	40.94	21.55

Notes: All figures are percentages. The countries reporting data to the OECD on technological skills and cloud dependency are the following: Australia, Austria, Belgium, Brazil, Canada, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Japan, Korea, Latvia. Lithuania, Luxembourg, Mexico, Netherlands, Norway, Poland, Portugal, Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Turkey and United Kingdom. The time period over which the data was averaged is 2005-2012.

Table V
Summary of OECD data

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Std. Dev.
Cloud	18.2	23.2	27.0	29.1	31.7	54.4	10.2
StaffTraining	12.2	19.9	23.3	28.9	32.1	59.1	15.7
PCUsers	31.3	41.3	44.1	56.9	73.7	89.8	21.2
SpecialistStaff	10.5	18.8	23.2	30.5	33.1	74.6	20.8

Table VI
Summary statistics for OECD technological skills and cloud variables

C. IT spending by sector

To evaluate whether certain sectors may be underinvesting in IT security, we use a dataset of IT spending gathered from the Information Week 500 (IW500) annual survey. The database provides a reflection on the investment in IT across sectors from firms, based largely in the US (and Canada). The original data is categorised by sectors as defined by IW500. Kennedy and Stratopoulos (2017) provide a mapping of these

self-defined categories to the NAICS categorisation.[15]

The data are available from 2005-2014 for most NAICS sectors. We take an average across the years for each sector to evaluate whether each sector has, on average during this period, overspent or underspent on IT resources.

IV. Identifying the drivers of cyber costs

Policy-makers, insurance companies and firms have a great interest in developing models that can explain the cost of cyber incidents. With respect to data breaches, using Ponemon data, Jacobs (2014) uses a regression framework to link the cost of a breach to the number of records that have been compromised. Romanosky (2016) uses the Advisen data to expand on this methodology, including a broader set of variables in the regression: firm size, a dummy for whether an incident was malicious, a dummy for the case in which a firm suffered multiple losses, the number of records stolen and the effect of lawsuits. We expand on this methodology in two ways. First, we widen the scope of the regression to include a wider set of incidents and not only data breaches. We then enrich the dataset further with sector-level explanatory variables that can be used to address the impact of IT investment across sectors and the dependency on cloud technology.

As noted above, there are some data limitations. One is the limited information on losses relative to the events in the sample. The full sample consists of over 100,000 incidents. However, much of the data is sparsely populated and, in particular, not every observation has information on the costs incurred.[16] Another limitation relates to the OECD data is used to complement our dataset, which is only provided for a limited number of sectors. Taking into account these limitations, the sample is reduced to 3,228 observations. We present below some additional information on the explanatory variables used in the benchmark regression.

Firm size. Curti et al. (2019) suggest that firm size positively correlated with the size of their operational losses, using data for financial institutions. This could have three main explanations: i) losses are correlated with the level of a firms' activity; ii) institutional complexity and ii) moral hazard incentives arising from "too-big-to-fail." Our paper evaluates if the positive correlation persist for other sectors beyond financial institutions. Evidence supporting larger losses at larger firms may warrant increased attention from regulators towards such firms. In the financial sector, negative externalities associated with institutions that are perceived as not being allowed to fail due to their size, interconnectedness, complexity, lack of substitutability or global scope are well documented. However, there is little empirical evidence that characterises a firm's risk-profile with respect to cyber incidents. Biener et al. (2015) note a U-shaped relationship between firm size and average losses. We test for the existence of this relationship by including a squared term in the regression. To proxy the size of the firm incurring the loss we use the log of revenues. We test the robustness of our results by performing regressions using number of employees as an alternative proxy for firm size.

Connections among cyber events. Cyber attacks and incidents in general could exhibit features of conta-

[15] For details on how this mapping was constructed, we refer the reader to the paper of Kennedy and Stratopoulos.

[16] As indicated by Jones and Freund (2014) other loss forms, such as productivity or certain response costs are often not reported publicly. This is a limitation of the Advisen database and data losses should be considered as a lower bound.

gion (Baldwin et al., 2017; Eisenbach et al., 2020). Incidents that impact multiple-firms could lead to greater costs in absolute value. On the other hand, the costs could be more distributed across firms. We include a variable that captures how many firms were linked to one specific cyber event in order to provide evidence on this effect.

The Advisen database provides information to link related incidents. For example, if hackers infiltrated one firm and were able to manoeuvre their way into the system of another, the losses would occur at two different firms but the incidents are linked to the same attack. We use this information to construct a variable that denotes how many losses were related to a specific incident. To return to the example, assuming the two firms were the only ones to be affected, the connection variable would be 2.[17]

Malicious events. 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 $HackerType$, which is equal to one if the event had malicious intent.

Control variables. In order to control for the possible existence of trends or anomalies in specific periods of time, we use year dummies. We also include a set of dummies for the sectors reported in Table V. Lastly, we include a set of dummy variables for different types of incident:

- *Security incident*: an incident that compromises or disrupts corporate IT systems (computers or networks) or their intellectual property. For example, hacking and consequently 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.
- *IT implementation errors*: faulty hardware or software containing bugs that lead to outages or losses.
- *Phishing/Skimming*: the sending of emails purporting to be from reputable companies 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*: unauthorised collection, use or sharing of personal information. For example, 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*: cyber-related losses that were not attributed to one of the above categories.

Table VII presents summary statistics for the baseline variables. The average cost of a cyber event is

[17] 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.

large, at $12.4 million. However, the distribution of the costs is highly skewed: the median is roughly 100,000 dollars. About 40% of the events are malicious (*HackerType*, caused by a cyber attack). The variable *Connections* indicates that, on average, the number of events connected to any given event is about 6, with a median of 2. This suggests that the distribution is skewed towards a longer tail, i.e. there are a few events with many connections.

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Std. Dev.
Cost ($m)	0.0	0.0	0.1	12.4	1.5	4,000.0	106.8
HackerType	-	-	-	0.4	-	-	-
Connections	1.0	2.0	2.0	6.1	3.0	79.0	11.4
Revenues ($m)	0.0	2.6	48.0	13,950.7	3,680.7	496,785.0	40,633.2

Notes: Costs and revenues are reported in millions of US dollars. HackerType is a dummy variable, hence only the mean is reported which corresponds to the proportion of events which were of a malicious nature.

Table VII
Summary statistics for the baseline variables

Our empirical strategy is similar to Romanosky (2016). The regression is performed at a cross sectional event level. That is, each observation corresponds to a different cyber event. The empirical model is given by the following equation:

$$\log(Cost) = \beta_0 + \beta_1 FirmSize + \beta_2 Connections + \beta_3 HackerType + \beta_4 FE + u \tag{2}$$

where, *Cost* is the total cost associated with the cyber event measured in US dollars,[18] *FirmSize* is the logarithm of the reported revenue for the firm which incurred the loss,[19] *Connections* are the number of events found to be related to the given loss event (that is, the events that have the same root cause), *HackerType* is a dummy variable which takes the value of 1 if the event was caused by a cyber attack with malicious intent or 0 otherwise. FE are the fixed effects: specifically, we use the year in which the event occurred as a time fixed effect, the sector of the firm which incurred the loss as a sectoral fixed effect and the type of incident (e.g. data breaches, phishing and skimming, etc). To sum up, equation 2 explains the average costs of cyber incidents as a function of: i) a firm size effect, related to the size of the company which suffered the loss; ii) a contagion effect, proxied by the number of related events; iii) a malicious/cyber attack dummy effect; iv) fixed effects that absorb time, sector and event type variation.

Table VIII contains the results from estimating equation 2. The different columns report specifications that include different sets of fixed effects. Specifically, Model I includes only a Year fixed effect; Model II only a Sector fixed effect; Model III only an incident type fixed effect and, finally, Model IV all fixed effects.

The larger the size of a firm, the higher the estimated cyber incident costs. The elasticity coefficients

[18]The costs are nominal. For robustness, we have also run the regressions using as dependent variable the logarithm of cyber costs expressed in real USD terms. The results were unchanged.

