Time dynamics of cyber risk*

Martin Eling† Rustam Ibragimov‡ Dingchen Ning†

June 16, 2023

ABSTRACT

This paper utilizes three large databases to better understand the characteristics of cyber loss events, especially how to deal with data biases and how cyber losses evolve over time. We first deal with the problem of report delay with an extended two-stage model in combination with detailed information in our data. Then we analyze the frequency and severity of different categories of cyber events (such as malicious and negligent events) using state-of-art statistical methods for the detection of structural changes. We document that the frequency is increasing rapidly with the malicious cyber events growing exponentially in the past two decades but there is no significant change in loss severity. We also explore the tail dynamics and find that the heavy-tailedness of cyber events is persistent over time. Finally, we develop a conceptual model with the documented empirical features (delayed information and heavy-tailedness) and show that they lead to significantly lower insurance demand. This might help explain the low volume of the cyber insurance market observed today.

JEL classification: C15, G22, G32.

*We thank Glenn Harrison, Marcel Tyrell, Julia Holzapfel, and the participants at the 2022 American Risk and Insurance Annual Conference, 2022 European Group of Risk Insurance Economists Annual Conference, 2022 German Finance Association Annual Conference, 2022 Asia-Pacific Risk and Insurance Conference, and PiF seminar at the University of St. Gallen for their comments.

†Institute of Insurance Economics, University of St. Gallen, Switzerland

‡Imperial College Business School, UK

I. Introduction

In 2007, an American department store chain, TJX, was hacked and nearly 94 million credit card information has been exposed (Swartz 2007). This was the largest recorded data breach incident at the time, but just several years later, more and more data breach incidents exceeding this magnitude occur. Among them, Yahoo's incident in 2013 was the largest, involving nearly 3 billion user accounts (Stempel & Finkle 2017). Not only the extreme cyber events becoming more and more frequent, but the overall frequency and severity are also changing quickly. For example, FBI (2020) reports a 300% increase in reported cybercrimes during the COVID-19 period. The recent report of Smith & Lostri (2020) estimates the cost of global cybercrime at $1 trillion, a more than 50% increase from the 2018 estimate ($600 billion). Also, recent academic research (e.g., Jamilov et al. 2021) emphasizes that cyber risks have increased significantly globally.

Despite the anecdotal evidence illustrating the increasing importance of cyber risk as well as its dynamic nature, the empirical evidence in the current literature is still relatively limited. The theoretical work on cyber risk and information security has begun as early as the beginning of this century (e.g., Gordon & Loeb 2002), but due to the limit of data, the empirical work is at least one decade lagging behind with Maillart & Sornette (2010) among the earliest works to use data breach loss information.[1] Therefore, we intend to provide a comprehensive analysis of cyber loss events by utilizing three large cyber databases and discuss the implications of our empirical results. The two research questions are:

(RQ1): What are the statistical properties of cyber risk and how they are changing over time?
(RQ2): What are the implications for cyber risk management given the evolving cyber threat landscape?

To address the first question, we focus on three dimensions of cyber risk: frequency, severity, and tail risk. Before analyzing the time trend of cyber frequency, we first consider the report delay bias. This relates to the structural delay between the occurrence date and the observation date of an event, and there is little literature studying report delay for cyber risk due to the limit of data. Using the unique information in our data, we are able to correct this bias by extending a two-stage statistical model based on Stoner & Economou (2020). The results show that after accounting for report delay, the trend of frequency is increasing much faster than what we see in raw data.

Building on the results of bias correction, we study cyber risk frequency, especially to understand whether there have been fundamental shifts over the years. More specifically, we apply recent statistical methods (Baranowski et al. 2019) to detect the unknown number of change points in the time series data of cyber risk. We find that malicious cyber risk has undergone exponential growth in the past two decades without significant structural change.

[1] We acknowledge that information security has been an evergreen IT topic before this century, but few of them are based on the economic (and risk management) perspective. Therefore, we refer to Gordon & Loeb (2002) as one of the earliest papers in this area.

We also analyze the dynamics of cyber risk severity. We start by discussing the potential selection bias and use the implementation of U.S. data breach notification laws to test the severity of this bias. There is no strong evidence indicating that selection bias is a serious concern in our data. Then we move on to the analysis of cyber risk severity. Traditionally, the analysis of loss severity focuses on the first moment of the distribution, but this leaves out useful information. Following recent advances in statistics (Dubey & Müller 2020), we consider the full distribution of cyber loss, which provides a more comprehensive understanding. We do not find significant changes that lead to more severe losses. Rather more and more events with small losses occur, which prevents the overall loss distribution from shifting to the right.

Given the extreme nature of cyber risks and manifold discussions around their insurability (e.g., Biener et al. 2015), the tail of the loss severity distribution requires a deeper look. We apply two commonly used methods to measure the tail index: Hill's estimator and OLS log-log rank-size estimator, together with the optimal threshold selection method. We show that cyber risk is extremely heavy-tailed with infinite mean and variance in most cases. In addition, we propose a new change point detection method for the tail index based on Ibragimov & Müller (2016) and show that the tail index for malicious cyber events is decreasing, while the index for negligent cases is unchanged or even increasing.

To address the second research question, we discuss the implications for cyber risk management in light of the empirical evidence of cyber risk. We build upon and expand the classical model of Ehrlich & Becker (1972) where insurance and self-protection are the standard risk management options for a firm and incorporate two stylized properties of cyber risk: delayed information and heavy-tailedness. Due to the issue of report delay, the firm may have less accurate and delayed information compared to the insurer. This can lead to an underestimation of its risk level as malicious cyber risk is increasing exponentially. In addition, extreme heavy-tailedness limits the supply of cyber insurance (Ibragimov et al. 2009), and the tail exposure is borne by the firm. Based on these features, we show that the volume of the cyber insurance market is reduced, which is consistent with the evidence from Cellerini et al. (2022) that over 90% of cyber losses are not covered by insurance. Figure 1 presents an overview of our research questions and the main insights when addressing these questions.

The contribution of this paper is threefold. First, we uncover the empirical properties of cyber risk and show the dynamics of these properties, advancing the understanding of cyber risk in addition to the works such as Maillart & Sornette (2010) and Edwards et al. (2016). Second, we connect the empirical evidence of cyber risk with the theoretical work on information security (e.g., Gordon & Loeb 2002; Böhme et al. 2010; and Zhao et al. 2013) and show how the documented empirical properties can influence the optimal investment in cyber risk management. Lastly, in many related studies (Maillart & Sornette 2010, Wheatley et al. 2016, Farkas et al. 2021) the authors have questioned the reliability of data and discussed the potential issues that this can bring about. But there has no empirical evidence on this issue and we are the first to deal with data bias related to cyber risk.

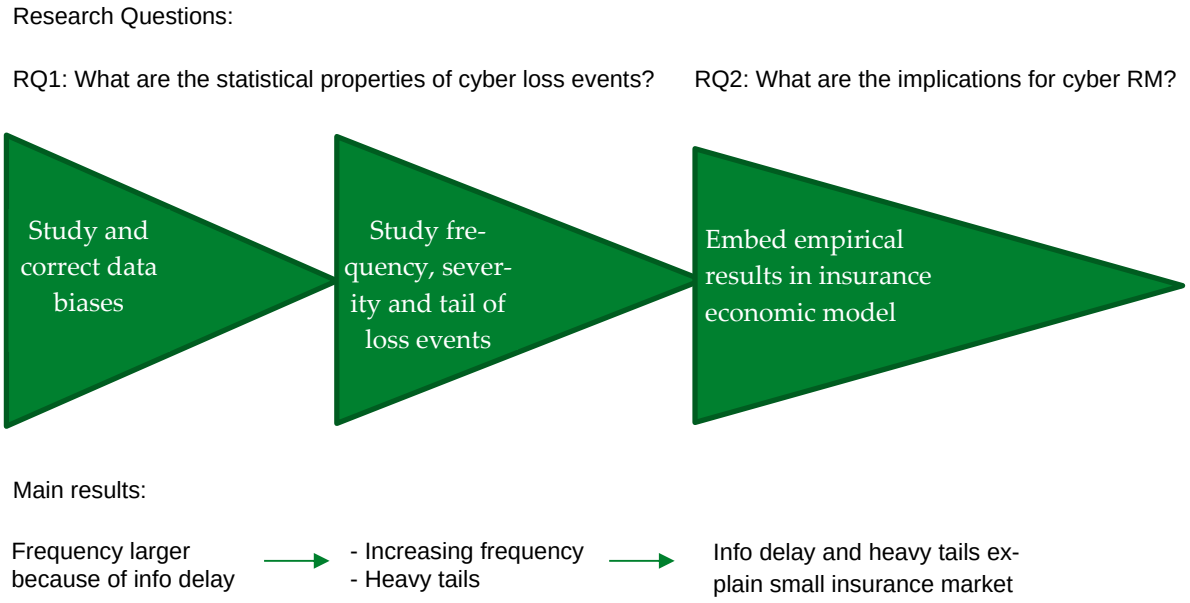

Figure 1. Overview of our paper

The remainder of the paper proceeds as follows. Section II provides more information on the background of cyber risk and related literature. Section III describes the data and the categorization of cyber risk. Section IV discusses the statistical methods used for the empirical analysis. Section V presents the main results on the dynamics of cyber risk. Section VI presents a basic model incorporating the documented properties of cyber risk and discusses the implications for cyber risk management. Finally, Section VII concludes.

II. Background

Cyber risk encompasses a broad range of risks to information and information systems, such as data breaches, ransomware, and system errors. We define cyber risk as "operational risks to information and technology assets that have consequences affecting the confidentiality, availability or integrity of information or information systems" (Cebula & Young 2010). Therefore, we consider cyber risk as a subcategory of operational risk, which enables us to distinguish cyber risk from other established risk categories and structure cyber risk in line with the classification of operational risk (Eling & Wirfs 2019).

A. Empirical work on cyber risk

Although there have been works with empirical data before 2010, the data are not actual cyber events but cyberattack attempts without information on the realization of such attempts (e.g., Böhme & Kataria 2006). Maillart & Sornette (2010) is the earliest empirical work on cyber risk analysis with actual cyber events data, to the best of our knowledge. In addition, we do not

include the empirical work on estimating the financial impact of cyber events based on event study approaches (e.g., Kamiya et al. 2021) since they do not focus on the statistical properties of cyber risk per se.

The early stage of the empirical work focuses on the general statistical properties of cyber risk, including correlation structure (Böhme & Kataria 2006; Wang & Kim 2009*a*; and Wang & Kim 2009*b*) and time trends (Maillart & Sornette 2010; Wheatley et al. 2016; Edwards et al. 2016; and Romanosky 2016). Starting from Eling & Loperfido (2017), more and more studies begin to study cyber risk frequency and severity by fitting existing statistical models (Eling & Wirfs 2019; and Woods et al. 2021) or proposing new frameworks to model cyber risk (Bessy-Roland et al. 2021; Farkas et al. 2021; Sun et al. 2021; Fang et al. 2021; and Zhang Wu et al. 2021). These works have exploited the available database to show the good performance of their models, and the basic consensus is that the modeling of severity should be based on heavy-tailed (at least highly right-skewed[2]) distributions, although the specific choice of the model is very diverse.[3]

Still, the study on time dynamics of cyber risk has been scarce and results are inconsistent. For example, with data period from 2000 to 2008, Maillart & Sornette (2010) show there is a strong non-stationary growth culminating in July 2006 followed by a stable period afterward. Edwards et al. (2016) find no evidence of an increasing trend for the size and frequency of data breaches for data from 2005 to 2015. However, Romanosky (2016) indicates an increasing trend for the number of cyber events in the same period. Wheatley et al. (2021) also observe an increasing trend for both frequency and severity in a similar time period, but only specific to hack-type events. More recently, Jung (2021) shows a breakpoint in 2014 for loss severity data with a stable trend before 2014 and rapid growth afterward. Overall, the results appear to be rather inconsistent and the difference might be largely driven by different datasets and methodologies. This motivates us to reconsider the empirical properties over a long horizon with the comparison of three main cyber databases. We also note that none of the above studies tries to incorporate the bias problems, which are inherent to all these datasets.

B. The economics of information security and insurance

To deal with cyber risk, it is critical to invest in information security. Gordon & Loeb (2002) is the first to consider an economic model for the optimal investment in information security and shows that firms should prioritize the information assets with midrange vulnerabilities as the cost of increasing safety level can be nonlinear. However, information security does not only depend on the efforts of one firm. There is an externality for the investment in security as the security level depends on the minimum effort any firm makes in the same system (Anderson & Moore 2006). This interdependence among firms is a key feature in shaping the security of information systems. Due to the complexity of the network structure in the system, game theory is a commonly used

[2]For example, the results of Woods et al. (2021) show the gamma distribution has better performance, which is not heavy-tailed distribution but exhibits high skewness.

[3]A detailed summary of each paper is presented in Appendix .A

tool to model the interactions among participants (Laszka et al. 2014).

Among the possible risk management mechanisms, insurance is the most studied remedy in the literature. Gordon et al. (2003) proposes a general framework for cyber insurance by incorporating the typical information asymmetry issues related to insurance. Later on, Böhme et al. (2010) provides a more comprehensive framework for considering all the peculiarities of cyber risk: interdependent security, correlated risk, and information asymmetries. In this framework, there has been extensive literature on the optimal security level with insurance, such as cyber insurance as an incentive (Bolot & Lelarge 2009), competitive market and security level (Shetty et al. 2010), self-insurance and self-protection (Johnson et al. 2011), managed security services (IT security outsourcing) as an alternative (Zhao et al. 2013), and fines and rebate corrective treatment (Naghizadeh & Liu 2014), etc. [4]

However, there is little research connecting the theoretical work on information security with the empirical work on cyber risk. This motivates us to fill the gap by first documenting the empirical properties of cyber risk and then discussing the implications for cyber risk management and information security.

III. Data

A. Data on cyber loss

We look at three sources of data for the analysis of cyber risk. All data focus on events that occur to legal entities (firms, public and non-profit institutions) rather than individuals, and contain two types of losses. The first type is the amount of information measured by the number of records or affected accounts. The second type is the monetary loss arising from the incident, such as first-party loss including the value of the lost records or the cost of business interruption, and third-party loss including the payment to affected customers and fines in case of violation of regulations.

The first data source is from Advisen. Their database collects information from multiple publicly available sources such as government websites (Securities & Exchange Commission, Federal Trade Commission, Federal Communications Commission, State data breach notification websites, etc.) and other sources including keyword-based alerts, official court and litigation sources, and other internet information. The magnitude of the observations in the database is over 150,000, while more than 80% of the cases are from the U.S. and the rest are from 177 different countries. Since the database creates different records for different kinds of losses arising from one incident such as direct damage and legal costs, we aggregate the original data resulting in 111,253 incidents for further analysis. Although the magnitude of cyber events in this database is large, the information on financial loss is relatively scarce. After cleaning the data and using the sample after 2001,[5] we have 5,789 incidents with known financial loss and 90,821 incidents with the known number of

[4] See Marotta et al. (2017) for a more comprehensive summary of cyber insurance literature.

[5] We restrict the sample to the time period from 2001 since cyber risk only becomes a serious issue in the 21st century and the data in the last century are very sparse. This also applies to other data sources.

affected accounts.

The second data source is SAS OpRisk Global data, which is the world's largest database on publicly reported operational losses. This database contains more than 35,000 operational events in excess of US$ 100,000 for different countries and industries. There is no classification for cyber risk, so we cannot extract cyber events directly from the data. Therefore, we follow Eling & Wirfs (2019) and exploit text mining to extract cyber-related events. This results in 2,123 observations of cyber events and 18,287 observations of non-cyber operational events after removing the sample before 2001 (we refer to this subsample as the operational risk sample afterward) in our analysis.

The last source of data is from the non-profit organization Privacy Rights Clearinghouse (PRC), a database that has been frequently used in the literature (e.g., Kamiya et al. 2021; Farkas et al. 2021; and Bessy-Roland et al. 2021). It collects information about breach events from government agencies and verifiable news sources starting from 2005. This data contains 6,733 records (with non-zero ID losses) up to the end of 2019. The major difference from the previous two data sources is that this database focuses only on data breach events and does not provide the financial loss amount for each case. Therefore, we will use this database for the analysis of risk frequency and the number of records breached.

Although there are three different databases, they are connected with each other as they all focus on the same area, cyber risk. We study these databases separately without merging them since we aim to find the general pattern of cyber risk that is persistent across different sources and categories, and the comparison of three data sources reduces the potential bias in each source of data.

B. Categorization and descriptive statistics

To disentangle cyber incidents of different kinds, we consider four categories following and expanding existing literature such as Edwards et al. (2016). The first category covers malicious cases, which are defined as cyber incidents that are initiated by individuals or institutions with malignant intentions against the victims, such as hackers infiltrating the system of the victim firms or internal employees stealing confidential information to gain profits. The second category is negligent cases and includes cyber incidents that do not involve a malicious third party, such as systems errors due to negligence or unintentional disclosure of consumer information. The third category is the violation of privacy (we refer to this as "privacy" afterward). This category is defined as the cyber incidents that involve the intentional misbehavior of the firm regarding the information of its counterparties (mostly their clients), either individuals or organizations. This type of incident does not have direct damages to the firms but may lead to lawsuits and fines as a violation of laws or regulations related to privacy. The last category includes all remaining cyber incidents, and we refer to them as "others". This category includes cases with unknown reasons or the ones that are directly caused by factors outside cyberspace such as natural disasters or physical damage. Details on the categories can be found in Appendix .E.

Table I summarizes the key statistics of each type of cyber incident from three databases

(and the operational risk sample in SAS). In Advisen, there are only around 5% to 10% of the incidents among each category where the information on financial losses is known, and around 80% of incidents where the information on the affected counts is available. The financial loss of malicious cases is more severe compared to the loss of negligent cases, but the affected counts are higher for negligent cases compared to malicious ones, although the difference is not large.