[19]Using the number of employees as an alternative measure of firm size brings to very similar results.

imply that a one percent increase in a firm's total revenues is associated with a 0.23-0.26 percent increase in the costs of a cyber event. This result is robust to different sets of fixed effects and significantly different from a unitary elasticity. The number of related events (connections), which proxies how systemic or contagious a cyber event is, is also positively correlated with the costs of a cyber event. The semi-elasticity coefficients imply that if a cyber incident is connected with another event the overall cost increases between 1.8-2.2%. This result as well is robust to different specifications of the fixed effects. Moreover, it is in line with the intuition that the more systemic events carry the highest impact and therefore are more costly. For example, the cyber event that is most connected in our database (connected with other 79 related cyber incidents) has a cost that is almost double, other things being equal.

	Dependent Var: $\log(Cost)$			
	I	II	III	IV
FirmSize	0.231***	0.259***	0.237***	0.227***
	(0.01)	(0.02)	(0.01)	(0.01)
Connections	0.020***	0.018***	0.018***	0.022***
	(0.01)	(0.01)	(0.01)	(0.01)
HackerType	-0.066	-0.323**	-0.617**	-0.511*
	(0.12)	(0.13)	(0.29)	(0.28)
Year	Y	N	N	Y
Sector	N	Y	N	Y
Incident Type	N	N	Y	Y
R^2	0.177	0.125	0.125	0.199
Obs	3228	3228	3228	3228

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$, standard errors in parentheses.

Table VIII
Drivers of cyber costs: baseline specification

Events associated with malicious behaviour (*HackerType*) are relatively less costly (up to 60% less) than other event types. Cyber attacks get significant press coverage and it is therefore reasonable to expect them to be more costly. However, the results show that there are other cyber events, such as IT implementation and processing errors, that are more costly. Moreover, cyber security actions adopted by many firms protect them from the effects of malicious cyber incidents. There are various tools, like anti-virus software, that are able to predict and manage cyber attacks more effectively than events that occur as a result of human error inside firms. All in all, this result indicates that while cyber attacks typically receive more media attention, other things being equal, they could be less damaging on average- than other cyber events. That being said, this should not be taken as a reason to gloss over the threat that is posed by malicious cyber attacks as we will show in the subsequent quantile analysis (Section IV.B). Moreover, well co-ordinated 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.

Figure 5 shows the evolution of the estimated year fixed effect coefficients over time. The error bars

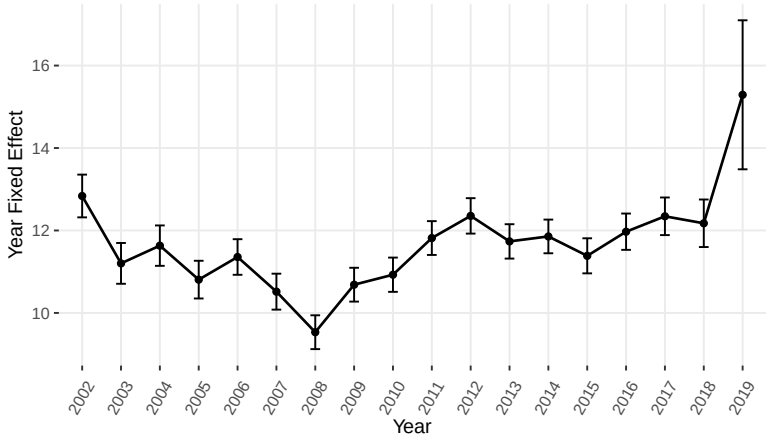


Figure 5
Year fixed effects from Table VIII

denote two standard deviations around the estimated coefficient. The average cost of an event each year (on a log scale) is increasing since 2008. This trend could be driven by a number of causes. The growth in IT complexity and interconnectivity of systems has led to new challenges regarding the use of IT solutions (Hanseth, 2007) and increasing sources of risks (Beck et al., 1994, 1992). Recent evidence suggests that it is now easier than ever to launch cyber attacks as cyber criminals have developed a lucrative market in selling do-it-yourself hacking kits and services. For example, in 2014, a Swedish citizen was charged with selling a ransomware and offering technical support to buyers, making as much as $350,000 in the process.[20]

Figure 6 shows the distribution of the estimated sectorial fixed effect coefficients. Accommodation and Food, Manufacturing and Real Estate have the highest average cost per event, whereas Administration and Support and ICT have the lowest.

The Finance and Insurance sector (circled in red in Figure 6) is below the average across sectors. It is important not to misinterpret this result as reflecting that the financial sector suffers fewer incidents; rather, risks in this sector seem to be better managed. Financial institutions indeed have always been a prime target for cyber-criminals due to what is often referred to as Sutton's rule.[21] While evidence suggests there is no shortage of attempts to penetrate the financial sector, attackers find it much harder to succeed. Due to the nature of their business, financial organisations need to develop security measures and actively manage cyber risks, also supported by regulation (we discuss more in detail this point below).[22] Financial firms

[20] https://www.bloomberg.com/news/articles/2014-05-19/u-s-malware-probe-said-to-yield-dozens-of-global-arrests.

[21] Sutton's rule states that, when diagnosing, one should first consider the obvious. The law is named after the famous bank robber Willie Sutton (1901 1980), who reputedly replied to a reporter's inquiry as to why he robbed banks by saying "because that's where the money is".

[22] Regulation of cyber security tends to vary by sector. In the US three main regulations address cyber security at the federal level: the 1996 Health Insurance Portability and Accountability Act, the 1999 Gramm-Leach-Bliley Act and the 2002 Homeland Security Act (aimed at healthcare organisations, financial institutions, and federal agencies respectively). At the international level, progress is constantly made to harmonise principles for banks. In December 2018 the Basel Committee on Banking Supervision published a report on the range of cyber-resilience practices (Basel Committee on Banking Supervision, 2018).

spend considerable sums on IT security systems and support from cyber security experts, which make them well equipped to reduce costs from cyber events, in particular cyber attacks.

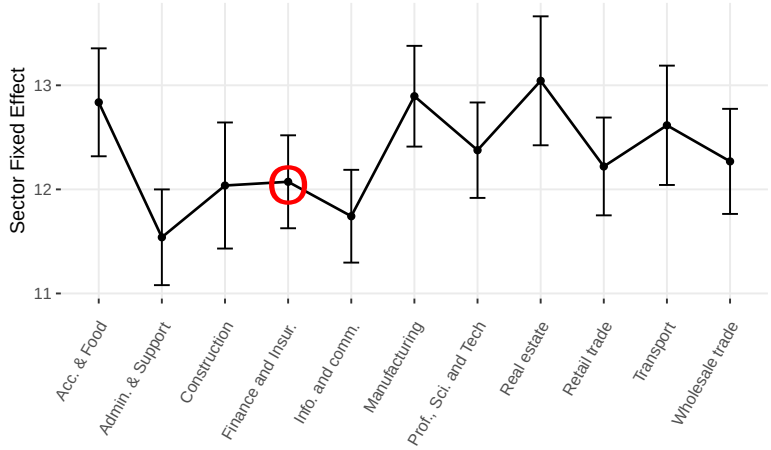


Figure 6
Sector fixed effects from Table VIII

Figure 7 shows the incident type fixed effects. Controlling for other factors, the more costly incident types are IT-related and Phishing/Skimming events. The IT category includes configuration, implementation and processing errors. The latter could be due to computer glitches, poorly designed software or failed software upgrades. As noted above, phishing is the fraudulent practice of sending emails purporting to be from reputable companies in order to convince individuals to reveal personal information, such as passwords and credit card numbers.[23] Other things being equal, these types of attacks can be particularly damaging in terms of monetary losses.

A. Testing the presence of non-linear relationship in firm size and cyber event connections

Biener et al. (2015) note that the costs of cyber incidents have a U-shaped relationships with firm size i.e. small firms and large firms incur larger losses. Moreover, and contrary to our result, they find that low connectivity events are associated with higher costs. The logic behind their finding is that as the number of firms connected to the event grows, the costs are distributed amongst more firms and tends to dilute. The different results could be due to the presence of a non-linear relationship that is not captured by the linear baseline regression (2). We check the robustness of our results in two ways.

First, we plot in Figure 8 the relationship between connections and cost (left hand panel) and between size and cost (right hand panel). In the first case, there is weak evidence of a positive relationship that is reduced in the last part of the distribution. In the second case, there are signs of a weak positive correlation but no sign of a U-shaped relationship.

[23] While phishing is conducted online, skimming is the in-person act of copying information from a card via a handheld tool the size of a lighter, or a device that fits neatly over a gas station pump or ATM. Both allow thieves to scan the account information off of a card's magnetic strip.

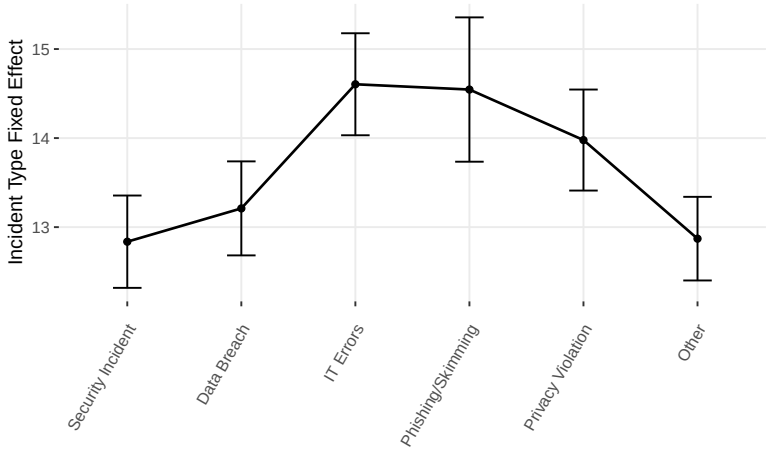


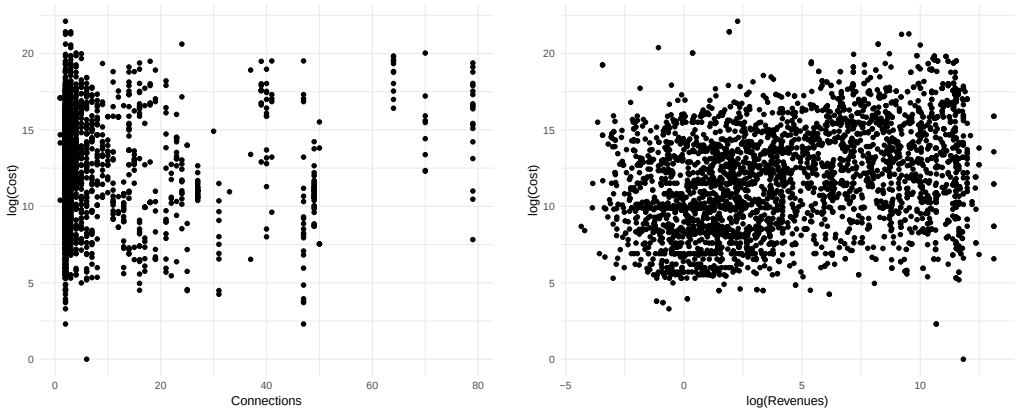
Figure 7
Incident type fixed effects from Table VIII

Second, we test for the presence of a non-linear relationship more formally, by including quadratic terms in equation 2. Table IX presents the results. There is no significant quadratic relationship for *FirmSize*. There is evidence of a quadratic relationship between the number of cyber event connections and costs. However, the maximum is reached for around 25 connections. Costs grow non-linearly with the number of entities connected to an incident and start to decline only for episodes with a very high number of connections.

	Dependent Var: $\log(Cost)$	
	I	II
FirmSize	0.259***	0.232***
	(0.04)	(0.01)
$FirmSize^2$	-0.003	
	(0.00)	
Connections		-0.050***
		(0.01)
$Connections^2$		0.001***
		(0.00)
HackerType	-0.388	-0.421
	(0.28)	(0.28)
Year	Y	Y
Sector	Y	Y
Incident Type	Y	Y
R^2	0.195	0.206
Obs	3228	3228

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$, standard errors in parentheses.

Table IX
Baseline regression with second order terms



(a) Connections versus the logarithm of costs **(b)** Revenues versus the logarithm of costs

Figure 8
Relationship of the logarithm of costs against select regressors

B. Quantile regressions

Losses stemming from cyber incidents could be characterised by a "heavy-tailed" distribution (Cohen et al., 2019). Therefore, it is reasonable to assume that the impact of the variables used may vary across different cost quantiles. In particular, the tails could be fundamentally different from the rest of the distribution. To check for the presence of different effects along the distribution of costs, we re-run our baseline regression at varying cost quantiles. The quantiles used are the 25%, 50%, 75%, and 90%. We present the results of this test in Table X.

The baseline regression is reported in the first column. *FirmSize* and *Connections* display a similar relationship when moving from lower to higher quantiles. The magnitude of the coefficients becomes larger towards the upper quantiles, although in both cases there is a decline between the 75th and 90th percentiles. Interestingly, the *HackerType* variable, which is originally observed to have a negative coefficient, changes sign at the 90th percentile. While the majority of malicious events are other things being equal – less damaging than those that are accidental, sophisticated hacks can actually exacerbate costs at the tail end of the distribution.[24]

C. Standard error clustering

Note that the residuals of equation 2 may be correlated across numerous dimensions and one useful robustness check is to use clustered standard errors. However, clustering usually reduces the precision of the standard errors and could automatically reduce the significance of the results (Cameron and Miller (2015)). The question is: which dimension should be chosen to cluster, if any? Abadie et al. (2017) argue that clustering is either a sampling design or an experimental design issue. From the sampling perspective,

[24]We repeat a similar quantile regression for each of the regressions run in the subsequent sections. However, there was no significant material change to the results presented.

	Dependent Var: log(*Cost*)				
	Mean	25%	50%	75%	90%
FirmSize	0.227***	0.167***	0.263***	0.289***	0.249***
	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)
Connections	0.022***	0.009	0.019**	0.027***	0.022***
	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)
HackerType	-0.511*	-1.007***	-0.512	-0.193	0.521**
	(0.28)	(0.20)	(0.34)	(0.21)	(0.22)
Year	Y	Y	Y	Y	Y
Sector	Y	Y	Y	Y	Y
Incident Type	Y	Y	Y	Y	Y
Obs	3228	3228	3228	3228	3228

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$, standard errors in parentheses.

Table X
Quantile regressions

we should consider whether there are clusters in the population of interest that are not represented in the sample. On the other hand, clustering is an experimental design issue if the assignment is correlated within the clusters.

Table XI takes a pragmatic approach and shows the results of the complete model when clustering along different dimensions: by sector, year, type of event (CaseType) or by connected events (Events). Results remain qualitatively similar.

	Dependent Var: log(*Cost*)			
	I	II	III	IV
FirmSize	0.227***	0.227***	0.227***	0.227***
	(0.04)	(0.03)	(0.04)	(0.02)
Connections	0.022***	0.022	0.022***	0.022
	(0.009)	(0.014)	(0.009)	(0.015)
HackerType	-0.511	-0.511	-0.511***	-0.511*
	(0.38)	(0.44)	(0.09)	(0.27)
Year	Y	Y	Y	Y
Sector	Y	Y	Y	Y
Incident Type	Y	Y	Y	Y
R^2	0.199	0.199	0.199	0.199
Obs	3228	3228	3228	3228
σ_c	Sector	Year	CaseType	Events

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$, clustered standard errors in parentheses.

Table XI
Alternative standard error clustering for baseline regression

V. Does the development of technological skills mitigate cyber costs?

Sectors that face shortages in IT skills and investment in staff training may leave themselves exposed to cyber risks. In this section, we use sector aggregate data from the OECD as a measure of technology skills across firms. While the frequency of cyber events may depend on exogenous factors, their impact (or severity) can be mitigated by maintaining sound IT practices. Based on this we postulate the following:

Hypothesis 1. *A higher development of technological skills can mitigate losses from cyber incidents.*

This hypothesis reflects our prior belief that the development of technological skills should reduce the severity of costs associated with cyber events. IT plays an integral role in day-to-day business across all sectors, hence those with greater human capital with respect to technological skills should manage the risks more effectively.