In SAS, we find that the magnitude of losses is higher than the case in Advisen, which is related to the fact that the SAS data only includes cases with losses of more than US$ 100,000. Also, negligent cases lead to higher loss severity compared to malicious cases, which is different from the Advisen data. This might be explained by the composition of the two samples, as the Advisen sample includes more cases related to unintentional disclosure which results in lower financial losses compared to others.

In the PRC data, malicious cases have a higher number of records lost or affected relative to negligent cases. This is different from the Advisen data as the PRC data focuses mainly on data breaches and thus has in general a higher loss related to the number of records. In addition, the term "affected counts" and "number of records" in this section are basically the same. There is only a slight difference in the case of privacy violation where "affected counts" includes the number of individuals whose personal information was not breached but misused. As we focus on the malicious and negligent cases in the following analysis, this difference should not affect our comparison. Albeit of the difference in the size and type of losses across different databases, the loss distributions are all heavily skewed as the median and mean values are different.

IV. Methodology

A. Report delay

Reliable data are crucial for the analysis of cyber risk, but the current databases are potentially biased (such as the database of Advisen and other commercial databases). Hence, empirical studies without bias correction may only lead to partial or even incorrect conclusions about cyber risk.

We aim to apply recent methods from the field of statistics to identify and correct the potential bias in the data before conducting further statistical analysis. One main problem is report delay, which is the case where the total observable number will only be available after a period of time. Therefore, before the total number becomes available, we can only observe incomplete data. This can be detrimental to the analysis of time dynamics and lead to misinterpretation of the actual number of events. In the case of cyber risk, this problem is common since many events are noticed and made public after a long time. Also, a delay may occur when the database cannot update the records in time due to limited resources invested in maintenance.

To model report delay, we follow the work of Stoner & Economou (2020) and extend their framework to include two stages that are unique in the Advisen dataset.[6] The Advisen dataset is

[6]The problem of report delay is closely related to the claims reserves problem in actuarial science. Two of the most common methods in the area are the distribution-free chain-ladder model (Mack 1993), and the overdispersed

Table I Summary statistics of three databases

	Sample size	Number of cases with known losses	First quantile	Mean	Median	Third quantile	Standard deviation
Advisen-loss amount ($Million)							
Malicious	53317	2476	0.02	17.83	0.15	1.11	271.39
Negligent	17845	357	0.02	15.21	0.12	0.61	123.01
Privacy	36285	2738	0.01	6.68	0.05	1.26	117.01
Others	9607	218	0.02	13.72	0.18	0.83	169.52
Advisen-affected counts (Million)							
Malicious	53317	38985	0.00	1.17	0.00	0.00	33.54
Negligent	17845	14500	0.00	1.81	0.00	0.00	84.89
Privacy	36285	29270	0.00	0.25	0.00	0.00	14.53
Others	9607	8066	0.00	0.08	0.00	0.00	1.43
SAS-loss amount ($Million)							
Malicious	1451	1451	0.40	27.21	1.23	4.34	260.86
Negligent	516	516	0.48	57.21	2.70	16.30	247.94
Privacy	80	80	0.80	20.48	4.48	21.22	35.42
Others	76	76	0.36	40.69	1.26	6.80	214.42
SAS (operational risk)-loss amount ($Million)							
Malicious	5736	5736	0.45	40.99	1.58	7.18	383.01
Negligent	12551	12551	1.09	116.85	5.08	28.41	1100.80
PRC-number of records (Million)							
Malicious	3207	2011	0.00	3.63	0.00	0.02	70.15
Negligent	1861	1553	0.00	1.69	0.00	0.01	36.61
Others	3858	3169	0.00	0.14	0.00	0.01	2.63

Note:
This table presents the summary statistics of different cyber categories from three databases (including the operational risk data from SAS as a comparison afterward).

the main focus in this part since it has the detailed timeline of each incident, from the event date to the date of the first notice, until the date of entry into the database. This unique feature allows us to capture two delay mechanisms.[7]

The reason we choose the method from Stoner & Economou (2020) is that it provides high accuracy by jointly modeling the delay mechanism and the total count number. Traditionally, the task of correcting the delayed reporting has been separated from the task of forecasting but this ignores the joint uncertainty in the incidence of the total count and the presence of delay. For example, a low number of cyber cases in month t may have resulted from a temporal decreasing trend or a low reported number in this period or both. Therefore, it is important to jointly model these two mechanisms.

Three models are considered in this paper, a generalized linear model (GLM) (Salmon et al. 2015), a generalized Dirichlet-multinomial hazard model (GDM hazard), and a generalized Dirichlet-multinomial survivor model (GDM survivor) (Stoner & Economou 2020). In the empirical part, we first compare the three models for their in-sample performance and then apply the best model for bias correction.

Let y_t be the total observable count at time t and after some delay unit (months in our case) a proportion of y_t, $z_{t,d}$, has been reported in this period, where d is the number of months delayed. This means that $\sum_{d=1}^{D} z_{t,d}$ gets close to y_t as the total number of months D increases.

The model based on the GLM framework starts with a negative-binomial (NB) distribution for y_t:

$$y_t \sim NB(\lambda_t, \theta); \quad log(p_{t,d}) = g(t,d),$$

where λ_t is the expected rate of occurrences and θ allows for overdispersion, the multinomial probability $p_{t,d}$, which is the expected proportion of y_t that will be reported at delay d, is modeled via a log-link, and $g(t,d)$ represents a combination of covariate effects. Therefore, the marginal distribution for z_i is also NB:

$$z_{t,d} \sim NB(\mu_{t,d} = p_{t,d}\lambda_t, \theta); \quad log(\mu_{t,d}) = \iota + \alpha_t + \eta_t + \psi_d + \beta_{t,d},$$

where α_t is a penalized cubic spline to capture nonseasonal variation, η_t is a penalized cyclic cubic spline to capture within-year temporal effect, $\beta_{t,d}$ is intended to allow for temporal changes of delay mechanism, and ι and ψ_d are fixed effects.

Poisson model (Renshaw & Verrall 1998). A more detailed summary of the literature in actuarial science can be found in (Taylor 2019). There are many works generalizing these two models, and it is easy to reach the GLM model we mention later from Mack's work. Therefore, the two areas are connected, but there are also differences. One of them is that the focus of actuarial science is on the aggregate claim amount which is the multiplication of the number of claims and severity of claims, while the report delay problem mostly focuses on the number or frequency of the events/cases. In our case, the information on the financial loss of the events is scarce compared to the number of events, thus we only focus on the report delay issue for the frequency data in this section.

[7]SAS OpRisk database only has the date of occurrence (the year when the incident started) and the date of entry, while the PRC database contains only the date of occurrence. Therefore, we choose Advisen data for the main analysis and SAS data for comparison.

Different from the GLM framework, the models based on GDM are designed to account for heterogeneity in the delay mechanism and appropriately separate variability and uncertainty in the delay mechanism from the model of count number. The GDM hazard model is defined by:

$$\begin{aligned} y_t \sim NB(\lambda_t, \theta); \quad & log(\lambda_t) = \iota + \alpha_t + \eta_t; \\ z_t \mid y_t \sim GDM(\boldsymbol{\nu_t}, \boldsymbol{\phi}, y_t); \quad & log(\frac{\nu_{t,d}}{1-\nu_{t,d}}) = \psi_d + \beta_{t,d}, \end{aligned}$$

where $\nu_{t,d}$ is the expected proportion of counts that will be reported at delay d out of those which are yet-to-be-reported and $\boldsymbol{\phi}$ controls for dispersion. In this model, the delay mechanism is modeled through the difference of temporal structure in the proportion of reported cases across delay levels.

The GDM survivor model applies a different way of modeling the delay mechanism:

$$\begin{aligned} y_t \sim NB(\lambda_t, \theta); \quad & log(\lambda_t) = \iota + \alpha_t + \eta_t; \\ z_t \mid y_t \sim GDM(\boldsymbol{\nu_t}, \boldsymbol{\phi}, y_t); \quad & probit(S_{t,d} = \psi_d + \beta_t); \\ & \nu_{t,d} = \frac{S_{t,d} - S_{t,d-1}}{1 - S_{t,d-1}}, \end{aligned}$$

where $S_{t,d}$ is the expected value of the cumulative proportion of cases at time t for delay level d. Compared with the hazard model that considers a structure for each delay level, this method models the delay structure for each time point, which allows for any number of delay levels.

The models above provide flexible ways of modeling delay structures for cyber risk, but how to connect two delay stages in our cyber risk data remains a problem. Given that the data we have are at the second stage as defined above, we could back trace the original trend with available data.

In the second stage, assume that for the time of first notice t, the number of total cases is a_t but is not fully available. Suppose after D months all the cases will be included in the database, but for now, we only have data for D' months. Therefore, after applying the methods defined above, we can estimate the number of total cases as

$$\hat{a_t} = \sum_{1}^{D'} a_{t,d} + \sum_{D'+1}^{D} \hat{a_{t,d}},$$

where $a_{t,d}$ is the number of cases reported in delay time d, while $\hat{a_{t,d}}$ is the estimated number of cases in delay time d.

Additionally, the correction ratio q_t is defined as the estimate of the actual total number divided by the available number at time t:

$$q_t = \hat{a_t} / \sum_{1}^{D'} a_{t,d}.$$

This correction ratio can be further applied to the first stage. When considering the delay structure between the accident date and the first notice date, the number of cases reported $b_{t,d}$ is biased due to the delay in the second stage. Therefore, we can adjust this bias with the correction ratio:

$\tilde{b_{t,d}} = b_{t,d} * q_{t+d}$. After the adjustment, we apply the models above to the database we have to account for first-stage bias, which provides us with the corrected results of cyber risk.

B. Time dynamics of loss frequency

We study loss frequency and in this context focus on the estimation of change points over the period. There is extensive literature on change points detection methods (Truong et al. 2020), which can be categorized based on their cost functions, search methods, and constraints. But the literature mostly focuses on the problem under the assumption of piecewise-constant parameters. However, cyber loss frequency is not likely to follow this assumption due to the increasing trend.

Therefore, we consider one newly proposed generic approach of detecting an unknown number of features occurring at unknown locations, narrowest-over-threshold detection (Baranowski et al. 2019). [8] This method shows low computational complexity, ease of implementation, and accuracy in the detection of the feature locations while allowing for non-constant time trends.

In this method, consider the model

$$Y_t = f_t + \sigma_t \epsilon_t, \quad t = 1, ...T,$$

where f_t is the signal, σ_t is the noise's standard deviation at time t, and ϵ_t follows standard normal distribution. We further assume that (f_t, σ_t) can be divided into $q+1$ segments with q unknown unique change points $0 = \tau_0 < \tau_1 < ... < \tau_q < \tau_{q+1} = T$. The structure of (f_t, σ_t) is modeled parametrically by a local real-valued d-dimensional parameter vector Θ_j, where d is known and typically small.

In the first step, we randomly draw subsamples such as $(Y_{s+1}, ..., Y_e)'$, where (s, e) is drawn uniformly from the set of pairs of indices in $\{0, ..., T-1\} \times \{1, ...T\}$. The generalized likelihood ratio (GLR) statistic for all potential single change points within the subsample is

$$\mathcal{R}^b_{(s,e]} = 2log[\frac{sup_{\Theta^1,\Theta^2}\{l(Y_{s+1}, ..., Y_b; \Theta^1) l(Y_{b+1}, ..., Y_e; \Theta^2)\}}{sup_{\Theta} l(Y_{s+1}, ...Y_e; \Theta)}],$$

where $l(Y_{s+1}, ...Y_e; \Theta)$ is the likelihood of Θ given $(Y_{s+1}, ..., Y_e)'$. Based on this statistic, we pick the maximum $\mathcal{R}_{(s,e]}(Y) = max_{b \in \{s+d, ...e-d\}} \mathcal{R}^b_{(s,e]}$.

In the next step, all $\mathcal{R}_{(s_m, e_m]}(Y)$ for $m = 1, ...M$ is tested against a given threshold and among the significant results, the one corresponding to the interval $(s_{m^*}, e_{m^*}]$ with the smallest length will be chosen. This step can be repeated recursively to find all the possible change points. For more technical details, we refer to Baranowski et al. (2019).

C. Time dynamics of loss severity

Traditionally, the analysis of loss amount in the time dimension is reduced to the analysis of univariate time series such as average loss severity. Although this is a simple and efficient way

[8] We compare the results of alternative methods in Appendix .D and show the main method is robust.

of understanding the dynamics of loss, we are leaving out too much information in this process. Therefore, in this paper, we adopt the recently developed method to analyze the change point in a sequence of distributions.

Dubey & Müller (2020) considers a sequence of independent random objects Y_t taking values in a metric space (Ω, d) rather than in $\mathbb{R}$ as in traditional methods (Niu et al. 2016). As in most practical situations, the differences in distributions are mostly in location or in scale. Therefore, this method aims to detect differences in means and variances which are in Fréchet type and provides a generalization of the notion of location and scale to metric spaces.

The test statistic for the change point can be written as:

$$T_n(b) = \frac{b(1-b)}{\hat{\sigma}^2}\{(\hat{V}_{[0,b]} - \hat{V_{[b,1]}})^2 + (\hat{V_{[0,b]}}^C - \hat{V_{[0,b]}} + \hat{V_{[b,1]}}^C - \hat{V}_{[b,1]})^2\},$$

where b is the possible value of the change point, $\hat{\sigma}$ is the asymptotic variance of the empirical Fréchet variance, $\hat{V_{[i,j]}}$ is the estimated Fréchet variance and lastly, $\hat{V_{[i,j]}}^C$ is the "contaminated" version of Fréchet variance obtained by plugging in the Fréchet mean from the complementary data segment.

Based on this test statistic, Dubey & Müller (2020) further provide an inference method for the identification of change points in a sequence of distributions. There are two ways of calculating critical values, Monte Carlo simulations or the bootstrap approximation. The latter one is shown to yield higher critical values and thus is more conservative. We refer to Dubey & Müller (2020) for more technical details.

D. Time dynamics of tail risk

Tail risk is an important part of the analysis for cyber risk, especially in the sense that extreme tail risk or heavy-tailedness has many unfavorable properties such as inducing non-diversification trap(Ibragimov et al. 2009).[9]

In models considering a heavy-tailed risk, the variable of interest r, cyber loss in our case, is usually assumed to have a distribution with power tails, such that $P(r > x) \sim \frac{C}{x^\zeta}$, $C > 0$, as $x \to +\infty$. The parameter ζ is the tail index. This index characterizes the heaviness of the tail of the distribution and the smaller the index, the greater the probability mass in the tail. The tail index is linked to the existence of the moments. For example, the variance of r is finite if and only if $\zeta > 2$, and the mean is only finite if and only if $\zeta > 1$.[10]

Estimation of tail index: we consider two basic non-parametric methods which are widely used in the literature. The first one is the Hill's estimator as follows (Hill 1975):

[9] When risk distributions have heavy left tails and insurance providers have limited liability, insurance providers may choose not to offer insurance for catastrophic risks and not to participate in reinsurance markets, even though there is a large enough market capacity.

[10] The definition of the tail index may differ from other papers, as this value is the inverse of the shape parameter in generalized Pareto distribution. To make sure both Hill and OLS log-log rank-size estimator report the same value, we define the tail index in this manner.

$$\zeta(k) = \{\frac{1}{k}\sum_{j=1}^{k} ln(x(n-j+1)) - ln(x(n-k))\}^{-1},$$

where $x(i)$ is the ith-order statistic such that $x(i) \geq x(i-1)$ for $i = 2, ...n$.

The second method is OLS log-log rank-size regression (OLS estimator). We use the revised version proposed by Gabaix & Ibragimov (2011) which is consistent in small samples:

$$log(Rank - 1/2) = a - \zeta log(Size).$$

The two methods above are applied to the tail of the distribution for the estimation, but a key issue remains: the selection of the threshold for the tail. There are many methods to select the optimal threshold, we consider the R package "tea" from Ossberger (2020), which includes 12 different approaches. We conduct the simulation to find the suitable approaches for our purpose and 2 methods ("dAMSE" and "hall") provide better performance compared to other methods in the package. But there is a downward bias in these two methods and we find using the OLS estimator significantly reduces the bias (details in Appendix .B.1). Therefore, we use these two methods in combination with the OLS estimator for the estimation of the tail index.

Change point detection: To further analyze the trend or potential change points in the extreme value index, we rely on Ibragimov & Müller (2016). The empirical strategy is to partition the sample into two periods, the period before a possible break point, i, and the period after the point, j. Then we divide each period into q groups chronologically, and compute the Behrens-Fisher statistic:

$$BF = \frac{\hat{\xi_1} - \hat{\xi_2}}{\sqrt{\frac{(s_1)^2}{q_1} + \frac{(s_2)^2}{q_2}}},$$

where $\hat{\xi_i} = q_i^{-1}\sum_{j=1}^{q_1}\xi_{i,j}$, $(s_i)^2 = (q_i - 1)^{-1}\sum_{j=1}^{q_i}(\xi_{i,j} - \hat{\xi_i})^2$, and $\xi_{i,j}$ is the tail estimator.

We then compare the BF statistic with the critical value of the Student-t distribution with $min(q_1, q_2) - 1$ degrees of freedom. This allows us to detect whether there is a change point for the time series data. Together with the optimal threshold selection methods, we can identify the possible change points for the tail index of cyber risk.

V. Empirical results

A. *Report delay*

To understand the problem of report delay, we first briefly compare our three databases. To ensure the comparability of different databases, we restrict the time period to start from 2005 and focus only on the malicious category in the U.S. (the PRC data starts from 2005 and covers only U.S. data). [11]

[11] There are still some issues affecting the reliability of comparison. First, there is no exact accident date in SAS data, so certain biases may exist when compared with other databases. For the PRC data, because of the compulsory

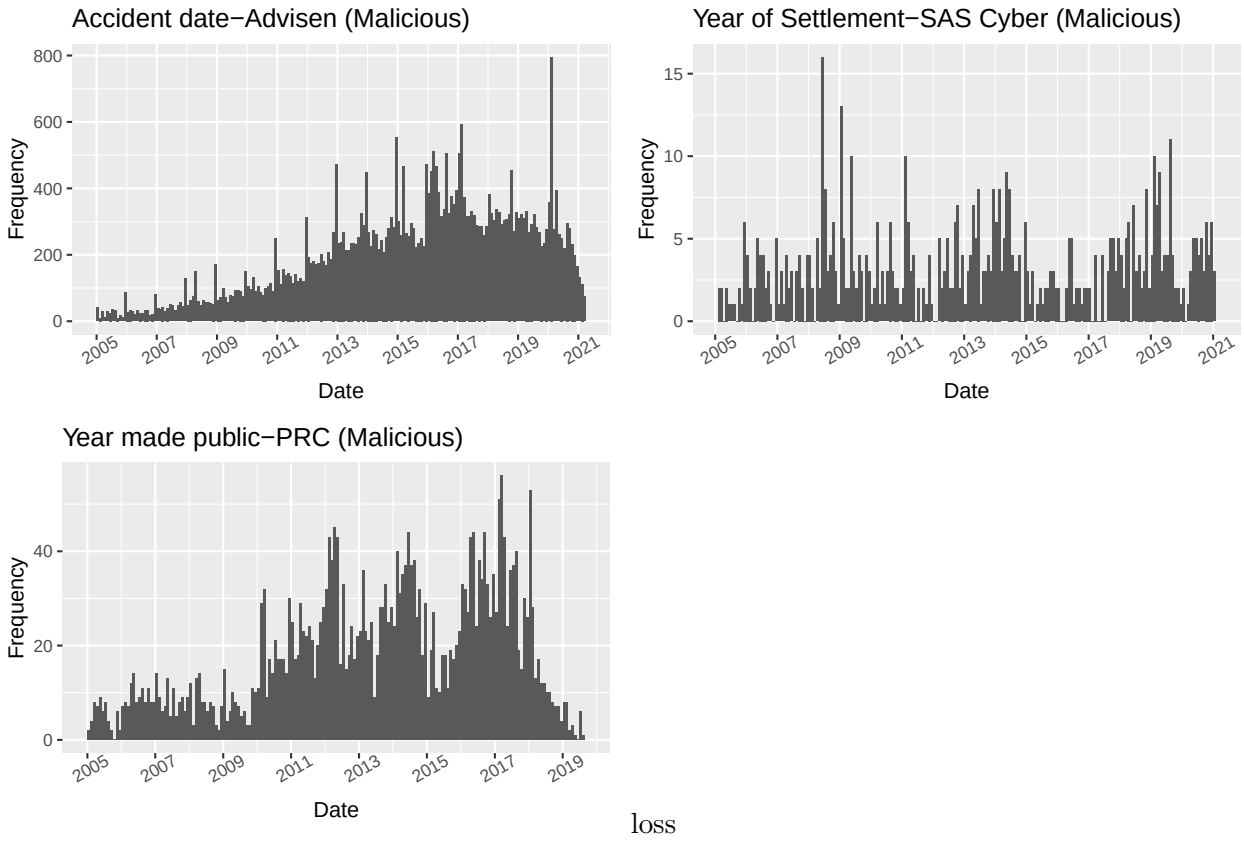

Figure 2. Different datasets of cyber risk

Notes: This figure reports the monthly frequency of malicious cyber events in three main databases. The abnormal and periodic peaks in the Advisen data are related to the inaccuracy of the accident date. For an event with only a known accident year, the database assigns the first day of the year as its estimated date.

Various sources and reports (Allianz 2021, Accenture 2021) suggest that cyber risk is increasing quickly over the years, but as shown in Figure 2, the increasing trend is not as obvious as we would expect. For example, the data from SAS show a steady trend, while the other two indicate an increasing trend during the early stage and then a steady trend in recent years. However, the sudden drop in the number of cases in 2018 for PRC and in 2020 for Advisen indicate that the problem of report delay may be one of the reasons behind this.

To look into the problem of report delay more deeply, we make use of the date of creation in Advisen to show how the trend evolves over the years in Figure 3. We plot the evolution of cyber risk based on four creation dates (every four years from 2009 to 2021) so that only cyber events before the creation date are included in each graph. This provides a clear comparison of different points in time and shows that at each point there is a clear decreasing trend which undoubtedly relates to delayed report.

disclosure of data breaches, the difference between the time when the event was made public and the accident date should not be large. Second, another point that may affect the comparison is that cyber events in PRC are mostly about data breaches while the other two include all kinds of cyber risk.

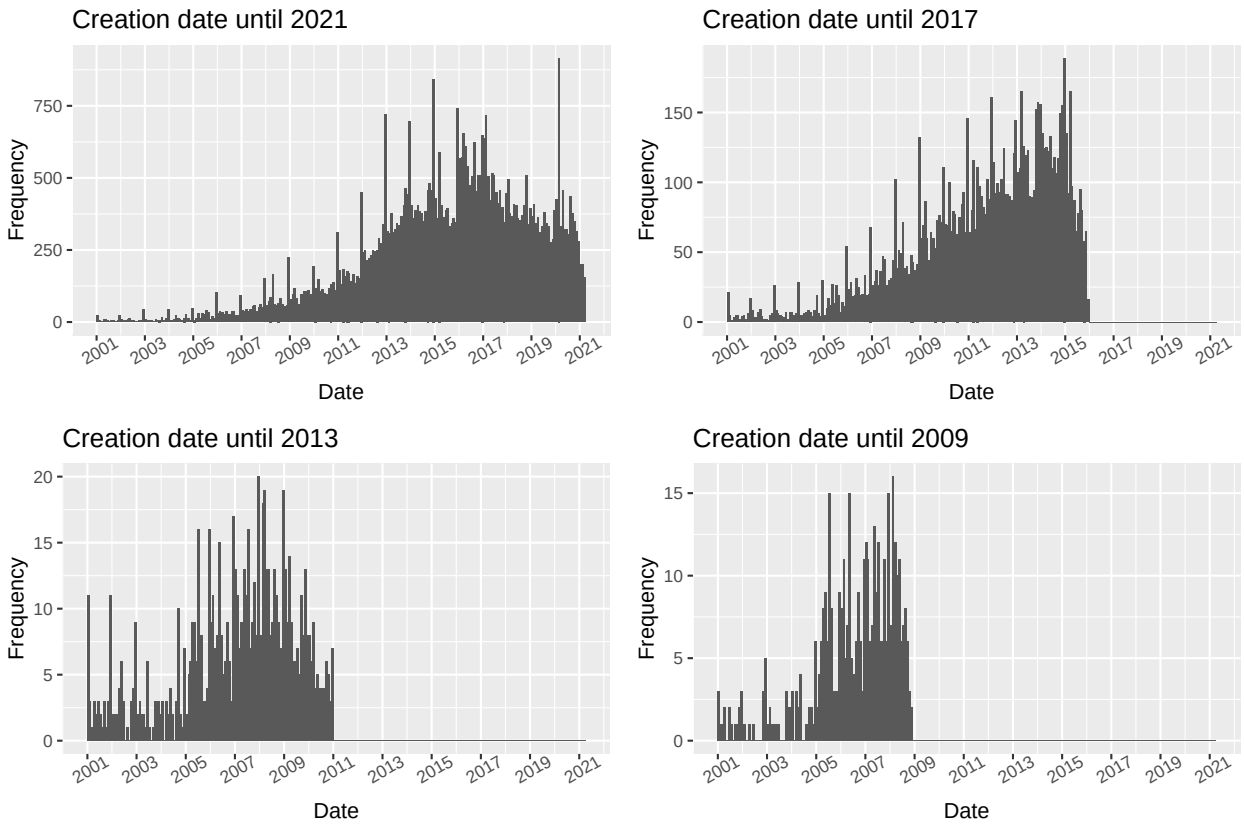

Figure 3. Different dates in Advisen (Malicious)

Notes: This figure reports the monthly frequency of malicious cyber events in Advisen, depending on the time when the events are included in this database.

In general, the process of collecting data related to cyber risk can be divided into two stages. The first stage is from the accident date to the date of the first notice. This period can be short for some types of events, such as cyber extortion or malfunction of devices, which the victims would notice almost immediately. But for other types including data breaches, the firms may take as long as months or years to find out that their data have been compromised. In general, the mean days of delay is 182 and the median is 33 days in the Advisen data.

The second stage starts with the date of the first notice and ends with the creation date in the database of concern. The time delay in this stage is mainly related to the efficiency of the database of concern, in some cases, the staff can update the data immediately but more likely there will be a moderate amount of delay in this stage, constrained by the investment of this database. In the Advisen data, the delay in this stage is much more severe than the first stage, with mean and median delayed days of 836 and 538. The major reason for this delay is that although the Advisen database begins to collect data in 2007, the majority of their events are created in recent years, especially during 2016-2018.

A.1. Bias correction for the Advisen data

We first conduct an in-sample analysis (see details in Appendix .C) to compare the performance of the three methods mentioned in the methodology part, and it is shown that the GDM hazard method has the best performance. Therefore, we apply the two-stage method with GDM hazard to the whole sample period. The result is shown in Figure 4. The increasing trend for the malicious, negligent, and "others" category is clear, although the number of malicious and negligent cases is increasing much faster compared to the "others" category. The exception is the privacy category. There is a peak around 2017 and then the number of cases decreases significantly. A more detailed analysis of the trends and change points will be discussed in the following section. After correcting the report delay problem, we can find for most of the cases the increasing trend becomes apparent compared to the raw data, indicating the necessity of our bias correction procedure.

A.2. Bias correction for the SAS data

The main analysis of the report delay problem is based on the Advisen data since it has detailed information on the time dimension. To validate the results from this database, we further apply the method above to another dataset-SAS. However, since SAS data only have information on the yearly level about the date of occurrence, we use this data as a robustness check only. As shown in Figure 5, only the malicious cyber cases exhibit a clearly increasing pattern, while the negligent cyber cases are relatively stable. This is different from the results we see in Advisen, and the possible reason is related to the fact that there are more incidents of unintentional disclosure in Advisen which are on the rise and affect the overall trend. In addition, we can find the trend of operational risk is decreasing even after the bias correction, which is in sharp contrast to the emerging cyber risk.

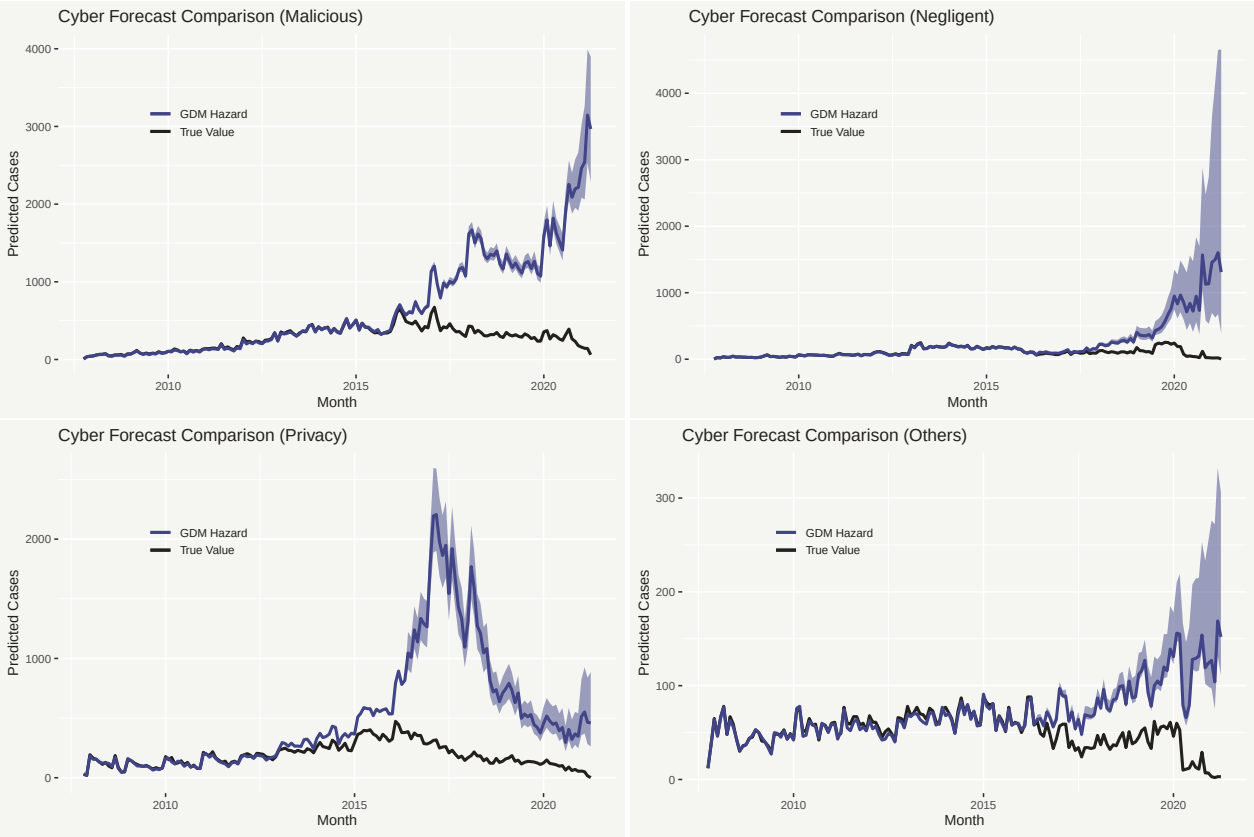

Figure 4. Bias correction for the Advisen data

Notes: This figure shows the forecast results of cyber incidents with the 95% confidence interval after adjusting the report delay problem for different categories in the Advisen data.

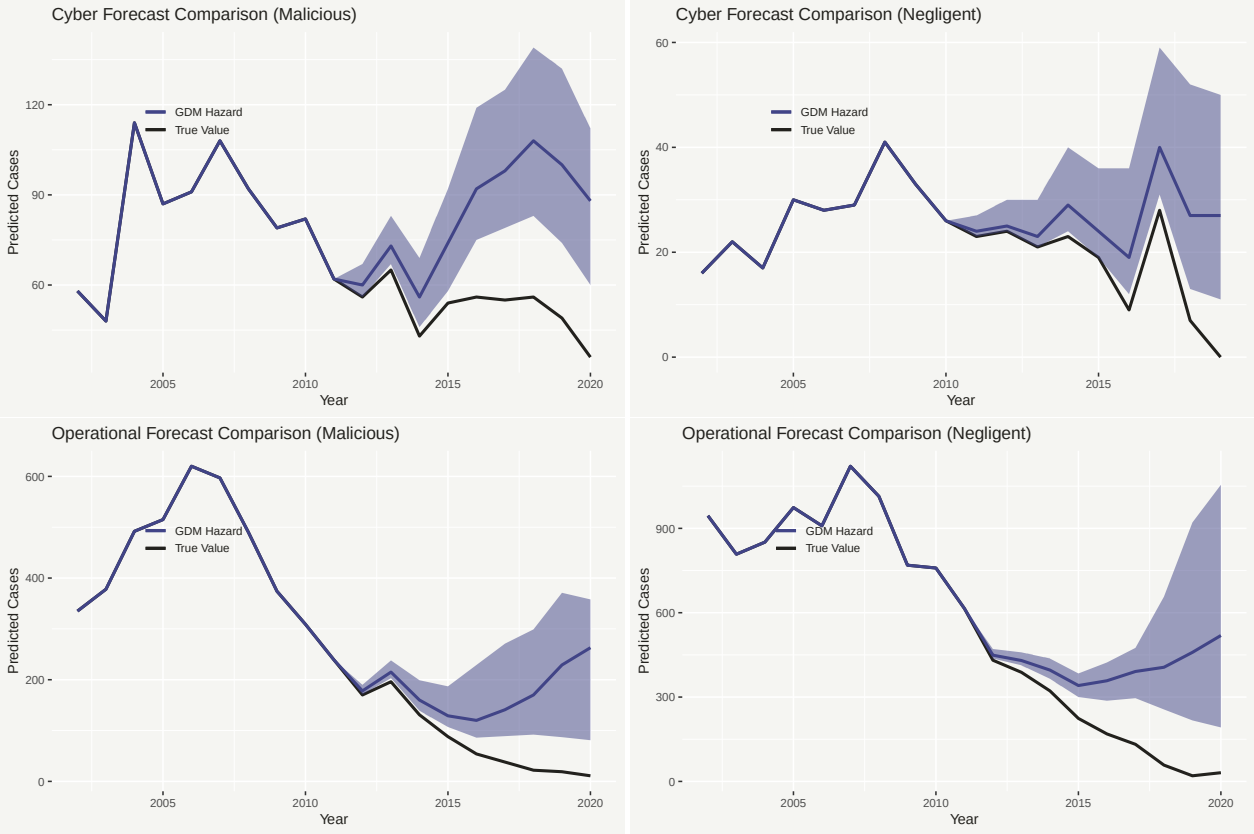

Figure 5. Bias correction for the SAS data

Notes: This figure shows the forecast results of cyber incidents with the 95% confidence interval after adjusting the report delay problem for different categories in the SAS data.

Overall, the results from SAS suggest the increasing trend we observe in Advisen data is not unique and data specific. Since the SAS data only include large events with loss amounts higher than $100,000, this also shows the increasing trend is not solely driven by a large number of incidents with small losses.