	Dependent Var: $\log(Cost)$			
	I	II	III	IV
$FirmSize$	0.335*** (0.04)	0.373*** (0.05)	0.307*** (0.03)	0.232*** (0.01)
$Connections$	0.022*** (0.01)	0.022*** (0.01)	0.022*** (0.01)	0.022*** (0.01)
$HackerType$	-0.519* (0.28)	-0.519* (0.28)	-0.520* (0.28)	-0.519* (0.28)
$FirmSize \times StaffTraining$	-0.002*** (0.00)			
$FirmSize \times PCUsers$		-0.002*** (0.00)		
$FirmSize \times SpecialistStaff$			-0.002*** (0.00)	
$FirmSize \times \Gamma$				-0.001*** (0.00)
Year	Y	Y	Y	Y
Sector	Y	Y	Y	Y
Incident Type	Y	Y	Y	Y
R^2	0.202	0.202	0.201	0.202
Obs	3228	3228	3228	3228

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$, standard errors in parentheses.

Table XII
The role of technological skills in mitigating cyber costs

Since the OECD skills variables are aggregated at the sectoral level, they are subsumed in sectoral fixed effects. For this reason, to observe the effect of technological skills on the severity of cyber losses we include them as interaction terms with the baseline regressors. In particular, we modify equation 2 by including one extra interaction term at a time. For example, if we consider the interaction term between $FirmSize$ and $StaffTraining$ we have the following:

$$\begin{aligned} \log(Cost) = \beta_0 + \beta_1 FirmSize + \beta_2 Connections + \\ \beta_3 HackerType + \beta_4 FirmSize \times StaffTraining + FE + u \end{aligned} \tag{3}$$

Table XII contains the estimated coefficients for the regression model in equation 3. Each of the spec-

ifications includes the interaction of *FirmSize* with one of the variables expressing technological skills. Model I includes the interaction of *FirmSize* with $StaffTraining$; Model II the interaction of *FirmSize* with $PCUsers$; Model III the interaction of *FirmSize* with $SpecialistStaff$; Model IV the interaction of *FirmSize* with a linear combination of the technology dependence variables, equivalent to the principal component corresponding to the largest eigenvalue, Γ.

From Table XII, consider first the baseline regressors *FirmSize*, *Connections* and $HackerType$. Their signs are the same as those in the baseline regression. The magnitudes and significance of the *Connections* and $HackerType$ variables remain consistent, whereas *FirmSize* has a higher magnitude making room for the significant negative coefficients estimated for the interaction term including the same variable. The coefficients of the interaction terms reveal that all of the technological skills variables have a statistically significant mitigating effect on costs.

Consider two events, occurring at firms of equivalent size, but of differing sectors. Our results suggest that a firm from the sector with better development of technological skills manages costs more effectively: there is a reduction in the severity of costs between 0.1-0.2% for every unit increase in the skills variables. To see what difference that can make, we compare the effect on costs for firms of the same size from the sector with highest development of technological skills to that of the lowest. For $StaffTraining$, the sectors with the highest and lowest values are Information and communication (59%) and Accommodation, food and beverage (12%), respectively. The expected costs of cyber losses are 9.5% [25] lower in the former than in the latter. For $PCUsers$ and $SpecialistStaff$, under the same assumptions, the difference in expected costs between the least and most technologically developed sectors are 11.7% (Information and communication versus Accommodation, food and beverage) and 12.4% (Information and communication versus construction) respectively.

Table XIII contains the estimated coefficients for the regression model in equation 3 with the inclusion of the number of connected events with different sectorial measures expressing technological dependence. The structure is similar to that of Table XII. The results for the baseline regressors (*FirmSize*, *Connections* and $HackerType$) are qualitatively similar, but *Connections* in this case is larger in magnitude, accommodating the significant negative coefficient in the interaction terms. The interaction terms in the different columns indicate a statistically significant mitigating (negative) effect on the expected costs of a cyber event. The magnitude of the interaction terms is 0.001 across all specifications. This reflects a reduction in the expected severity of costs between 0.1% for every unit increase in the skills variables, given a constant number of connections. A similar comparison of most versus least technologically advanced sectors for our three variables yields to expected losses that are 4.7%, 5.9% and 6.2% lower for the sectors with higher technological skills (for a given level of connections).

[25] This is computed as: $(59-12)\times 0.002 \approx 0.095$.

	Dependent Var: $\log(Cost)$			
	I	II	III	IV
$FirmSize$	0.227*** (0.01)	0.226*** (0.01)	0.227*** (0.01)	0.227*** (0.01)
$Connections$	0.069*** (0.01)	0.086*** (0.02)	0.066*** (0.01)	0.027*** (0.01)
$HackerType$	-0.556** (0.28)	-0.554** (0.28)	-0.564** (0.28)	-0.560** (0.28)
$Connections \times StaffTraining$	-0.001*** (0.00)			
$Connections \times PCUsers$		-0.001*** (0.00)		
$Connections \times SpecialistStaff$			-0.001*** (0.00)	
$Connections \times \Gamma$				-0.001*** (0.00)
Year	Y	Y	Y	Y
Sector	Y	Y	Y	Y
Incident Type	Y	Y	Y	Y
R^2	0.203	0.202	0.203	0.203
Obs	3228	3228	3228	3228

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$, standard errors in parentheses.

Table XIII
Technological skills and connected events

Table XIV presents the results of interacting the Hacker type variable with the technological skills variables. The structure is similar to XII and XIII. The coefficients of the baseline regressors *FirmSize* and *Connections* are similar to those obtained in Tables XII and XIII. By contrast, when the interaction terms are included, $HackerType$ becomes statistically not significant, at all levels. Moreover, the interaction terms are not statistically significant.

	Dependent Var: $\log(Cost)$			
	I	II	III	IV
$FirmSize$	0.228*** (0.01)	0.227*** (0.01)	0.228*** (0.01)	0.227*** (0.01)
$Connections$	0.022*** (0.01)	0.022*** (0.01)	0.021*** (0.01)	0.022*** (0.01)
$HackerType$	-0.067 (0.40)	-0.372 (0.49)	-0.203 (0.37)	-0.513* (0.28)
$HackerType \times StaffTraining$	-0.011 (0.01)			
$HackerType \times PCUsers$		-0.002 (0.01)		
$HackerType \times SpecialistStaff$			-0.007 (0.01)	
$HackerType \times \Gamma$				-0.004 (0.00)
Year	Y	Y	Y	Y
Sector	Y	Y	Y	Y
Incident Type	Y	Y	Y	Y
R^2	0.2	0.199	0.2	0.2
Obs	3228	3228	3228	3228

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$, standard errors in parentheses.

Table XIV
Technological skills and incidents with malicious intent

VI. Does cloud technology alter cyber risks?

Cloud technology enables firms to rent computing power and storage from service providers, which gives them flexibility in their storage costs. In this way, some fixed costs become marginal costs of production. Moreover, cloud computing poses positive externalities such as the reduction of energy consumption and carbon emissions (Etro, 2015).

However, cloud computing also faces some risk, and in particular involves firms inherently placing a lot of trust in vendors of cloud technology (Urquhart, 2009). The presence of a market failure through information asymmetry between buyer and vendor is rather well recognised. Often users of cloud services may not know the exact location of their data or the other sources of the data collectively stored with theirs. Cloud computing can be a double-edged sword. On one hand, firms are implicitly purchasing the cyber security expertise of vendors like Microsoft, Google and Amazon. This can be especially beneficial for firms with low IT budgets and a general lack of IT skills, which as was highlighted in the previous section can be detrimental when it comes to the expected costs of cyber losses. That said, firms are still responsible for maintaining sound IT practices, as individual servers can still be at risk of compromise through poor security. Cloud computing can also be a target for cyber-criminals. Clouds of multiple entities (providers) could be formed; such configuration is only as strong as its weakest link. If this link can be identified and exploited, this renders the entire cloud vulnerable. Consequently, an attacker can do considerable damage across multiple entities by exploiting a weakness in one firm (Kaufman, 2009).[26]

Cloud computing could pose a concern in terms of systemic risk. Providers of cloud services, undoubtedly have some of the best cyber security experts and ultimately provide highly secure services, but tail-risks could lead to substantial losses and potentially bring the economy to a halt (Danielsson and Macrae (2019)). Moreover, the market for cloud services is highly concentrated, and there are warnings about increased homogeneity and the greater risk of single points of failure.[27] Through shared software, hardware and vendors, incidents could, in principle, spread more quickly, leading to higher overall costs. The impact of the use of cloud services in the case of cyber attacks can thus go both ways, and clearly depends on the benefit-risk analysis. Based on the discussion above, we make the following hypothesis.

Hypothesis 2. *A higher dependency on cloud technologies can alter losses from cyber events. However, the net benefit depends on the connectivity of the cyber incidents and the size of the shock.*

We test our hypothesis by using data collected from the OECD on purchases of cloud technology by sector. The rationale is that firms with higher investment in cloud technologies are consequently more exposed to the benefits and risks that come with it.

We look at two regressions including the cloud variable. In the first one, the cloud dependency variable is introduced as a simple regressor and – due to the multicollinearity issue previously discussed – the sector fixed effect is omitted. We then include it as an interaction with the Connections variable and keep the fixed

[26] For a full discussion of the benefits and risks of cloud computing see Catteddu (2009).

[27] For a discussion of cloud service providers as critical infrastructures and potential policy responses, see Carr et al. (2019).

effects. Cloud computing increases the potential for connectivity between firms and thus it is interesting to evaluate the impact of events with a high connectivity and high usage of cloud services. If we consider the second specification with the interaction between Cloud and Connections we have the following equation:

$$\begin{aligned} \log(Cost) = \beta_0 + \beta_1 FirmSize + \beta_2 Connections + \\ \beta_3 HackerType + \beta_4 Cloud \times Connections + \beta_5 FE + u \end{aligned} \tag{4}$$

We present the results of the regressions in Table XV. The higher the investment in cloud services, the lower the cost of cyber events: the coefficient on cloud dependence in the second column is negative and significant. A similar negative coefficient occurs for the interaction term between *Connections* and *Cloud* in the third columns. In other words, increasing cloud dependence reduces the impact of connected cyber attacks. This result, however, needs a qualification. Whether, and under what conditions, clouds absorb or spread losses depends on the size of the shock and the connectivity of sectors. As long as the negative shocks are sufficiently small, clouds could act as a shock absorber. Ideally, cyber attacks would be fully neutralised by the cloud's cyber security technologies and cloud users would not be affected. However, once the attack is able to overcome the clouds' security barriers, the same features that make a cloud more resilient may become sources of instability.

	Dependent Var: $\log(Cost)$		
	I	II	III
FirmSize	0.227*** (0.01)	0.223*** (0.01)	0.228*** (0.01)
Connections	0.022*** (0.01)	0.022*** (0.01)	0.076*** (0.02)
HackerType	-0.511* (0.28)	-0.527* (0.28)	-0.572** (0.28)
Cloud		-0.015*** (0.00)	
Connections × *Cloud*			-0.002*** (0.00)
Year	Y	Y	Y
Sector	Y	N	Y
Incident Type	Y	Y	Y
R^2	0.199	0.191	0.203
Obs	3228	3228	3228

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$, standard errors in parentheses.

Table XV
Dependence on cloud technology and cyber costs

VII. Does the price of Bitcoin motivate hackers?

Cryptoassets have grown in interest and trading volumes since the advent of Bitcoin (Nakamoto, 2008). Proponents of cryptoassets suggest they have the potential to improve financial inclusion, user experience, and reduce transaction costs. However, they could also amplify cyber risks. In particular, intermediaries

like market exchanges where cryptoassets are traded have been a target for cyber-criminals. Exchanges handle billions of dollars (or equivalents) in investor assets as well as providing various other services. While financial regulation establishes a separation between trading platforms and custodians to mitigate possible conflicts of interest, there are no similar requirements for intermediaries operating in the crypto space (Massad, 2019).

Financial regulators also mandate financial institutions to maintain high security practices. Cyber scenarios are now being factored into stress testing.[28] At the moment, these crypto firms fall out of the remit of such regulations as well.[29] This lack of regulation may also leave these intermediaries vulnerable to cyber risks.

The incentive to attack firms that hold cryptoassets is a function of the price of cryptoassets. As this price increases, the potential pay-off and hence the incentive to launch an attack on firms that store large quantities of cryptocurrencies becomes greater. To investigate this relationship, we operate in different steps. First, we filter the database for events that are related to cryptocurrencies and create a dummy variable for these events. By including this dummy in the cost regression, we are able to verify if these events have an additional impact on cyber costs. In a second step, we look at the relationship between the frequency of crypto-related attacks and the price of Bitcoin.

The frequency of attacks correlates with the price of Bitcoin. The upper panel of Figure 9 shows the frequency of attacks on crypto exchanges over the past four years. The lower panel reports the price of Bitcoin (averaged across the top four exchanges). The increase in price and frequency of attacks up until late 2017 indicates that monetary incentives could be motivating hackers. On the flipside, the price decline following this peak in attacks may be the consequence of wavering confidence of investors due to evidence of exploits. However, this decline may also be the result of improved security at exchanges. It should be noted that the frequency of attacks is a lower bound of the real number, as the number of unsuccessful attacks is not reported.

Based on the above discussion, we formulate the following three hypotheses:

Hypothesis 3. *Events related to cryptocurrencies are, on average, more costly.*

Hypothesis 4. *Rising valuations in the price of cryptocurrencies can lead to an increase the frequency of crypto-related cyber events involving.*

Hypothesis 5. *Cyber incidents involving crypto exchanges can have a negative impact on cryptocurrency prices.*

To test these hypotheses, we employ price data collected from crypto exchanges, and we build three

[28]The Bank of England has piloted some cyber-attack scenarios as part of its stress testing of banks. See https://www.reuters.com/article/us-boe-cyber-tests/bank-of-england-to-test-banks-resilience-to-cyber-attacks-idUSKCN1QM1H7.

[29]There are some specific regulations, such as the BitLicense in New York that is granted to firms that are operating in the crypto space. These regulations are not as restrictive as the benchmark financial regulations, but do require firms to prove they have a sound risk management framework. However, these regulations are applied at state level, hence a firm wishing to circumvent them need only shift its operations outside of the state of New York.

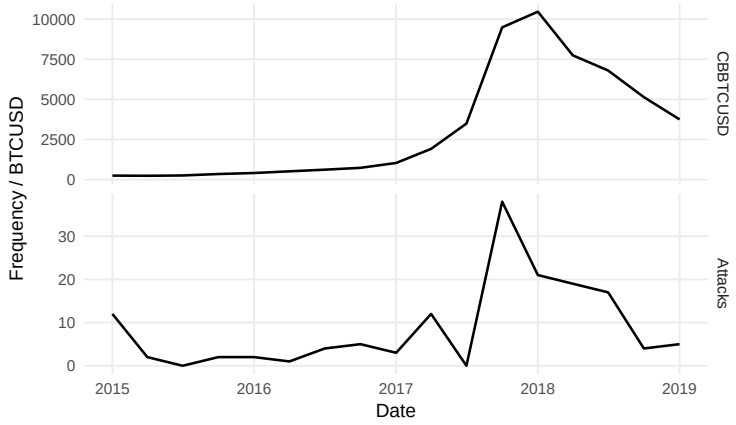


Figure 9
Frequency of attacks on crypto exchanges and the price of bitcoin

different regression models. In the first, we modify equation 2 by including a dummy variable that equals one for crypto related events. The regression equation becomes

$$\begin{aligned} \log(Cost) = \beta_0 + \beta_1 FirmSize + \beta_2 Connections + \\ \beta_3 HackerType + \beta_4 Cryptorelated + u \end{aligned} \tag{5}$$

Results are presented in Table XVI, where for convenience we also report the baseline model in the first column. The coefficient of the *CryptoRelated* variable in the second column indicates approximately a 128% higher expected cyber cost when the incident is crypto-related, i.e. more than twice as high as a non-crypto related event. This effect is statistically significant.