B. Time dynamics of loss frequency

To better understand the dynamics of loss frequency, we apply the narrowest-over-threshold method to the bias-adjusted time series data of cyber risk in Advisen.

The top left graph in Figure 6 plots the dynamics of cyber incidents of the malicious type with change points as the grey vertical lines. Four change points are detected, the first one is in November 2015, and after this point, the increasing rate became higher. At the second change point in March 2018, the increasing trend was replaced by a downward pattern. Starting in June 2019, the number of cyber incidents began to increase rapidly and the last change point in July 2020 led to an even higher increase rate.

Given the fact that we are working with time series data, serial dependence can be a problem

of concern as it may induce bias for the estimation of the variance of noise and thus the accuracy of the change point detection. Therefore, following Baranowski et al. (2019), we add additional IID Gaussian noise to the original data with a mean of 0. The standard deviation is chosen to be the standard deviation of the residuals after fitting the original data. The top right graph of Figure 6 plots the result after adjusting serial dependence and we can find the overall pattern is consistent with the original results.

However, the current results are based on the assumption of a linear pattern, but the fact that there are multiple phases with increasing slopes in the previous results indicates that the trend of cyber risk may follow a non-linear pattern. As the method is not designed for non-linear change points, we transform the original data into the log scale and then present the results of change points in the bottom left graph. There is a nearly linear increase in cyber risk in the log scale, which means that cyber risk has undergone exponential growth. The only exception is in the period between March 2018 and December 2019, when there is a short decreasing phase. To control for the potential bias from serial dependence, we add additional IID Gaussian noise with zero mean and the standard deviation of the residuals to the data. As the noise is random, there might be different results depending on the actual distribution of the Gaussian noise. In our paper, we consider the most conservative results with the least change points among different possibilities. This is shown in the bottom right graph of Figure 6. This provides further evidence that the linear pattern assumption is not likely to hold and that cyber risk of the malicious type is in fact undergone exponential growth in the past two decades. There might be certain disturbances during this period such as the period from March 2018 to December 2019 as in the bottom left graph, but these disturbances do not fundamentally change the overall pattern of malicious cyber incidents.

For other types of cyber risk in Advisen, we present the results in Figure 7 with log transformation and serial dependence adjustment as above. The top left graph is the result for the malicious type. The top right graph presents the possible change points for the negligent type. After trying different kinds of Gaussian noise to the data, the decreasing phase between July 2014 and November 2016 is still present. Therefore, for the negligent type, the exponential growth is not continuous but subject to certain disruptions.

For the privacy type, the time pattern is significantly different than other categories. The number of cases began to increase rapidly after April 2015 and peaked around the beginning of 2017. Then there is a significant drop in the number of cases until now. In the data, the incidents are mainly related to the violation of two acts (Telephone Consumer Protection Act and Fair Debt Collection Practices Act) by contacting the consumers without permission. The declining trend after 2017 might be driven by the possibility that more and more firms learn to comply with the acts after a significant number of lawsuits is filed by the consumers. Furthermore, the European Union adopted the general data protection regulation (GDPR) in April 2016, which enhances individuals' control and rights over their personal data. This also might contribute to the decline of cases related to the violation of privacy in our data.

For the "others" category, there is also a clear exponential growth over time but the growth

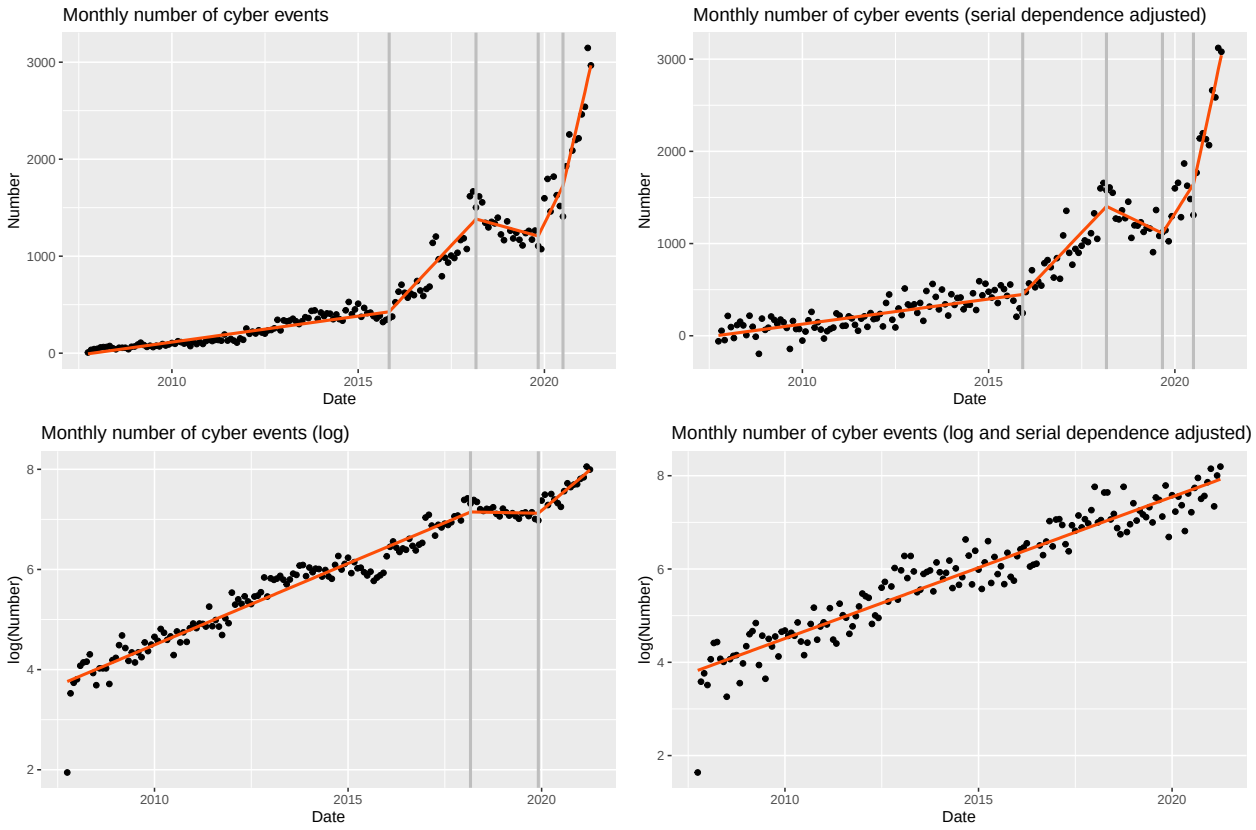

Figure 6. Change points for loss frequency (Malicious)

Notes: This figure reports the results of the change point detection method for malicious cyber incidents in the Advisen data. The grey vertical lines are the dates of change points and the black dots are the monthly number of malicious cyber incidents based on the forecast estimation after correcting the report delay problem.

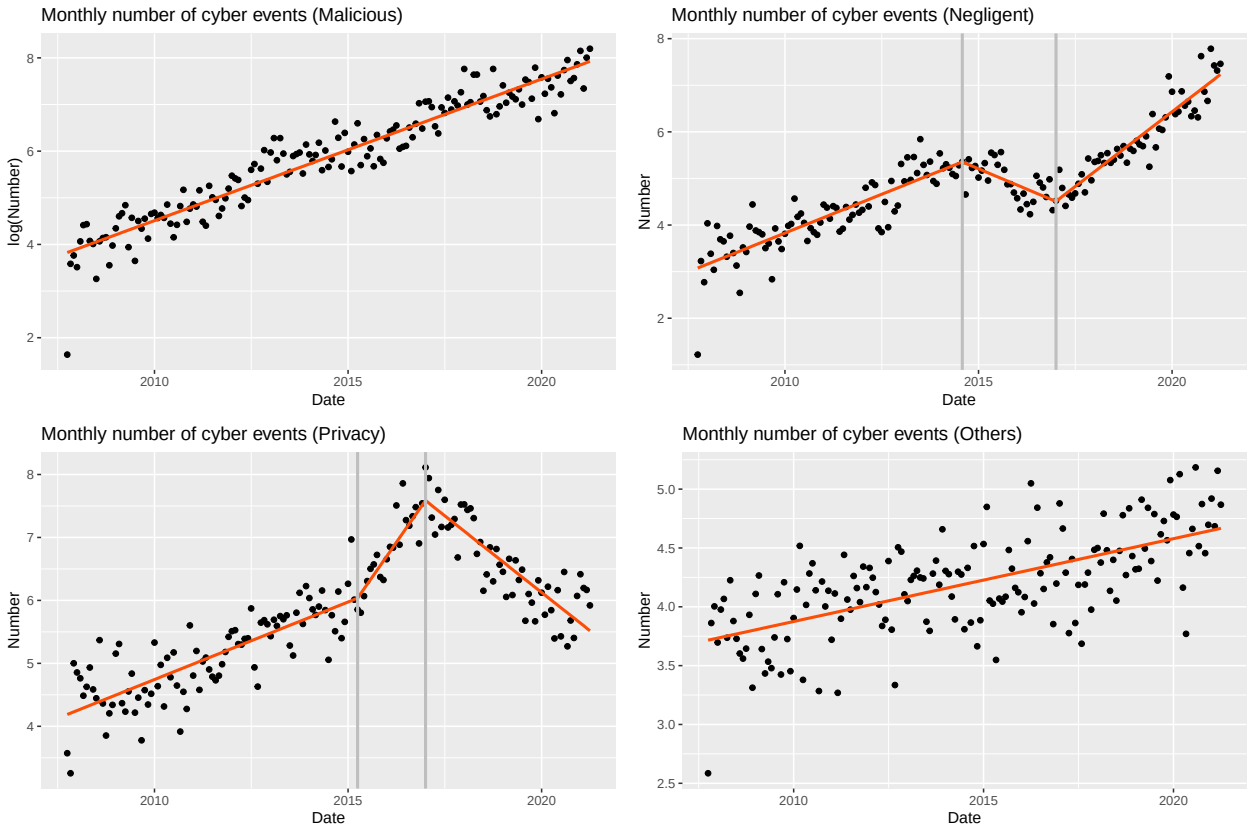

Figure 7. Change points for loss frequency by type

Notes: This figure reports the results of the change point detection method for different kinds of data after log transformation and serial dependence adjustment, based on the forecast estimation when correcting the report delay problem.

rate is much lower than in the case of the malicious and negligent type. Therefore, we will not focus on this category in the following analysis. The comparison of different types of cyber risk indicates that the malicious type is the biggest threat as the number of malicious incidents is growing exponentially and has no sign of slowing down in the observed period.

There are also several other papers looking at the time dynamics of cyber risk, although from different perspectives with different data sources. Jamilov et al. (2021) collect a complete set of transcripts from quarterly earnings conference calls of public firms from 85 countries over the 2002-2020 period and construct a cyber risk exposure measure for each quarter, as shown in the top right graph in Figure 8. The time pattern of their results is very much similar to our bias-corrected pattern in the top left graph for the malicious type. Jamilov et al. (2021) also highlights some notable events related to cyber risk in the figure. In addition, Florakis et al. (2022) builds a cyber risk exposure measure based on the "Risk Factor" section of the SEC filings and presents the yearly average of this measure from 2011 to 2018 (bottom left graph). Although they have less granular results, the increasing pattern is basically the same as what we show.

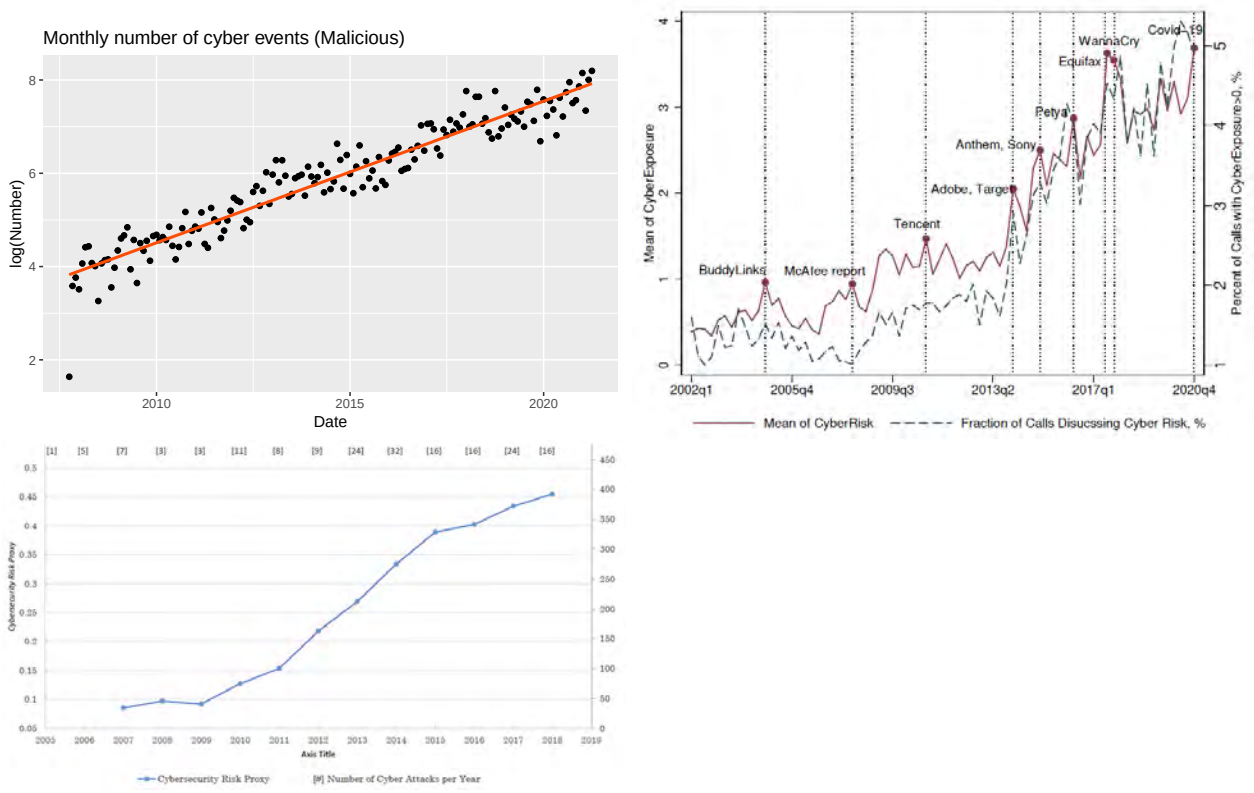

Figure 8. Cross comparison of multiple sources

Notes: This figure compares the time trend of cyber risk frequency from three different sources. The top left graph is from this paper; the top right graph is from Jamilov et al. (2021), based on the cyber risk measure from quarterly earnings conference calls of public firms from 85 countries over the 2002-2020 period; the bottom left graph shows the annual average cyber risk measure based on the "Risk Factor" section of the SEC from 2011 to 2018 (Florakis et al. 2022).

C. Time dynamics of loss severity

C.1. Possible selection bias

Although we address one type of data bias (report delay) in the previous section, there are still other types of bias that might be present in the data. One kind of bias that is of special concern when we study the distribution of cyber risk, is selection bias. This is a common issue in many fields and in our case, it might be that the cyber incidents of certain characteristics are selected for our sample. For example, Amir et al. (2018) indicates that public firms may choose not to disclose certain incidents with small losses and withhold information on more severe cases. Therefore, our sample can be biased in the direction that only the incidents with large losses are included, which will lead to an overestimation of the damage caused by cyber risk.

To test whether our data have this kind of selection bias, we make use of the introduction of data breach notification laws across the U.S. These laws require institutions to notify their customers and other relevant parties in the case of a data breach or unauthorized access to data within a reasonable amount of time. This also means the information about the incident will be reported to the authorities at the same time. Therefore, if there is selection bias in our data, we should observe a significant change in the profile of cyber incidents in the short period before and after the introduction of the notification law. Furthermore, we restrict our sample to cases involving personal identity, financial, health, or record information in the U.S. as only these are required to disclose in a timely manner.

In the U.S., the introduction of data breach notification law is different across states and thus we collect the information on the effective date of this law for each state (PerkinsCoie 2021). For the characteristics of cyber risk, we have information on the total financial losses and affected counts of the incident (as we show above, this information is not frequently populated, especially for the financial losses). If these two variables are smaller after the law, this indicates that incidents with large losses are more likely to be included without regulation. In addition, we have information on the victim firms, such as the number of employees and total revenues. If these two variables are higher before the implementation of the notification law, this provides support that there is selection bias and large firms are likely to be included without regulation.

We utilize three methods to detect possible changes before and after the notification law. The first one is the Kolmogorov–Smirnov test, which is a commonly used method for the equality of continuous and one-dimensional distributions. The second one is the Student's t-test, which is also a standard tool to compare two samples. We conduct the tests with equal variance and unequal variance assumptions. The last one is from Ibragimov & Müller (2016), the one introduced in Section IV.D. This method is specially adjusted for the small sample problem, which is relevant for variables such as financial losses in our data. We choose a half-year period before and after the introduction of the law for each state (the states with at least 10 incidents for each period).[12]

Figure 9 plots the distribution of p values of the tests for different states. The grey horizontal

[12]We also extend the period to one year before and after the law, and the result is consistent with the main result.