	Dependent Var: $\log(Cost)$					
	Baseline	All	Security Incidents	Data Breaches	Phishing/Skimming	Privacy
FirmSize	0.227***	0.230***	0.195***	0.170***	0.126**	0.303***
	(0.01)	(0.01)	(0.06)	(0.02)	(0.06)	(0.02)
Connections	0.022***	0.021***	0.016	0.027***	-0.003	-0.016
	(0.01)	(0.01)	(0.02)	(0.01)	(0.03)	(0.01)
HackerType	-0.511*	-0.527*		-0.468		
	(0.28)	(0.28)		(0.30)		
CryptoRelated		1.285**	-1.976*	2.614***	3.095*	5.060*
		(0.55)	(1.08)	(0.82)	(1.58)	(2.97)
Year	Y	Y	Y	Y	Y	Y
Sector	Y	Y	Y	Y	Y	Y
Incident Type	Y	Y	N	N	N	N
R^2	0.199	0.201	0.418	0.136	0.556	0.304
Obs	3228	3228	250	1334	169	1452

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$, standard errors in parentheses.

Table XVI
Impact on costs of crypto-related cyber incidents: Romanosky definitions

Is the higher cost of attacks on crypto exchanges directly linked to the type of attack these exchanges suffer? To answer this question, we perform regressions on subsets of the data that have been partitioned by the type of incident, following Romanosky's definitions (see Table IV). Interestingly, incidents involving cryptocurrencies are particularly costly for categories associated to events likely to yield monetary gains. For example, hackers attacking crypto exchanges are more likely to engage in phishing/skimming schemes to try and obtain user passwords and steal funds rather than launch a DDoS attack that result in damages to the exchange and its users but is less likely to derive monetary gains. Similar results are obtained dividing the cyber incidents on the base of the Bouveret (2018) definitions (see Table III).

To evaluate the second hypothesis, we look at the relationship between the likelihood of a cyber event and the price of Bitcoin. In particular, we investigate whether price increases are associated with a higher probability of future attacks on crypto-exchanges. To this end we use a Probit regression with, respectively, the weekly and fortnightly lag of Bitcoin price:

$$CryptoEvent_t = \Phi(\beta_0 + \beta_1 \log(Price_{t-k})) \tag{6}$$

where, $CryptoEvent_t$ is a binary variable indicating whether a crypto-related event occurred, $\Phi(\cdot)$ is the cumulative distribution function of the standard normal, and $Price_{t-k}$ is the daily price of Bitcoin at lag k averaged from the leading four market exchanges. Table XVII shows a positive correlation between the 7 and 14-day lags of the price of bitcoin and the probability that a cyber event occurs. The effect is significant, indicating that the probability of an event occurring following a price rise is more likely. The coefficients imply an increase of the z-score by 0.13-0.12, for 7 and 14 day lags respectively. To better interpret this effect we can evaluate the derivative:

$$\frac{\partial \Phi(X_i\beta)}{\partial \log(Price_{t-k})} = \Phi(\beta_0 + \beta_1 \log(Price_{t-k}))\beta_1 \tag{7}$$

Taking the derivative and evaluating at the mean of $\log(Price_{t-k})$:

$$\frac{\partial \Phi(X_i\beta)}{\partial \bar{p}_{t-7}} = 0.014, \quad \frac{\partial \Phi(X_i\beta)}{\partial \bar{p}_{t-14}} = 0.013 \tag{8}$$

where, $\bar{p}_{t-k}$ denotes the mean of $\log(Price_{t-k})$. These estimates are the marginal effects, i.e. the effect on the dependent variable (probability of a crypto-related cyber-event occurring) in response to a change in the independent variable (log of bitcoin price) evaluated at its mean value. We interpret the result as follows: a 1 unit increase at the logarithm of the mean price of Bitcoin leads to a 1.4% (1.3%) increase in the likelihood of there being a crypto-related cyber event one (two) week(s) later.

Finally, to evaluate our third hypothesis and observe the impact of a cyber-event on the price of Bitcoin we look at the inverse relationship. We regress a lag of the binary variable on the first difference (returns) of the bitcoin price

$$\Delta Price_t = \beta_0 + \beta_1 ExchangeEvent_{t-k} \tag{9}$$

	Dependent Var: *CryptoEvent*	
	I	II
Intercept	-2.09***	-2.1***
	(0.164)	(0.16)
$Price_{t-7}$	0.13***	
	(0.023)	
$Price_{t-14}$		0.12***
		(0.023)
AIC	1536	1536
Obs	2237	2230

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$, standard errors in parentheses.

Table XVII
Hackers respond to the economic incentives of a rising Bitcoin price

The results are presented in Table XVIII. The coefficients are not significant: there is no effect of cyber events occurring at crypto exchanges on the subsequent price of bitcoin. This test has to be treated with caution. To observe this relationship may indeed require price data at a higher frequency, and also the precise time at which the information regarding the cyber event became public, which is not available in our dataset.

	Dependent Var: $\Delta Price$	
	I	II
Intercept	0.56	1.3
	(4.96)	(4.96)
$CryptoEvent_{t-1}$	9.38	
	(14.85)	
$CryptoEvent_{t-7}$		3.1
		(14.8)
R^2	0.000	0.000
Obs	2243	2237

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$, standard errors in parentheses.

Table XVIII
Crypto-related cyber attacks do not affect the price of Bitcoin

VIII. Is the financial sector more resilient?

The financial sector handles trillions of dollars in transactions every day, assuring the correct functioning of the payment system in the economy. Due to their pivotal role, banks and other financial institutions are inherently interconnected and heavily reliant on information technology. External events like a cyber attack or an IT failure can therefore lead to severe consequences for the financial system: if a single node in the

network of financial firms fails, a domino effect and cascades of failures could take place. The shocks could propagate in the network through various mechanisms.

In principle, a more interconnected financial system could help to absorb shocks. Especially in the case of small shocks, a strongly connected system can act as a loss absorber, insulating balance sheet exposures. However, as discussed for the case of cloud dependency, if the size of the shock is large, a dense connectivity could amplify its effects (Acemoglu et al., 2015; Haldane and May, 2011).

In terms of cyber-physical connections, financial market infrastructures tend to exhibit a "hub-and-spoke" architecture, where many transactions between banks are facilitated by intermediaries, e.g. payment systems, clearing houses. This structure is not particularly robust to shocks (Albert et al., 2000). These intermediaries are essentially hubs in the network and affecting them (potentially through a cyber attack) could lead to cascading failures over numerous dimensions e.g. financial and cyber-physical.

The financial system is heavily regulated and a substantial share of its revenues are invested in protecting itself from the fragilities discussed above. Has this large investment in cyber security been effective so far? In Section IV, we tried to answer this question by analysing the fixed effects in the cost equation 2. In particular, we showed that the financial sector, other things being equal, has an expected cyber cost that is lower than other (less protected) sectors.

In this section, we investigate this relationship more explicitly. Specifically, we introduce a dummy variable that indicates whether the event occurred in the financial sector and then look at the interaction between this term and other variables across several regressions. To isolate the effect from the financial sector, we use the more granular NAICS definitions to construct a dummy that is specific to financial sector firms (excluding insurance). The dummy is introduced as an interactive term with the variables in our regressions.

The results from interacting the term with the technological skills variables are contained in Table XIX. While the three baseline regressors confirm the previous general findings, the (negative) impact of technological skills variables for the financial sector (see interaction terms) is stronger than in other sectors. All measures of investment in technological skills decrease the expected cost in the financial sector by more (reductions range between 0.3% and 1.3%, depending on the measure used). All in all, these results indicate that investment into information technology in the financial sector is money well spent.