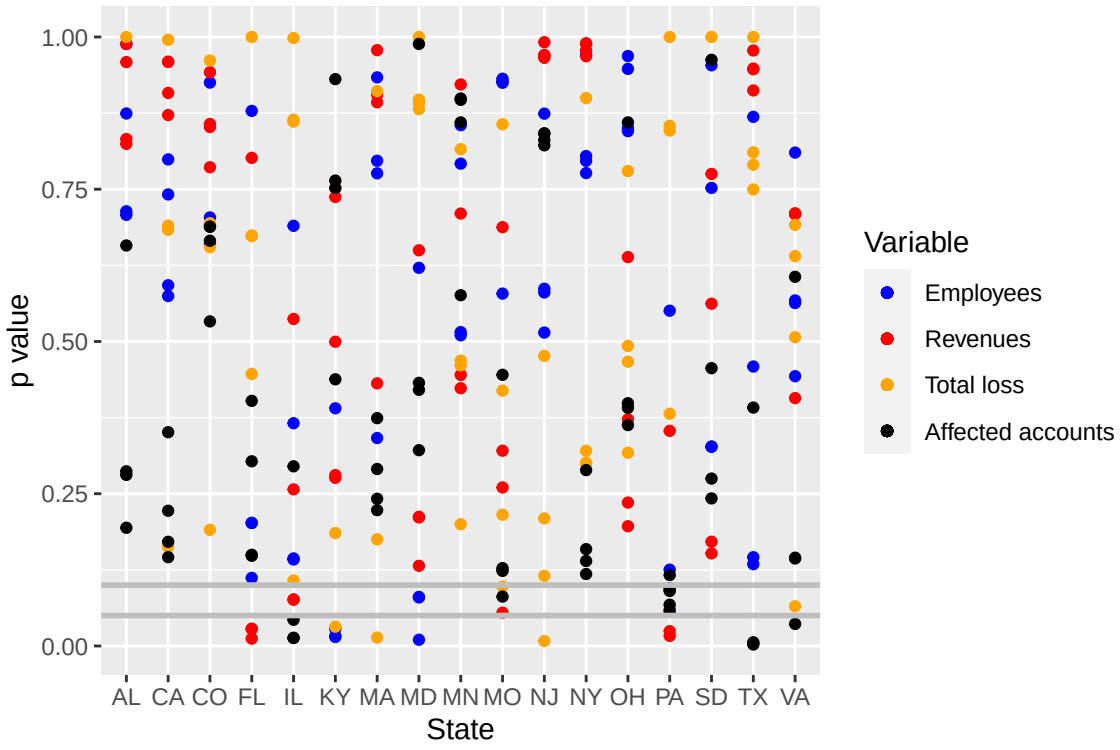

Figure 9. Distribution of p-value for different variables

Notes: This figure reports the distribution of p values of three tests (Kolmogorov–Smirnov test, Student's t-test with equal variance and unequal variance, and small sample adjusted t-test from Ibragimov & Müller (2016). In the figure, we only distinguish the p values for different variables and different states, but not for different tests. The abbreviation of states used in this figure is standard two-letter abbreviations in the U.S, see also https://www.ssa.gov/international/coc-docs/states.html.)

lines represent the 5% and 10% significance levels. The results for different variables and different states are not significant in most of the tests. Although there are certain exceptions, there is no clear pattern among them and the evidence is more consistent with the argument that there is no significant selection bias in our data. Despite the fact that the test is simple and only applies to incidents involving personal data in the U.S., this is still valuable as it reduces the concern of selection bias to a certain extent.

C.2. Change point detection

Figure 10 shows the dynamics of financial loss distributions (log-transformed) for malicious and negligent cases in Advisen and SAS. There are gaps among distributions in the top right graph as there are not enough data points for the density plot. We can find a common pattern in these two databases: the distribution for the malicious cases is shifting to the right, while the distribution for the negligent cases is shifting to the left. In Figure 11, the distribution of the number of records over the years is presented. The malicious cases appear to lead to more losses compared to the negligent cases, and the right tail of malicious cases is fatter than the tail of negligent cases. Still,

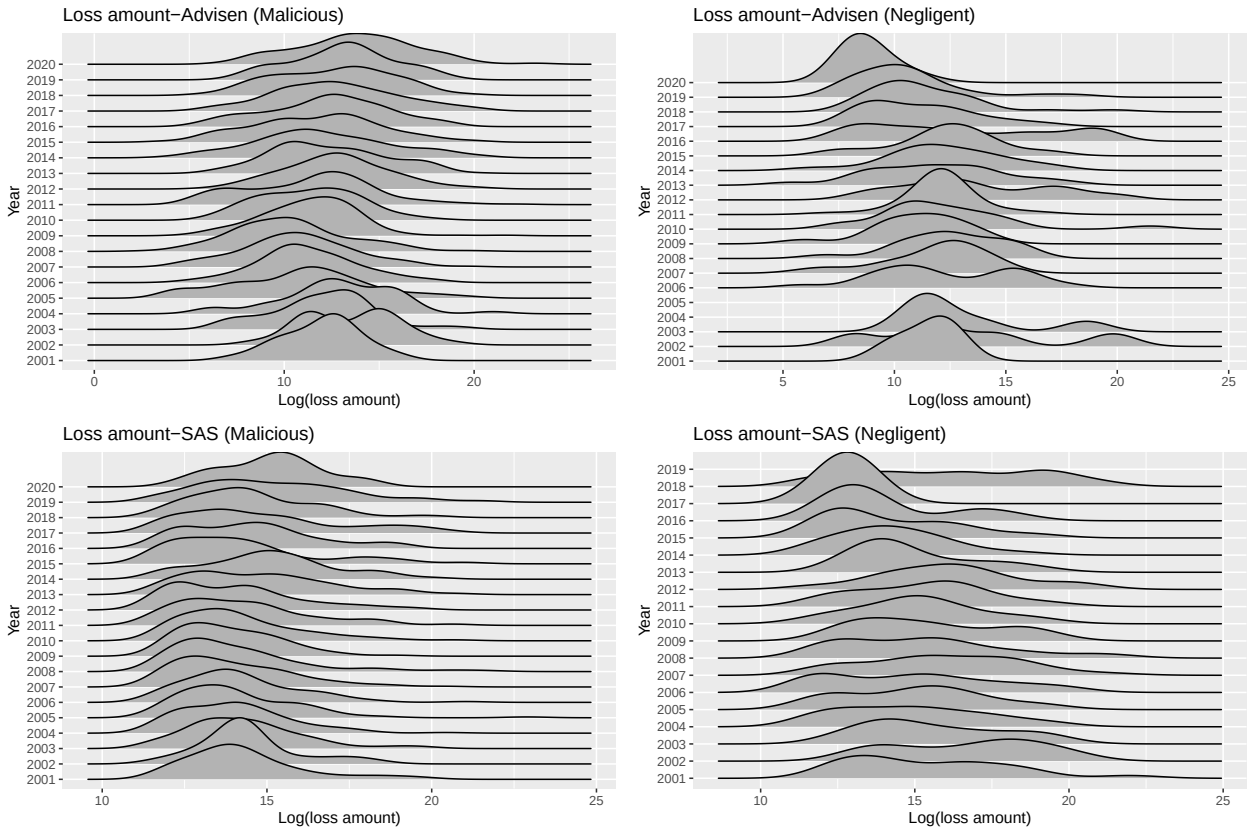

Figure 10. Dynamics of distributions

Notes: This figure presents the dynamics of distributions for financial loss (log scale) in Advisen and SAS. The left panel shows the results for malicious cases, and the right panel shows the results for negligent cases. There are several years with no plotted distribution due to limited data points.

this is only descriptive evidence and we will use the method from Dubey & Müller (2020) to detect change points across distributions.

When we apply the method to the distributions of cyber risk, the significance level of the change point differs when using two ways of calculating the critical value, as discussed in Section IV.C. In fact, the change points detected in most of our cases are only significant when using the asymptotic p-value but not the bootstrap p-value. As the bootstrap p-value is more conservative and indicates larger differences compared to the asymptotic p-value, the change points we discuss later are only present under weaker conditions. In our application, the change point frequently occurs at the beginning or the end of the period. To make sure there is no additional change point for the rest of the period, we apply the method for a second time. As our sample has at most a size of 20 distributions, it is sufficient to apply the method twice because the sample size will decrease significantly and the validity of the change point no longer holds. Therefore, this is a small extension from Dubey & Müller (2020) and it is not especially necessary to provide further analysis of the validity of the method.

Figure 12 compares the average distribution of financial loss before and after the change point.

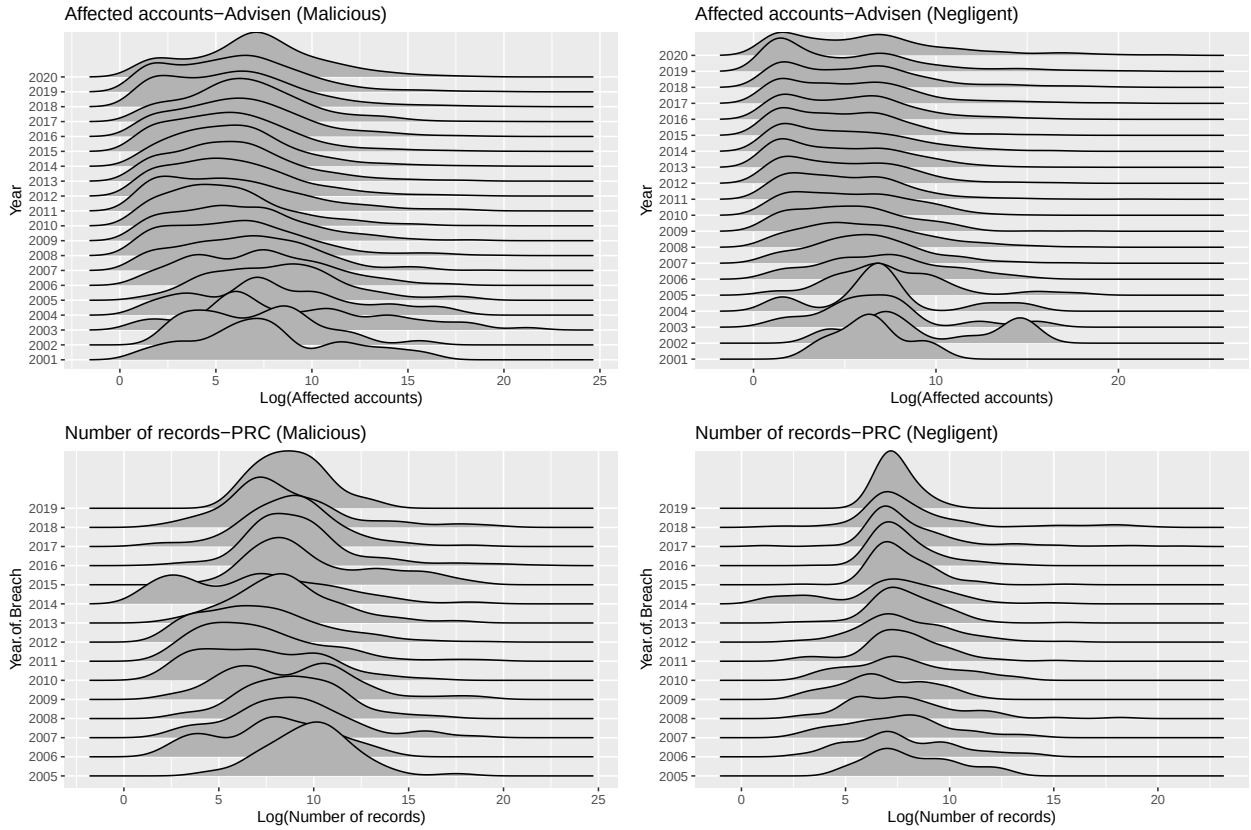

Figure 11. Dynamics of distributions

Notes: This figure presents the dynamics of distributions for the loss of personal records (log scale) in Advisen and PRC. The left panel shows the results for malicious cases, and the right panel shows the results for negligent cases.

The left panel shows the distributions for malicious cyber risk but the results are mixed. In Advisen, we find the distribution first shifts to the left after 2003 and moves to the right after 2018. This indicates the financial losses from malicious cases are getting more severe in the recent period. However, the results from SAS show there is a shift to the left after 2003 but no change point is detected afterward. Therefore, there is no strong evidence indicating the loss distribution of malicious cyber risk is changing in a worrisome direction. In contrast, the loss distribution for negligent cases in the right panel shows a consistent pattern that the distribution is shifting to the left after 2018. In general, the results in Figure 12 show the financial loss distribution from malicious cases has not changed significantly after 2003 but the distribution for negligent cases has undergone a shift in the direction of lower severity.

Figure 13 presents the results for the distributions of the number of records or affected counts. The general impression is that for both malicious and negligent cases the distributions before and after the change point are not as different as the case of financial loss distributions. Still, for the malicious case, we can find that the distribution is shifting to the left during the first decade of this century. The possible reason is that with the development of IT and related technology, all firms, not only the large ones, are exposed to cyber risk. Therefore, the losses come from both the large and small firms, but the small firms usually have less number of records or consumer information. Therefore, the overall loss profile shifts to the left in the second decade. More generally, we can identify a fatter right tail in the recent period for both malicious and negligent cases, which might indicate more and more extremely damaging cases are happening in the recent period. To further understand the tail risk from malicious and negligent cyber incidents, we will provide a more detailed analysis in the following section.

D. Time dynamics of tail risk

D.1. Basics of tail index

We first provide a detailed comparison of the tail index in Table II using two methods ("dAMSE" and "hall") mentioned in Section IV.D with two tail estimators (Hill's and OLS estimator). Although these two methods are chosen using small sample simulations, they also provide good performance for a large sample. Consistent with the simulation results, the tail estimation with the OLS estimator provides higher values, as a result of correcting the downward bias. Therefore, we focus on the estimation results with the OLS estimator afterward. For most of the cases, the results using two threshold selection methods ("dAMSE" and "hall") are reasonably close, but there is a significant difference for the tail index with the Advisen data with respect to the affected counts. This is driven by a large discrepancy between the optimal thresholds selected by these two methods. "hall" chooses a higher threshold for the malicious case compared to "dAMSE", but an extremely lower threshold for the negligent case compared to "dAMSE". To address this issue, we consider the next two best methods in the pool of threshold selection methods, "eye" and "mindist". For the malicious case, the tail index is 0.84 and 1.29, respectively. Therefore, the estimation from "hall" is more consistent with the rest. For the negligent case, the tail index is 0.62 and 0.88 using

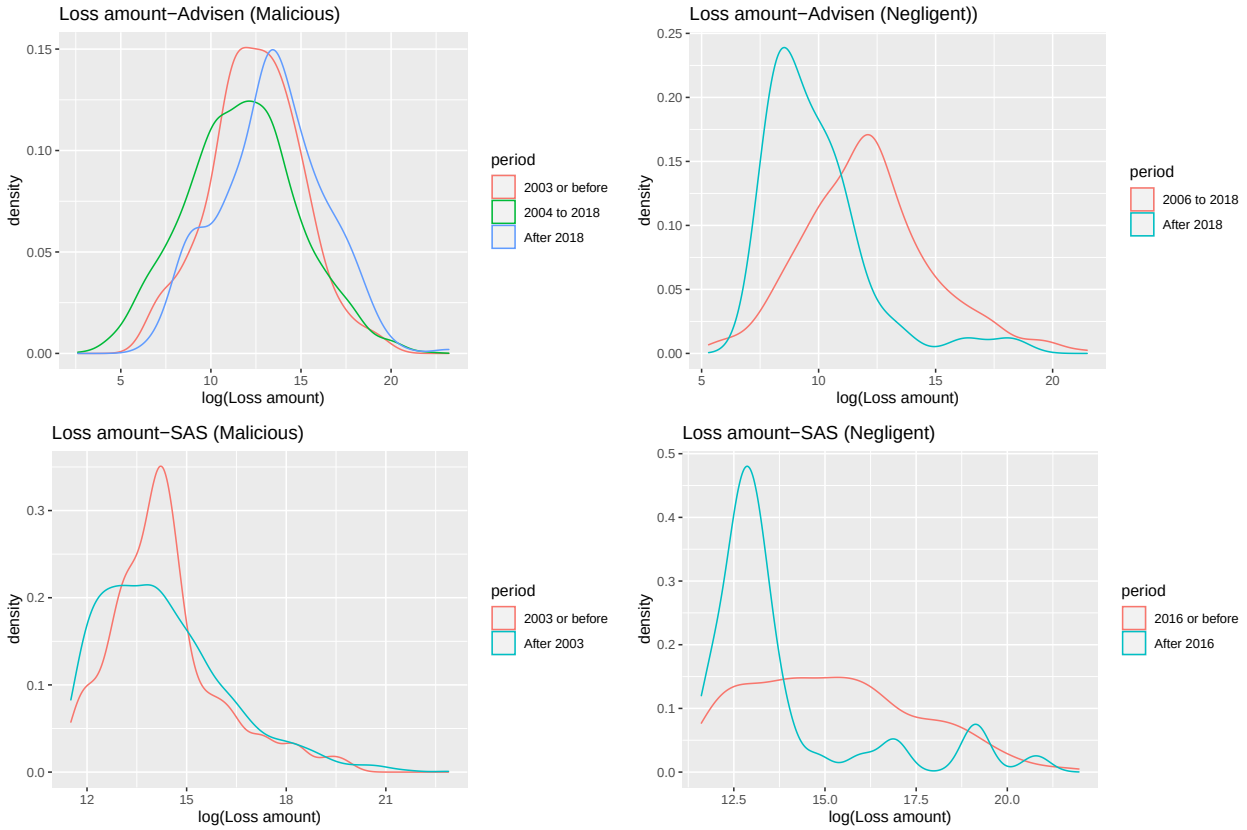

Figure 12. Change points of distributions (loss amount)

Notes: This figure presents the comparison of distributions (log scale) for financial loss in Advisen and SAS. The left panel shows the results for malicious cases, and the right panel shows the results for negligent cases.

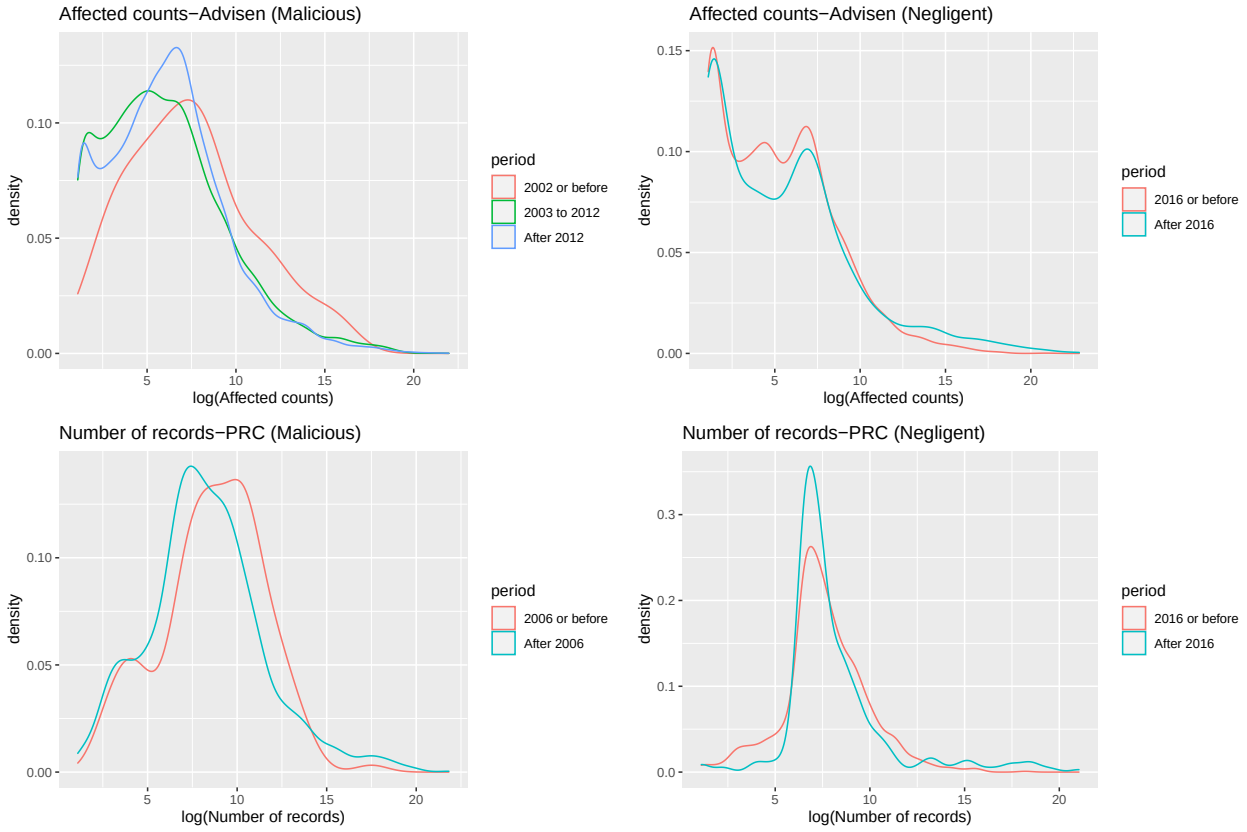

Figure 13. Change points of distributions (number of records)

Notes: This figure presents the comparison of distributions for the loss of personal records (log scale) in Advisen and PRC. The left panel shows the results for malicious cases, and the right panel shows the results for negligent cases

"eye" and "mindist". Hence, the result from "dAMSE" is more reliable. In the following analysis, we will mainly focus on the results from "hall" with the OLS estimator (except the negligent case in Advisen (affected counts) where "dAMSE" with the OLS estimator is used).

The estimation in Table II shows that different kinds of cyber risk all yield severe heavy-tailedness, with a tail index lower than 1 for most of the cases. This is a serious concern as there is no finite mean for cyber loss distributions. Furthermore, the monetary losses have a less heavy tail compared to the non-monetary losses such as the number of records breached. This is reasonable in the sense that the capacity for data storage of any device is increasing rapidly in recent periods and thus the exposure to cyber risk. Therefore, the probability of extreme observations is increasing and this can lead to higher tail risk.

D.2. Change point detection

As the tail of cyber risk exhibits extreme heavy-tailedness, it is of special interest to understand whether this is a stable or dynamic feature. Figure 14 plots the rolling window estimation of the tail index for financial loss of cyber risk and shows the most probable change point in the time period. For malicious cases in Advisen and SAS, the tail index becomes slightly higher and more volatile after the change point, which means that the tail for cyber financial losses is getting less heavy. But for the negligent cases, the general trend is relatively stable with a tendency to decline.

Figure 15 presents the results for the cyber loss as measured by the number of records or accounts affected. For the malicious cases, there is mixed evidence as the trend from Advisen and PRC is different, but the tail index remains below 1 over time. For the negligent cases, there is a clear and consistent declining trend after 2016. This leads to an even heavier tail for negligent cyber risk. In fact, there are more extreme cyber events caused by negligence rather than by malicious third parties in our data. The reason is likely to be that the negligent behaviors are unexpected and arise from the internal process that affects a wide range of data in the system, while the malicious attacks typically have a pre-defined target and focus on a specific set of data rather than all the data in a firm. Although there is no strong academic evidence, this is consistent with various anecdotal evidence such as Shred-it (2018) and CybelAngel (2020).

Overall, the heaviness of the tail for malicious cases is relatively stable with a tendency to decline, while the heaviness of the tail for negligent cases is increasing, especially when the loss is measured by the number of records or accounts affected.

VI. Implications for cyber risk management

A. Optimal risk management with delayed information

One distinct feature of cyber risk is its dynamic nature, and this imposes a serious challenge for the management of cyber risk. We document exponential growth for malicious cyber risk after adjusting report delay. This is significantly different from the trend shown in the raw data. Therefore, depending on the ability to collect and analyze data, it can be very different with respect

Table II Comparison of tail index

Sample	Number after truncation	Hill's estimator			OLS estimator		
		Tail index	95% CI (lower)	95% CI (higher)	Tail index	95% CI (lower)	95% CI (higher)
Advisen (loss amount)–Malicious							
hall	87	0.84	0.67	1.02	0.90	0.63	1.17
dAMSE	156	0.76	0.64	0.88	0.85	0.66	1.04
Advisen (loss amount)–Negligent							
hall	34	0.55	0.37	0.74	0.64	0.33	0.94
dAMSE	61	0.45	0.33	0.56	0.54	0.35	0.73
SAS (loss amount)–Malicious							
hall	67	0.86	0.66	1.07	0.97	0.64	1.29
dAMSE	136	0.70	0.58	0.82	0.84	0.64	1.04
SAS (loss amount)–Negligent							
hall	43	1.12	0.79	1.46	1.15	0.66	1.64
dAMSE	27	1.06	0.66	1.46	1.18	0.55	1.80
SAS Operational risk (loss amount)–Malicious							
hall	127	0.98	0.81	1.15	1.05	0.79	1.30
dAMSE	192	0.89	0.77	1.02	1.00	0.80	1.20
SAS Operational risk (loss amount)–Negligent							
hall	353	1.13	1.01	1.25	1.11	0.95	1.28
dAMSE	1179	0.85	0.80	0.90	0.98	0.90	1.05
Advisen (affected counts)–Malicious							
hall	332	0.64	0.57	0.71	0.82	0.70	0.95
dAMSE	687	0.52	0.48	0.56	0.64	0.57	0.71
Advisen (affected counts)–Negligent							
hall	2055	0.39	0.37	0.40	0.38	0.36	0.41
dAMSE	27	0.68	0.42	0.93	0.74	0.35	1.14
PRC (number of records)–Malicious							
hall	572	0.43	0.39	0.46	0.43	0.38	0.48
dAMSE	166	0.43	0.36	0.49	0.48	0.38	0.59
PRC (number of records)–Negligent							
hall	1218	0.50	0.47	0.53	0.50	0.46	0.54
dAMSE	172	0.48	0.41	0.56	0.44	0.35	0.53

Note:
The truncation is made for the right tail based on the threshold selected by "hall" and "dAMSE". The number after truncation indicates the sample size of observations used for the estimation of the tail index.

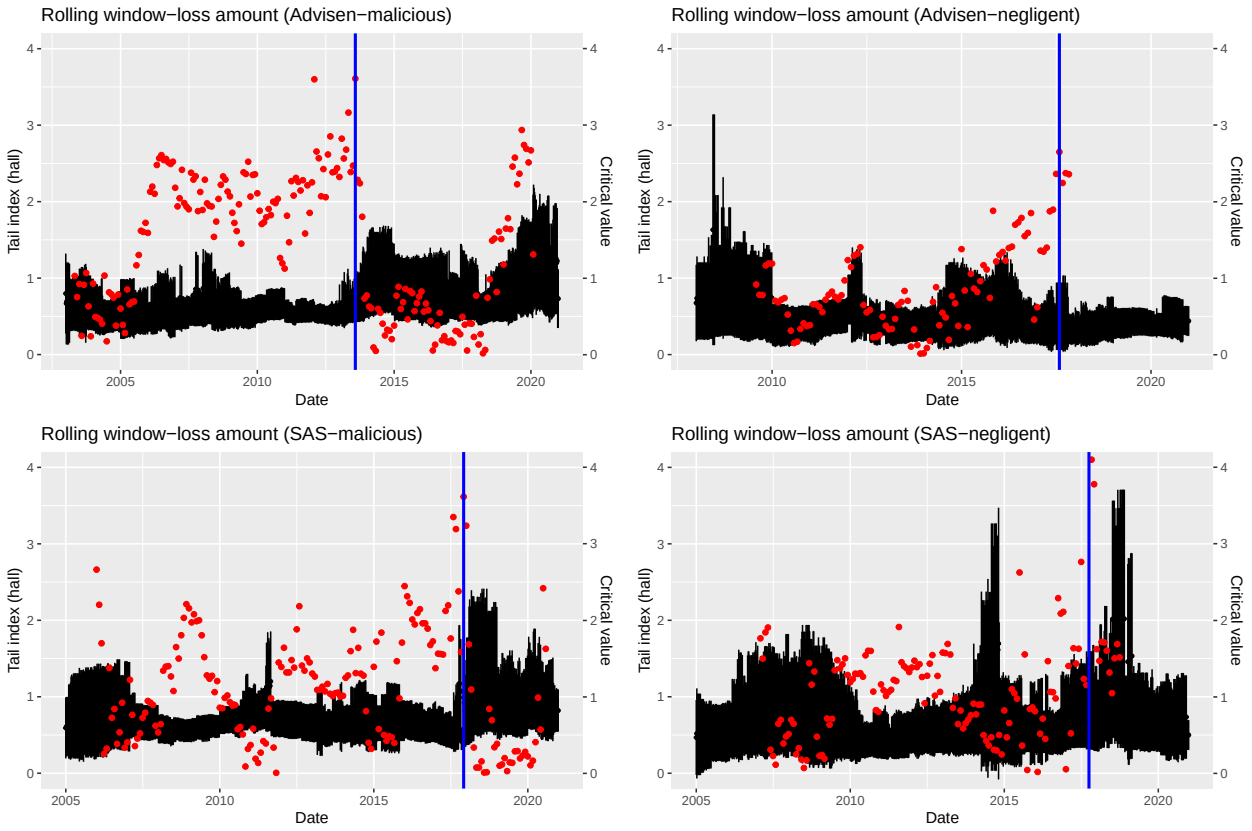

Figure 14. Change points for tail index (loss amount)

Notes: This figure presents the possible change point together with the rolling window estimation of the tail index for financial loss data in Advisen and SAS. The black area in each graph is the estimated tail index with a two-year rolling window and the "hall" method for threshold selection. The red dots are the BF values from Ibragimov & Müller (2016). The blue line is the point with the highest BF value, which indicates the most possible change point in the whole period.

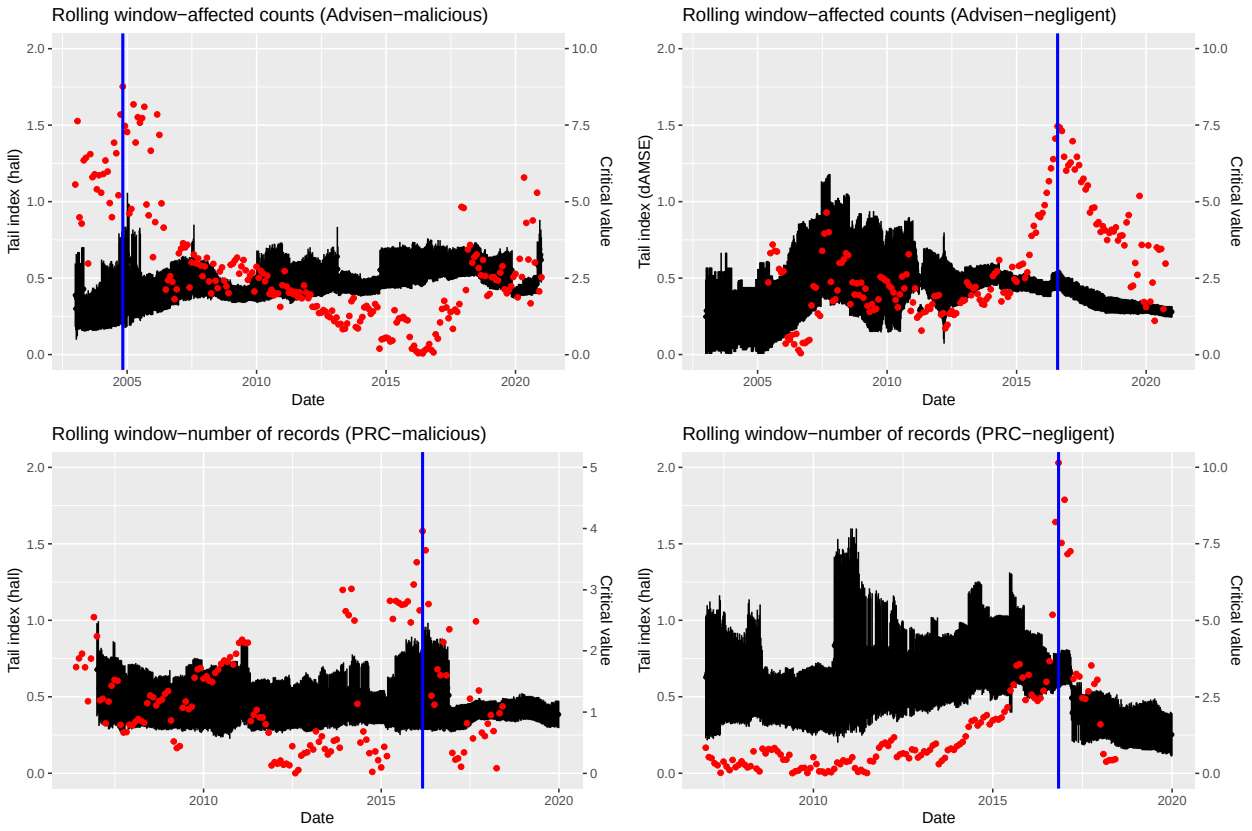

Figure 15. Change points for tail index (number of records)

Notes: This figure presents the possible change point together with the rolling window estimation of the tail index for the number of records affected in Advisen and PRC. The black area in each graph is the estimated tail index with a two-year rolling window and the "hall" method for threshold selection (except for the negligent cases in Advisen, where we rely on the "Advisen" method). The red dots are the BF values from Ibragimov & Müller (2016). The blue line is the point with the highest BF value, which indicates the most possible change point in the whole period.

to the perception and estimation of cyber risk of different parties. In a simple setting of a firm and an insurer, the insurer is more likely to have an information advantage since it specializes in the area of risk management, and dealing with report delay is a standard task as it needs to calculate the claim reserves constantly. The decision-makers in firms may have more information about their own vulnerability related to cyber risk, but may not be experts in analyzing the dynamics of cyber risk. Therefore, there is an information gap between the firm and the insurer, which leads to a difference in the estimation of cyber risk probability. To understand how this difference may affect the optimal decision of cyber risk management, we adopt the widely used framework from Ehrlich & Becker (1972) and discuss the implications in this section.

We consider the situation where the firm has delayed information about the level of cyber risk and thus has a downward biased belief about its risk probability. Consider the firm with initial wealth w_0 subject to a loss of L ($0 < L < w_0$). The probability of loss L is believed to be $p_0 < 1$, while the true probability is $p_1 > p_0$. The firm has a standard utility function u ($u' > 0$ and $u'' < 0$). The concavity of the function comes from the imperfect capital market and tax reasons (Greenwald & Stiglitz 1990; Doherty & Smetters 2005). The options for cyber risk management for the firm include market insurance and self-protection. The price of insurance $\pi(p) = p$ is actuarially fair and thus the premium with the coverage level α is $\alpha\pi(p)L$. The firm could reduce the probability of cyber risk by self-protection measures such as improving the defense system with a cost $x > 0$. The relationship between the risk probability and the self-protection cost is a convex function, with $p'(x) < 0$, $p''(x) < 0$ and $\lim_{x\to\infty} p(x) = 0$.

In case of delayed information, the firm will maximize its expected utility based on the biased probability, facing the actuarially fair insurance premium from the insurer. Therefore, the maximization problem is

$$EU = (1 - p(p_0, x))u(w_0 - x - \alpha\pi(p(p_1, x))L) + p(p_0, x)u(w_0 - x - L + \alpha L - \alpha\pi(p(p_1, x))L), \tag{1}$$

where $p(p_0, x) < p(p_1, x)$.

The first-order condition for a maximum with respect to the coverage α is

$$\frac{\partial EU}{\partial \alpha} = (1 - p(p_0, x))u_1'(\cdot)(-\pi(p(p_1, x))L) + p(p_0, x)u_2'(\cdot)(1 - \pi(p(p_1, x)))L = 0, \tag{2}$$

where $u_1'(\cdot) = u'(w_0 - x - \alpha\pi(p(p_1, x))L)$ and $u_2'(\cdot) = u'(w_0 - x - L + \alpha L - \alpha\pi(p(p_1, x))L)$.