We next assess whether the main drivers of cyber costs are different for the financial sector (Table XX). Model I includes an interaction term between the financial sector and each of the baseline regressors. Model II also includes the *CryptoRelated* dummy and its interaction with the financial sector dummy. Lastly, Model III includes our *Cloud* variable and the corresponding interaction with the financial sector dummy.

	Dependent Var: $\log(Cost)$			
	I	II	III	IV
$FirmSize$	0.227*** (0.01)	0.227*** (0.01)	0.227*** (0.01)	0.222*** (0.01)
$Connections$	0.022*** (0.01)	0.022*** (0.01)	0.022*** (0.01)	0.021*** (0.01)
$HackerType$	-0.511* (0.28)	-0.511* (0.28)	-0.511* (0.28)	-0.538* (0.28)
$Financial \times StaffTraining$	-0.013** (0.01)			
$Financial \times PCUsers$		-0.009** (0.00)		
$Financial \times SpecialistStaff$			-0.012** (0.00)	
$Financial \times \Gamma$				-0.003 (0.00)
Year	Y	Y	Y	Y
Sector	Y	Y	Y	Y
Incident Type	Y	Y	Y	Y
R^2	0.199	0.199	0.199	0.188
Obs	3228	3228	3228	3228

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$, standard errors in parentheses.

Table XIX
Technological skills are particularly helpful in reducing costs for the financial sector

The interaction of the financial sector dummy with *FirmSize* is significant and negative. While the elasticity between cyber costs and total revenues is 0.27 for the other sectors, it drops to 0.15 for the financial sector. Interestingly, the $HackerType$ dummy is not significant alone but the interaction shows that malicious events that take place in the financial sector are less costly. These results support the conjecture that the financial sector due to regulation and higher investment in cyber security is less exposed to cyber costs and well equipped to ride out cyber attacks. In Model II, the interaction of the financial sector dummy with $CryptoRelated$ events is insignificant, indicating that the financial sector (as well as the other economic sectors) do not suffer larger costs in the case of crypto-related incidents. Interestingly, the interaction term between the financial sector dummy and the *Cloud* indicator is positive, halving the coefficient on the *Cloud* alone. To interpret this, think of two events that occurred in two different sectors that have an equivalent level of cloud use, one the sectors being the financial sector. Our results imply that the expected costs of the event in the financial sector would be higher than in the other sector.

IX. Are we investing enough in IT?

In this section we analyse the distribution of IT security spending across sectors with respect to a benchmark model. To do this we use a database constructed by Kennedy and Stratopoulos (2017) based on the IW500 survey that aggregates IT spending across sectors of firms predominantly located in the US. The benchmark model is constructed using the Gordon and Loeb (GL) model. In particular, the GL model, detailed in Gordon and Loeb (2002), is typically used to assess optimal security investment at the firm level. Here we will generalise the model to perform analysis at the sector level.

	Dependent Var: $\log(Cost)$		
	I	II	III
FirmSize	0.273*** (0.02)	0.277*** (0.02)	0.273*** (0.02)
Financial × *FirmSize*	-0.125*** (0.03)	-0.127*** (0.03)	-0.125*** (0.03)
Connections	0.019*** (0.01)	0.018*** (0.01)	0.019*** (0.01)
Financial × *Connections*	0.006 (0.01)	0.008 (0.01)	0.006 (0.01)
HackerType	-0.059 (0.30)	-0.088 (0.30)	-0.059 (0.30)
Financial × *HackerType*	-0.998*** (0.27)	-0.973*** (0.27)	-0.998*** (0.27)
CryptoRelated		1.371** (0.62)	
Financial × *CryptoRelated*		-0.617 (1.28)	
Cloud			-0.052 (0.05)
Financial × *Cloud*			0.039** (0.02)
Year	Y	Y	Y
Sector	Y	Y	N
Incident Type	Y	Y	Y
R^2	0.209	0.21	0.209
Obs	3228	3228	3228

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$, standard errors in parentheses.

Table XX

Drivers of cyber cost: does the financial sector differ?

Each sector faces the problem of how much should be invested in IT security, given a breach probability function, $S(c,v)$. The breach probability function can be considered the probability of a cyber incident, occurring in a given period. The function is dependent on two factors, the security level of the sector (i.e. how much it invests in security), and the inherent vulnerability of that sector.

$$S(c,v) = \frac{v}{(\alpha c + 1)^{\beta}} \tag{10}$$

where c denotes the level of IT investment per sector; $v \in [0,1]$ is the exogenous vulnerability of suffering an incident in the given time period; and α and β are parameters that govern the efficiency of the investment, we use relatively standard parameterisations of $\alpha = 0.4$ and $\beta = 1$.

We consider the time period to be one year, hence v translates as the probability of a sector suffering a cyber incident in a year. In our database, every sector suffered some form of cyber incident in each year of available data, thus we set $v = 1$.

We can now define the expected benefit of security investment:

$$\Omega = \lambda(v - S(c,v)) \tag{11}$$

where, λ is the expected losses for a given year. The expected benefit can be thus interpreted as the expected losses without investment in security minus the expected loss given there was investment in security. The expected net benefit is simply the expected benefit minus the cost of the investment, defined as:

$$\Omega_B = \lambda(v - S(c,v)) - c \tag{12}$$

To compute the optimal IT security investment for one year, lastly we need to provide an estimate for λ across each of our sectors. To do this we take a monte carlo approach and compute the annual loss distribution in each sector.[30] We assume that the frequency of incidents per year in each sector is distributed according to a Poisson function, while the losses are Pareto-distributed (heavy-tailed). We fit these distributions to the data in each sector. With the fitted estimates, we perform a convolution of the distributions to get an estimate for the annual loss distributions in each sector. The expected loss is simply the mean of the annual loss distribution.

We then need to find the optimal level of investment, c^*. Which is simply to maximise the expected net benefit of investment with respect to c, as follows:

$$\max_c \Omega_{NB}(c) \tag{13}$$

the solution to which is,

$$c^* = \frac{1}{\alpha}(\lambda\alpha\beta)^{\frac{1}{\beta+1}} - 1 \tag{14}$$

In Table XXI we report the optimal level of investment by sector. In particular, the table reports the actual average IT expenses in each sector according to the IW500 data, our estimates of the optimal spending c^*, and the difference between the two. The last column gives an indication of what are the sectors that are overspending (positive difference) and which sectors are underspending (negative differences). To control for size, the figures are reported as a percentage of firms' revenue in each sector.

The bottom of Table XXI reports a measure for the total and indicates an overall underspending. The financial sector overspends relative to the size of its revenues. However, the model does not take into account the centrality of each sector with respect to the others i.e. how systemically important is each sector. Therefore the model only allows us to analyse what the optimal level of private investment rather than what is the optimal level for the society. We would expect that sectors that are critically important for the functioning of society should invest more to account for the implications of losses beyond their own. The high level of IT security investment contributes to explain our result that the financial sector shows a good resilience against cyber incidents relative to other sectors. Other spending surplus sectors include agriculture, manufacturing and utilities. Two sectors stand out in terms of underspending: Arts, entertainment and recreation, and Professional scientific and technical. The entertainment industry has indeed suffered some high-profile cyber-attacks and it is often the target of piracy criminals. Another sector to highlight, which is often cited as a target for hackers, is the healthcare sector. The sensitive information that is held by healthcare companies is highly sought after by criminals. While not appearing to be in

[30] This approach is typically used in operational risk literature and referred to as the loss distribution approach, for details on the computation we refer the reader to Cruz et al. (2015).

such a large deficit relative to other sectors, this sector requires increasing attention because of the potential negative effects on individuals' privacy.