After rearranging the equation, we have

$$\frac{u_1'(\cdot)}{u_2'(\cdot)} = \frac{p(p_0, x)}{1 - p(p_0, x)} \cdot \frac{1 - p(p_1, x)}{p(p_1, x)}. \tag{3}$$

In the standard situation where $p(p_0, x) = p(p_1, x)$, the right-hand side equals 1 and α should satisfy $u_1'(\cdot) = u_2'(\cdot)$. This means $\alpha = 1$, as shown by Mossin (1968). In our case, as $p(p_0, x) < p(p_1, x)$, the right-hand side is smaller than 1, α should be such that $u_1'(\cdot) < u_2'(\cdot)$. Since $u' > 0$

and $u'' < 0$, α has to be smaller than 1. Therefore, the firm would purchase less insurance with the presence of delayed information.

Next, the first-order condition (FOC) with respect to the self-protection x is

$$\frac{\partial EU}{\partial x} = -p'(p_0,x)u_1(\cdot) + (1-p(p_0,x))u_1'(\cdot)(-1-\alpha p'(p_1,x)L) + p'(p_0,x)u_2(\cdot) + p(p_0,x)u_2'(-1-\alpha p'(p_1,x)L) = 0, \quad (4)$$

where $u_1'(\cdot) = u'(w_0 - x - \alpha\pi(p(p_1,x))L)$ and $u_2'(\cdot) = u'(w_0 - x - L + \alpha L - \alpha\pi(p(p_1,x))L)$. By arranging the equation, we have

$$\frac{1+\alpha p'(p_1,x)L}{u_1-u_2} = \frac{-p'(p_0,x)}{(1-p(p_0,x))u_1'(\cdot)+p(p_0,x)u_2'}. \quad (5)$$

Again, consider the standard situation without delayed information, the equilibrium condition will be $\frac{1+\alpha p'(p_1,x)L}{u_1-u_2} = \frac{-p'(p_1,x)}{(1-p(p_1,x))u_1'(\cdot)+p(p_1,x)u_2'}$. Compared to equation 5, the left-hand side will be equal for both situations with a fixed x. However, the right-hand side in equation 5 would be smaller because $-p'(p_0,x) < -p'(p_1,x)$ due to the convexity of the function and $(1-p(p_0,x))u_1'(\cdot)+p(p_0,x)u_2' > (1-p(p_1,x))u_1'(\cdot)+p(p_1,x)u_2'$. To ensure the equality of FOC, x needs to be reduced. Therefore, optimal self-protection will be lower in the case of delayed information.

Combining the results from optimal coverage and self-protection, we find that if the firm has a downward biased belief about its cyber risk, this will lead to underinvestment in cyber risk management, including lower insurance and lower self-protection (thus higher risk).

B. Optimal risk management with delayed information and tail risk

In addition to the problem of delayed information, we also provide evidence for the extreme heavy-tailedness of cyber risk. As mentioned earlier, Ibragimov et al. (2009) has shown this feature might induce the non-diversification trap, resulting in no market for cyber risk in the special situation. In practice, the insurance market exists and has been increasing rapidly, but insurers mostly offer contracts with coverages lower than $1 million and avoid providing high coverage which might severely undermine the financial stability of the company in extreme scenarios. Although this strategy can be useful for protecting the insurers from extreme tail risk arising from cyber insurance lines, this level of coverage is not enough to protect businesses with increasingly high exposure to cyber incidents, and thus limits the value of insurance in the management of cyber risk in the whole society.

Building on this evidence, we introduce tail risk to our model by considering a small probability ϵp of bankruptcy for the firm. The maximum insurance coverage is still L since the insurer does not insure the tail risk due to the non-diversification trap. Therefore, in this bankruptcy case, the remaining value of the firm will be zero even with full insurance. Hence, the maximization problem

of the firm is

$$EU = (1 - p(p_0, x))u(w_0 - x - \alpha\pi(p(p_1, x))L) \\ + (1 - \epsilon)p(p_0, x)u(w_0 - x - L + \alpha L - \alpha\pi(p(p_1, x))L) + \epsilon pu(0), \quad (6)$$

where $\epsilon pu(0) = 0$ and $\pi(p(p_1, x)) = p(p_1, x)$.

Therefore, FOC with respect to the optimal coverage α is

$$\frac{\partial EU}{\partial \alpha} = (1 - p(p_0, x))u_1'(\cdot)(-\pi(p(p_1, x))L) + (1 - \epsilon)p(p_0, x)u_2'(\cdot)(1 - \pi(p(p_1, x)))L = 0, \quad (7)$$

where $u_1'(\cdot) = u'(w_0 - x - \alpha\pi(p(p_1, x))L)$ and $u_2'(\cdot) = u'(w_0 - x - L + \alpha L - \alpha\pi(p(p_1, x))L)$.

Rearranging the equation, we have

$$\frac{u_1'(\cdot)}{u_2'(\cdot)} = \frac{(1 - \epsilon)p(p_0, x)}{1 - p(p_0, x)} \cdot \frac{1 - p(p_1, x)}{p(p_1, x)}. \quad (8)$$

Compared to equation 3, the right-hand side is smaller as $(1-\epsilon)p(p_0, x) < p(p_0, x)$. This means that with the presence of tail risk, the optimal insurance coverage will be even lower. The reason is that the value of insurance is reduced as the reimbursement in the extreme scenario would not benefit the firm.

For optimal self-protection, FOC is shown as

$$\frac{\partial EU}{\partial x} = -p'(p_0, x)u_1(\cdot) + (1 - p(p_0, x))u_1'(\cdot)(-1 - \alpha p'(p_1, x)L) \\ + (1 - \epsilon)p'(p_0, x)u_2(\cdot) + (1 - \epsilon)p(p_0, x)u_2'(-1 - \alpha p'(p_1, x)L) = 0, \quad (9)$$

where $u_1'(\cdot) = u'(w_0 - x - \alpha\pi(p(p_1, x))L)$ and $u_2'(\cdot) = u'(w_0 - x - L + \alpha L - \alpha\pi(p(p_1, x))L)$.

Simplifying the equation yields

$$\frac{1 + \alpha p'(p_1, x)L}{u_1 - (1 - \epsilon)u_2} = \frac{-p'(p_0, x)}{(1 - p(p_0, x))u_1'(\cdot) + (1 - \epsilon)p(p_0, x)u_2'}. \quad (10)$$

Compared to equation 5, the right-hand side is larger because $(1 - \epsilon)p(p_0, x) < p(p_0, x)$ and the left-hand side is smaller because $u_1 - (1 - \epsilon)u_2 > u_1 - u_2$. Therefore, the optimal x should be adjusted upward to ensure the equality of FOC. This result is obtained due to the different effects of delayed information and tail risk as delayed information would induce lower self-protection while tail risk will lead to a higher protection level. However, whether the optimal x is higher than the standard situation without delayed information and tail risk is unclear and depends on the severity of delayed information and the exposure of tail risk.

Overall, this basic model illustrates how the empirical properties of cyber risk influence the demand for cyber insurance and the optimal investment in self-protection. In particular, we show how delayed information and tail risk reduce the demand for insurance, which is consistent with

the evidence from Cellerini et al. (2022) that more than 90% of the cyber loss is not covered by insurance. Previous work such as Böhme et al. (2010) has discussed some of the special properties of cyber risk, we provide further empirical evidence on this and connect these features to the standard insurance economics model for understanding their impacts on risk management.

This also has implications for policy. As delayed information is a serious problem for cyber risk, it is possible to address this issue by establishing a public center that serves to provide up-to-date information about cyber risk to the public. For the issue of limited insurance supply due to tail risk, the government can step in and provide support for insurers as the lender of last resort. This would partially alleviate the insurer's concern about tail risk and improve the market capacity for underwriting cyber risk.

VII. Conclusion

To better understand the dynamic nature of cyber risk, this paper exploits three main databases and studies the fundamental empirical properties of this increasingly important type of risk. We first deal with the problem of report delay that is inherent to the datasets used in empirical research. Then we analyze the frequency and severity of cyber risk using state-of-art statistical methods for the detection of structural changes. We show that the threat of cyber risk comes more from the fact that the frequency is increasing rapidly rather than each cyber incident getting more severe as malicious cyber risk is growing exponentially in the past two decades but there is no significant change for the loss distribution. Moreover, we explore the dynamics of tail risk and find that the heavy-tailedness of cyber risk is persistent over time. Based on these results, we incorporate two empirical features (delayed information and tail risk) into basic insurance economics and show that they lead to significantly lower insurance demand.

Our research also highlights various challenges for analyzing and managing cyber risk in the future. First, we show that the cyber risk landscape is constantly evolving, with new threats emerging all the time. It is of paramount importance to diligently monitor these changes to ensure updated and timely responses. Second, researchers face the obstacle of limited data availability for cyber risk, related to the bias problems encountered in our study. This scarcity arises partly due to organizations' reluctance to disclose information about cyber attacks and breaches, and partly due to the challenging detection and quantification of cyber attacks. Third, cyber risk is a complex and multifaceted problem that involves technical, organizational, and human factors. This complexity makes it challenging for researchers to develop comprehensive and effective solutions and thus interdisciplinary work is needed (Falco et al. 2019). However, such collaborations are often limited due to disciplinary boundaries and institutional structures. Fourth, there is a lack of standardization in the way that cyber risk is measured and assessed. This makes it difficult to compare research findings and develop a common understanding of the problem. Addressing these challenges will require greater collaboration among researchers, practitioners, and policymakers, as well as increased investment in research and development.

REFERENCES

Accenture (2021), '2021 cyber threat intelligence report'. https://www.accenture.com/lu-en/insights/security/cyber-threat-intelligence-report-2021 (accessed December 14, 2022).

Allianz (2021), 'Managing the impact of increasing interconnectivity: Trends in cyber risk', *Allianz Global Corporate & Specialty* . https://www.agcs.allianz.com/news-and-insights/reports/cyber-risk-trends-2020.html (accessed December 14, 2022).

Amir, E., Levi, S. & Livne, T. (2018), 'Do firms underreport information on cyber-attacks? evidence from capital markets', *Review of Accounting Studies* **23**(3), 1177–1206.

Anderson, R. & Moore, T. (2006), 'The economics of information security', *science* **314**(5799), 610–613.

Arnold, T. B., Tibshirani, R. J., Arnold, M. T. & ByteCompile, T. (2020), 'Package 'genlasso'', *Statistics* **39**(3), 1335–1371.

Bai, J. & Perron, P. (2003), 'Computation and analysis of multiple structural change models', *Journal of applied econometrics* **18**(1), 1–22.

Baranowski, R., Chen, Y. & Fryzlewicz, P. (2019), 'Narrowest-over-threshold detection of multiple change points and change-point-like features', *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **81**(3), 649–672.

Bessy-Roland, Y., Boumezoued, A. & Hillairet, C. (2021), 'Multivariate hawkes process for cyber insurance', *Annals of Actuarial Science* **15**(1), 14–39.

Biener, C., Eling, M. & Wirfs, J. H. (2015), 'Insurability of cyber risk: An empirical analysis', *The Geneva Papers on Risk and Insurance-Issues and Practice* **40**(1), 131–158.

Böhme, R. & Kataria, G. (2006), Models and measures for correlation in cyber-insurance., *in* 'WEIS', Vol. 2, p. 3.

Böhme, R., Schwartz, G. et al. (2010), Modeling cyber-insurance: towards a unifying framework., *in* 'WEIS'.

Bolot, J. & Lelarge, M. (2009), Cyber insurance as an incentive for internet security, *in* 'Managing information risk and the economics of security', Springer, pp. 269–290.

Caeiro, F. & Gomes, M. I. (2015), 'Threshold selection in extreme value analysis', *Extreme value modeling and risk analysis: Methods and applications* pp. 69–82.

Cebula, J. L. & Young, L. R. (2010), A taxonomy of operational cyber security risks, Technical report, Carnegie-Mellon Univ Pittsburgh Pa Software Engineering Inst.

Cellerini, E., Finucane, J., Lanci, L. & Holzheu, T. (2022), 'Cyber insurance: strengthening resilience for the digital transformation', *Swiss Re Institute* . https://www.swissre.com/institute/research/topics-and-risk-dialogues/digital-business-model-and-cyber-risk/cyber-insurance-strengthening-resilience.html (accessed December 14, 2022).

CybelAngel (2020), 'Negligence drives 90% of data breaches'. https://cybelangel.com/negligent-data-breaches/ (accessed December 14, 2022).

Doherty, N. & Smetters, K. (2005), 'Moral hazard in reinsurance markets', *Journal of Risk and Insurance* **72**(3), 375–391.

Dubey, P. & Müller, H.-G. (2020), 'Fréchet change-point detection', *The Annals of Statistics* **48**(6), 3312–3335.

Edwards, B., Hofmeyr, S. & Forrest, S. (2016), 'Hype and heavy tails: A closer look at data breaches', *Journal of Cybersecurity* **2**(1), 3–14.

Ehrlich, I. & Becker, G. S. (1972), 'Market insurance, self-insurance, and self-protection', *Journal of Political Economy* **80**(4), 623–648.

Eling, M. & Jung, K. (2018), 'Copula approaches for modeling cross-sectional dependence of data breach losses', *Insurance: Mathematics and Economics* **82**, 167–180.

Eling, M. & Loperfido, N. (2017), 'Data breaches: Goodness of fit, pricing, and risk measurement', *Insurance: mathematics and economics* **75**, 126–136.

Eling, M. & Wirfs, J. (2019), 'What are the actual costs of cyber risk events?', *European Journal of Operational Research* **272**(3), 1109–1119.

Falco, G., Eling, M., Jablanski, D., Weber, M., Miller, V., Gordon, L. A., Wang, S. S., Schmit, J., Thomas, R., Elvedi, M. et al. (2019), 'Cyber risk research impeded by disciplinary barriers', *Science* **366**(6469), 1066–1069.

Fang, Z., Xu, M., Xu, S. & Hu, T. (2021), 'A framework for predicting data breach risk: Leveraging dependence to cope with sparsity', *IEEE Transactions on Information Forensics and Security* **16**, 2186–2201.

Farkas, S., Lopez, O. & Thomas, M. (2021), 'Cyber claim analysis using generalized pareto regression trees with applications to insurance', *Insurance: Mathematics and Economics* **98**, 92–105.

FBI (2020), '2020 internet crime report'. https://www.fbi.gov/news/pressrel/press-releases/fbi-releases-the-internet-crime-complaint-center-2020-internet-crime-report-including-covid-19-scam-statistics (accessed December 14, 2022).

Florakis, C., Louca, C., Michaely, R. & Weber, M. (2022), 'Cybersecurity risk', *The Review of Financial Studies* .

Gabaix, X. & Ibragimov, R. (2011), 'Rank- 1/2: a simple way to improve the ols estimation of tail exponents', *Journal of Business & Economic Statistics* **29**(1), 24–39.

Gordon, L. A. & Loeb, M. P. (2002), 'The economics of information security investment', *ACM Transactions on Information and System Security (TISSEC)* **5**(4), 438–457.

Gordon, L. A., Loeb, M. P. & Sohail, T. (2003), 'A framework for using insurance for cyber-risk management', *Communications of the ACM* **46**(3), 81–85.

Greenwald, B. C. & Stiglitz, J. E. (1990), 'Asymmetric information and the new theory of the firm: Financial constraints and risk behavior'.

Hall, P. (1990), 'Using the bootstrap to estimate mean squared error and select smoothing parameter in nonparametric problems', *Journal of multivariate analysis* **32**(2), 177–203.

Hill, B. M. (1975), 'A simple general approach to inference about the tail of a distribution', *The Annals of Statistics* pp. 1163–1174.

Ibragimov, R., Jaffee, D. & Walden, J. (2009), 'Nondiversification traps in catastrophe insurance markets', *The Review of Financial Studies* **22**(3), 959–993.

Ibragimov, R. & Müller, U. K. (2016), 'Inference with few heterogeneous clusters', *Review of Economics and Statistics* **98**(1), 83–96.

Jamilov, R., Rey, H. & Tahoun, A. (2021), The anatomy of cyber risk, Technical report, National Bureau of Economic Research.

Johnson, B., Böhme, R. & Grossklags, J. (2011), Security games with market insurance, *in* 'International Conference on Decision and Game Theory for Security', Springer, pp. 117–130.

Jung, K. (2021), 'Extreme data breach losses: An alternative approach to estimating probable maximum loss for data breach risk', *North American Actuarial Journal* pp. 1–24.

Kamiya, S., Kang, J.-K., Kim, J., Milidonis, A. & Stulz, R. M. (2021), 'Risk management, firm reputation, and the impact of successful cyberattacks on target firms', *Journal of Financial Economics* **139**(3), 719–749.

Kim, S.-J., Koh, K., Boyd, S. & Gorinevsky, D. (2009), '\ell_1 trend filtering', *SIAM review* **51**(2), 339–360.

Laszka, A., Felegyhazi, M. & Buttyan, L. (2014), 'A survey of interdependent information security games', *ACM Computing Surveys (CSUR)* **47**(2), 1–38.

Mack, T. (1993), 'Distribution-free calculation of the standard error of chain ladder reserve estimates', *ASTIN Bulletin: The Journal of the IAA* **23**(2), 213–225.

Maillart, T. & Sornette, D. (2010), 'Heavy-tailed distribution of cyber-risks', *The European Physical Journal B* **75**(3), 357–364.

Marotta, A., Martinelli, F., Nanni, S., Orlando, A. & Yautsiukhin, A. (2017), 'Cyber-insurance survey', *Computer Science Review* **24**, 35–61.