	Actual	c^*	Difference
Accommodation, food and beverage	0.02	0.05	-0.03
Administrative and support service	0.05	0.16	-0.12
Agriculture, Forestry, Fishing and Hunting	0.03	0.00	0.03
Arts, Entertainment, and Recreation	0.03	0.20	-0.17
Construction	0.02	0.09	-0.07
Educational Services	0.04	0.03	0.014
Financial and insurance activities	0.06	0.01	0.05
Health Care and Social Assistance	0.04	0.05	-0.01
Information and communication	0.05	0.04	0.01
Management of Companies and Enterprises	-	-	-
Manufacturing	0.03	0.01	0.02
Mining, Quarrying, and Oil and Gas	0.02	0.00	0.01
Other Services	0.04	0.18	-0.14
Professional, scientific and technical	0.04	0.29	-0.25
Public Administration	-	-	-
Real estate activities	0.04	0.04	0.00
Retail trade	0.02	0.02	0.00
Transportation and storage	0.03	0.05	-0.02
Utilities	0.02	0.00	0.02
Wholesale trade	0.02	0.08	-0.06
Total	0.03	0.07	-0.04

Notes: The figures in the table refer to the fraction of revenues spent on IT security investment. Data were unavailable for the Management of companies and enterprises, and Public administration sectors.

Table XXI
Summary of IT spending versus optimal spending across sectors

X. Conclusions

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

Despite the large and growing exposure to cyber risks, cyber costs are difficult to quantify. Using a unique database of more than 100,000 cyber events across sectors, 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), especially if the attack is not conducted on a large scale.

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 developing technological skills 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.

Crypto-related activities, which are largely unregulated, are associated with higher losses. We observe the existence of a positive correlation between the price of bitcoin and the intensity of crypto-related cyber events. With rising Bitcoin prices there is an increased incentive to attack exchanges, a vulnerable part of the cryptoasset ecosystem. Stronger regulation of the activities of intermediaries that operate in cryptoasset markets is likely necessary as the expected costs for cyber-related events are also significantly higher than for other events.

Finally, we do a first evaluation of over- or underspending on IT budgets by sector. 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.

References

Abadie A, Athey S, Imbens GW, Wooldridge J. 2017. When should you adjust standard errors for clustering? NBER Working Paper 24003.

Acemoglu D, Ozdaglar A, Tahbaz-Salehi A. 2015. Systemic risk and stability in financial networks. *American Economic Review* **105**: 564–608.

Albert R, Jeong H, Barabási AL. 2000. Error and attack tolerance of complex networks. *Nature* **406**: 378.

Aldasoro I, Gambacorta L, Giudici P, Leach T. 2020. Operational risks in the financial sector. BIS Working Papers 840, Bank for International Settlements.

Baldwin A, Gheyas I, Ioannidis C, Pym D, Williams J. 2017. Contagion in cyber security attacks. *Journal of the Operational Research Society* **68**: 780–791.

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

Beck U, Giddens A, Lash S. 1994. *Reflexive modernization: Politics, tradition and aesthetics in the modern social order*. Stanford University Press.

Beck U, Lash S, Wynne B. 1992. *Risk society: Towards a new modernity*, volume 17. London: Sage.

Biener C, Eling M, Wirfs JH. 2015. Insurability of cyber risk: An empirical analysis. *The Geneva Papers on Risk and Insurance-Issues and Practice* **40**: 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 AC, Miller DL. 2015. A practitioners guide to cluster-robust inference. *Journal of Human Resources* **50**: 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 EW. 2013. Effects of the financial crisis on banking operational losses. *The Journal of Operational Risk* **8**: 3.

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

Chande N, Yanchus D. 2019. The cyber incident landscape. Working Paper 32, Bank of Canada.

Chernobai A, Jorion P, Yu F. 2011. The determinants of operational risk in US financial institutions. *Journal of Financial and Quantitative Analysis* **46**: 1683–1725.

Cohen RD, Humphries J, Veau S, Francis R. 2019. An investigation of cyber loss data and its links to operational risk. *Journal of Operational Risk* **14**: 1–25.

Committee on Payments and Market Infrastructures. 2014. Cyber resilience in financial market infrastructures, Basel.

Cruz MG, Peters GW, Shevchenko PV. 2015. *Fundamental aspects of operational risk and insurance analytics: A handbook of operational risk*. John Wiley & Sons.

Curti F, Mihov A, Frame WS. 2019. Are the largest banking organizations operationally more risky? *Available at SSRN 3210206* .

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

Dreyer P, Jones T, Klima K, Oberholtzer J, Strong A, Welburn JW, Winkelman Z. 2018. Estimating the global cost of cyber risk. Research Reports RR-2299-WFHF, Rand Corporation.

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

Eisenbach TM, Kovner A, Lee MJ. 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, 2135–2148.

Etzioni A. 2011. Cybersecurity in the private sector. *Issues in Science and Technology* **28**: 58–62.

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

Goldstein J, Chernobai A, Benaroch M. 2011. An event study analysis of the economic impact of it operational risk and its subcategories. *Journal of the Association for Information Systems* **12**: 1.

Gordon LA, Loeb MP. 2002. The economics of information security investment. *ACM Transactions on Information and System Security (TISSEC)* **5**: 438–457.

Haldane AG, May RM. 2011. Systemic risk in banking ecosystems. *Nature* **469**: 351.

Hanseth O. 2007. Introduction: Integrationcomplexityrisk the making of information systems out of control. In *Risk, Complexity and ICT*, chapter 1. Edward Elgar Publishing.

Jacobs J. 2014. Analyzing ponemon cost of data breach. *Data Driven Security* **11**.

Jones J, Freund J. 2014. *Measuring and Managing Information Risk: A FAIR Approach*. Butterworth-Heinemann.

Kamiya S, Kang JK, Kim J, Milidonis A, Stulz RM. 2018. What is the impact of successful cyberattacks on target firms? NBER Working Paper 24409.

Kashyap AK, Wetherilt A. 2019. Some principles for regulating cyber risk. *AEA Papers and Proceedings* **109**: 482–87.

Kaufman LM. 2009. Data security in the world of cloud computing. *IEEE Security & Privacy* **7**: 61–64.

Kennedy DB, Stratopoulos TC. 2017. Mapping it spending across industry classifications: An open source dataset. *Available at SSRN 3073236* .

Kopp E, Kaffenberger L, Jenkinson N. 2017. Cyber risk, market failures, and financial stability. IMF Working Paper WP/17/185, International Monetary Fund.

Kshetri N, Voas J. 2017. Do crypto-currencies fuel ransomware? *IT professional* **19**: 11–15.

Lloyds. 2017. Counting the cost: Cyber exposure decoded. Emerging risks report, Lloyds.

Lloyds. 2018. Cloud down: Impacts on the us economy. Emerging risks report, Lloyds.

Makridis C, Dean B. 2018. Measuring the economic effects of data breaches on firm outcomes: Challenges and opportunities. *Journal of Economic and Social Measurement* **43**: 59–83.

Massad TG. 2019. Its time to strengthen the regulation of crypto-assets. M-RCBG Associate Working Paper 112, Mossavar-Rahmani Center for Business and Government.

Mell P, Grance T. 2011. The NIST definition of cloud computing (draft). *NIST Special Publication* **800**: 145.

Moore T, Christin N. 2013. Beware the middleman: Empirical analysis of bitcoin-exchange risk. In *International Conference on Financial Cryptography and Data Security*. Springer, 25–33.

Pretty D. 2018. Reputation risk in the cyber age: The impact on shareholder value. Technical report, Aon and Pentland Analytics.

Romanosky S. 2016. Examining the costs and causes of cyber incidents. *Journal of Cybersecurity* **2**: 121–135.

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

Rowe BR, Gallaher MP. 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. *Operational Risk* **2**: 21–22.

Urquhart J. 2009. The biggest cloud-computing issue of 2009 is trust.
URL `http://web.archive.org/web/20080207010024/http://www.808multimedia.com/winnt/kernel.htm`

Welburn JW, Strong A. 2019. Systemic cyber risk and aggregate impacts. Working Paper Series WR-1311, RAND Corporation.

Wolff J. 2014. Models for cybersecurity incident information sharing and reporting policies. In *TPRC 43: The 43rd Research Conference on Communication, Information and Internet Policy Paper*. TPRC.

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.

Wolff J, Lehr W. 2018. When cyber threats loom, what can state and local governments do? *Georgetown Journal of International Affairs* **19**: 67–75.

Zhu K, Weyant JP. 2003. Strategic decisions of new technology adoption under asymmetric information: a game-theoretic model. *Decision sciences* **34**: 643–675.