Mossin, J. (1968), 'Aspects of rational insurance purchasing', *The Journal of Political Economy* pp. 553–568.

Naghizadeh, P. & Liu, M. (2014), Voluntary participation in cyber-insurance markets, *in* 'Workshop on the Economics of Information Security (WEIS)'.

Niu, Y. S., Hao, N. & Zhang, H. (2016), 'Multiple change-point detection: a selective overview', *Statistical Science* pp. 611–623.

Ossberger, J. (2020), 'Package 'tea''.

PerkinsCoie (2021), 'Security breach notification chart'. https://www.perkinscoie.com/en/news-insights/security-breach-notification-chart.html (accessed December 14, 2022).

Renshaw, A. E. & Verrall, R. J. (1998), 'A stochastic model underlying the chain-ladder technique', *British Actuarial Journal* **4**(4), 903–923.

Romanosky, S. (2016), 'Examining the costs and causes of cyber incidents', *Journal of Cybersecurity* **2**(2), 121–135.

Salmon, M., Schumacher, D., Stark, K. & Höhle, M. (2015), 'Bayesian outbreak detection in the presence of reporting delays', *Biometrical Journal* **57**(6), 1051–1067.

Shetty, N., Schwartz, G., Felegyhazi, M. & Walrand, J. (2010), Competitive cyber-insurance and internet security, *in* 'Economics of information security and privacy', Springer, pp. 229–247.

Shred-it (2018), '2018 state of the industry report: Information security'. https://www.shredit.com/content/dam/shred-it/global/images/legacy/Shred-it-2018-North-America-State-of-the-Industry.pdf.coredownload.inline.pdf (accessed December 14, 2022).

Smith, Z. & Lostri, E. (2020), 'The hidden costs of cybercrime', *McAfee* . https://www.mcafee.com/enterprise/en-us/assets/reports/rp-hidden-costs-of-cybercrime.pdf (accessed December 14, 2022).

Stempel, J. & Finkle, J. (2017), 'Yahoo says all three billion accounts hacked in 2013 data theft', *Reuters* . https://www.reuters.com/article/us-yahoo-cyber-idUSKCN1C82O1 (accessed December 14, 2022).

Stoner, O. & Economou, T. (2020), 'Multivariate hierarchical frameworks for modeling delayed reporting in count data', *Biometrics* **76**(3), 789–798.

Sun, H., Xu, M. & Zhao, P. (2021), 'Modeling malicious hacking data breach risks', *North American Actuarial Journal* **25**(4), 484–502.

Swartz, J. (2007), 'Tjx data breach may involve 94 million credit cards', *ABC News* . https://abcnews.go.com/Technology/story?id=3773782 (accessed December 14, 2022).

Taylor, G. (2019), 'Loss reserving models: Granular and machine learning forms', *Risks* **7**(3), 82.

Truong, C., Oudre, L. & Vayatis, N. (2020), 'Selective review of offline change point detection methods', *Signal Processing* **167**, 107299.

Wang, G., Gu, Z., Li, X., Yu, S., Kim, M., Wang, Y., Gao, L. & Wang, L. (2021), 'Comparing and integrating us covid-19 data from multiple sources with anomaly detection and repairing', *Journal of Applied Statistics* pp. 1–27.

Wang, J., Chaudhury, A. & Rao, H. R. (2008), 'Research notea value-at-risk approach to information security investment', *Information Systems Research* **19**(1), 106–120.

Wang, Q.-H. & Kim, S. H. (2009*a*), Cyber attacks: Cross-country interdependence and enforcement, WEIS.

Wang, Q.-H. & Kim, S.-H. (2009*b*), Cyber attacks: Does physical boundary matter?, AIS.

Wheatley, S., Hofmann, A. & Sornette, D. (2021), 'Addressing insurance of data breach cyber risks in the catastrophe framework', *The Geneva Papers on Risk and Insurance-Issues and Practice* **46**(1), 53–78.

Wheatley, S., Maillart, T. & Sornette, D. (2016), 'The extreme risk of personal data breaches and the erosion of privacy', *The European Physical Journal B* **89**(1), 1–12.

Woods, D. W., Moore, T. & Simpson, A. C. (2021), 'The county fair cyber loss distribution: Drawing inferences from insurance prices', *Digital Threats: Research and Practice* **2**(2), 1–21.

Zeileis, A., Leisch, F., Hornik, K., Kleiber, C. & Hansen, B. (2022), 'strucchange: Testing, monitoring, and dating structural changes'.

Zhang Wu, M., Luo, J., Fang, X., Xu, M. & Zhao, P. (2021), 'Modeling multivariate cyber risks: deep learning dating extreme value theory', *Journal of Applied Statistics* pp. 1–21.

Zhao, X., Xue, L. & Whinston, A. B. (2013), 'Managing interdependent information security risks: Cyberinsurance, managed security services, and risk pooling arrangements', *Journal of Management Information Systems* **30**(1), 123–152.

Appendices

Appendix A. Literature review

Table A.1 summarizes the empirical work on cyber risk, especially the ones from a statistical or actuarial perspective.

Table A.1 Empirical work on cyber risk

Number	Title	Author (Year)	Dataset	Time period	Study Focus	Main Empirical Result/Implication
1	Models and Measures for Correlation in Cyber-Insurance	Böhme & Kataria (2006)	Honeypot data on attack intensity of network exploits	Feb 2003 to Sep 2005	Correlation of internal and global network structure	The existence of correlation and the result for global correlation is more robust than internal correlation.
2	A Value-at-Risk Approach to Information Security Investment	Wang et al. (2008)	Daily data from a large financial institution	Jan 2004 to Mar 2005	Value-at-risk of daily losses an organization faces	The firms can make a better security investment choice based on their proposed approach.
3	Cyber Attacks: Cross-Country Interdependence and Enforcement	Wang & Kim (2009*a*)	Attack data from Internet Storm Center (ISC)	Jan 2003 to Dec 2007	The impact of the first international treaty against cybercrimes on cyber attacks	The treaty lowers the cyber attacks by around 20% and affects the interdependency across countries.
4	Cyberattacks: Does Physical Boundry Matter?	Wang & Kim (2009*b*)	Attack data from Internet Storm Center (ISC)	Jan 2003 to Dec 2007	Spatial correlation of cyber attacks	Strong evidence of spatial correlation over time.
5	Heavy-tailed Distribution of Cyber-risks	Maillart & Sornette (2010)	Identity Data (ID) loss event data from Open Security Foundation	Jan 2000 to Nov 2008	Heavy-tailedness of ID losses	The presence of a stable heavy-tailed distribution of personal identity losses per event with a strong non-stationary growth of ID losses culminating in July 2006 followed by a more stationary phase.
6	Insurability of Cyber Risk: An Empirical Analysis	Biener et al. (2015)	Cyber losses extracted from SAS OpRisk database	Mar 1971 to Sep 2009	Insurability of cyber risk	The distinct characteristics of cyber risk compared to other operational risk and discuss the main insurability limitations.
7	The Extreme Risk of Personal Data Breaches and the Erosion of Privacy	Wheatley et al. (2016)	ID loss event data from Open Security Foundation and Privacy Rights Clearinghouse	Jan 2007 to Apr 2015	Projection of extreme risk	The maximum breach size is expected to grow by 50% and the total amount is to double in 5 years.

8	Hype and Heavy Tails: A Closer Look at Data Breaches	Edwards et al. (2016)	Privacy Rights Clearinghouse	Jan 2005 to Sep 2015	Trend of data breach	No evidence of the increasing trend for size or frequency of data breaches.
9	Examining the Costs and Causes of Cyber Incidents	Romanosky (2016)	Advisen	Jan 2004 to Dec 2015	Statistical analysis of costs of cyber risk	While there is an increase in the number of events and legal actions, the estimates of firm costs are not of large magnitude.
10	Data Breaches: Goodness of Fit, Pricing and Risk Measurement	Eling & Loperfido (2017)	Privacy Rights Clearinghouse	Jan 2005 to Dec 2015	Model fitting for cyber risk	Log-skew-normal is a good distribution for data breach amount.
11	Copula Approaches for Modeling Cross-sectional Dependence of Data Breach Losses	Eling & Jung (2018)	Privacy Rights Clearinghouse	Jan 2005 to Dec 2016	Cross-sectional dependence of data breach	The presence of a significant asymmetric tail dependence among risk factors.
12	What are the Actual Costs of Cyber Risk Events?	Eling & Wirfs (2019)	Cyber losses extracted from SAS OpRisk database	Jan 1995 to Mar 2014	Model fitting for cyber risk	Extreme value theory is needed for the modeling of severity and cyber risk is inherently dynamic.
13	Addressing Insurance of Data Breach Cyber Risks in the Catastrophe Framework	Wheatley et al. (2021)	ID loss event data from Open Security Foundation and Privacy Rights Clearinghouse	Jan 2005 to Sep 2017	Catastrophic cyber risk and the dynamics	The rate of breaches in excess of 50k ids is constant but an increasing trend for both frequency and severity of hack type events.
14	Multivariate Hawkes Process for Cyber Insurance	Bessy-Roland et al. (2021)	Privacy Rights Clearinghouse	Jan 2010 to Dec 2017	A multivariate Hawkes framework for modelling and predicting attack frequency	The proposed method has good performance.
15	Cyber Claim Analysis Using Generalized Pareto Regression Trees with Applications to Insurance	Farkas et al. (2021)	Privacy Rights Clearinghouse	Jan 2005 to Jan 2019	Analyzing cyber claims with regression trees	Some sectors (such as healthcare, education, and nonprofit organization) have lighter tails than the others, and it is important to separate typical and extreme claims.
16	Modeling Malicious Hacking Data Breach Risks	Sun et al. (2021)	Privacy Rights Clearinghouse	Jan 2005 to Mar 2019	Modeling data breach risk with a frequency-severity framework	The proposed framework captures the nonlinear dependence of data breach risk.

17	Extreme Data Breach Losses: An Alternative Approach to Estimating Probable Maximum Loss for Data Breach Risk	Jung (2021)	Cowbell Cyber	Jan 2005 to Dec 2018	Projection of extreme data breach losses	A significant increase with a break in 2014 for loss severity and substantially larger loss in 5 years compared to the estimate of Pareto model
18	A Framework for Predicting Data Breach Risk: Leveraging Dependence to Cope With Sparsity	Fang et al. (2021)	Privacy Rights Clearinghouse	Jan 2005 to Dec 2018	Predicting the frequency of data breach at enterprise level	Data breach sizes exhibit large variability and large skewness, and consecutive breaches are unlikely to occur to an enterprise within a short period of time.
19	Modeling Multivariate Cyber Risks: Deep Learning Dating Extreme Value Theory	Zhang Wu et al. (2021)	Honeypot data on attack intensity of network exploits	Mar 2013 to Aug 2013	Modeling cyber risk with deep learning and extreme value theory	The proposed method has high accurate prediction power.
20	The County Fair Cyber Loss Distribution: Drawing Inferences from Insurance Prices	Woods et al. (2021)	Insurers' pricing information from SERFF Filing System	Jan 2008 to Dec 2018	Inferring cyber loss distribution from prices	Gamma and Lognormal distributions have better fitting performance.

Appendix B. Multiple change point detection for tail risk

Appendix B.1. Comparison of optimal threshold selection methods

As the first step of detecting change points for tail risk, the reliable estimation of tail risk is crucial. One key issue about tail risk estimation is the choice of threshold. Therefore, we consider the R package "tea" from Ossberger (2020), which contains 12 different ways of selecting the optimal threshold for the estimation of tail risk. There are four methods that are not used in our simulation as they are either not designed for small sample estimation or the running time is significantly longer than other methods due to the coding structure.

To find out which ones to use, we have done the simulation to compare the 8 methods from the package. The basic idea of the simulation is to first generate a heavy-tailed distribution similar to the real data of cyber risk. We use two common distributions, the generalized Pareto distribution (GPD) and the Fréchet distribution. As the original cyber data show extreme heavy-tailedness, we use 0.5, 1, and 1.5 as the tail index, and run 10,000 simulations for each case. In addition, we face the problem of small samples when estimating the rolling window tail index, therefore the sample size (N) of each simulation is set to be 100 and 500 to reflect the special characteristic of our data. We report the mean bias between the estimated and actual index and its variance.

Table B.1 reports the results when the sample size equals 100, and Table B.2 reports the results when $N = 500$. In both cases, two methods provide better results: "dAMSE" from Caeiro & Gomes (2015) and "hall" from Hall (1990). The first method is based on the concept of minimizing the AMSE (average mean squared error) criterion with respect to k (the optimal number of upper-order statistics). The second one uses the bootstrap procedure to simulate the AMSE criterion of the Hill estimator. The unknown theoretical parameter of the inverse tail index gamma is replaced by a consistent estimation using a tuning parameter for the Hill estimator. Minimizing this statistic gives a consistent estimator of the sample fraction k/n with k.

However, these two methods have a systemic downward bias, as shown in Table B.1 and B.2. In other words, these two methods tend to estimate lower tail indices compared to the actual value, especially when the true value is high (see the case when the tail index is 1.5). This is partly related to the small sample issue in our data. As all the methods in the package use Hill's estimator for the estimation of the tail index which is not suitable for the small sample, we consider changing the estimation method to the OLS estimator from Gabaix & Ibragimov (2011) that is specially adjusted for small sample bias. Table B.3 presents the comparison of simulation results for "dAMSE" and "hall" based on both Hill's estimator and the OLS estimator. It can be seen that there is a significant improvement when using the second method, especially when the sample is generated by the Fréchet distribution.

Table B.1 Comparison of optimal threshold selection methods (N=100)

Tail_index	Value	dAMSE	eye	GH	hall	Himp	HW	PS	mindist
GPD									
1.5	Mean bias	-0.3165	0.0817	0.0559	-0.2798	0.9716	-0.3833	-0.5355	0.0261
	Variance	0.0771	0.9135	2.2626	0.0844	517.4994	0.3413	0.0748	0.2554
1	Mean bias	-0.1018	0.1443	0.1526	-0.0723	0.2365	-0.0445	-0.2451	0.1493
	Variance	0.0433	0.5027	3.2748	0.0480	11.9436	0.3962	0.0327	0.1652
0.5	Mean bias	0.0012	0.1211	0.1063	0.0119	0.0719	0.2083	-0.0497	0.1521
	Variance	0.0109	0.1819	0.4037	0.0141	1.8916	4.2911	0.0074	0.0677
Fréchet									
1.5	Mean bias	-0.0774	0.2341	0.2808	-0.0327	0.2207	0.0049	-0.2337	0.2186
	Variance	0.0928	1.2333	6.1720	0.1187	32.4339	12.6251	0.0530	0.3447
1	Mean bias	-0.0477	0.1691	0.2045	-0.0192	0.4161	0.1761	-0.1614	0.1821
	Variance	0.0412	0.5501	5.3721	0.0529	231.9303	10.2856	0.0238	0.1814
0.5	Mean bias	-0.0255	0.1277	0.1029	-0.0110	0.6179	0.4850	-0.0865	0.1436
	Variance	0.0102	0.1951	0.6882	0.0134	1406.7142	66.6224	0.0060	0.0687

Note:
The table reports the comparison of 8 methods for optimal threshold selection. The sample size is 100 for each simulation. Two distributions (GPD, Fréchet) and three tail indices are used.

Table B.2 Comparison of optimal threshold selection methods (N=500)

Tail_index	Value	dAMSE	eye	GH	hall	Himp	HW	PS	mindist
GPD									
1.5	Mean bias	-0.2033	0.0103	0.2035	-0.1708	-0.1339	-0.2310	-0.4608	-0.0632
	Variance	0.0300	0.2537	2.5362	0.0420	4.8221	0.1361	0.0264	0.0483
1	Mean bias	-0.0605	0.0715	0.1926	-0.0392	-0.0341	-0.0293	-0.1997	0.0531
	Variance	0.0163	0.1512	1.6429	0.0219	1.4480	0.0315	0.0108	0.0402
0.5	Mean bias	-0.0054	0.0642	0.1181	0.0042	-0.0043	0.1551	-0.0415	0.0927
	Variance	0.0029	0.0517	0.6448	0.0052	0.0038	28.7445	0.0017	0.0243
Fréchet									
1.5	Mean bias	-0.0699	0.0899	0.3548	-0.0228	-0.0560	-0.0273	-0.1978	0.0737
	Variance	0.0261	0.3050	4.9652	0.0502	0.1587	0.3037	0.0156	0.0610
1	Mean bias	-0.0494	0.0773	0.2277	-0.0164	-0.0408	0.0780	-0.1399	0.0669
	Variance	0.0116	0.1526	2.5970	0.0223	0.1063	6.2950	0.0068	0.0410
0.5	Mean bias	-0.0240	0.0640	0.1019	-0.0080	-0.0204	0.1549	-0.0769	0.0923
	Variance	0.0029	0.0509	0.3327	0.0055	0.0042	5.1581	0.0017	0.0251

Note:
The table reports the comparison of 8 methods for optimal threshold selection. The sample size is 500 for each simulation. Two distributions (GPD, Fréchet) and three tail indices are used.

Table B.3 Comparison of optimal threshold selection methods–Hill's and OLS estimator

		N=100				N=500			
Tail_index	Value	dAMSE	hall	dAMSE-OLS	hall-OLS	dAMSE	hall	dAMSE-OLS	hall-OLS
GPD									
1.5	Mean bias	-0.3216	-0.2872	-0.2047	-0.2010	-0.2029	-0.1704	-0.1343	-0.1229
	Variance	0.0753	0.0790	0.1173	0.1269	0.0310	0.0430	0.0526	0.0642
1	Mean bias	-0.0980	-0.0699	-0.0340	-0.0288	-0.0592	-0.0373	-0.0256	-0.0172
	Variance	0.0435	0.0488	0.0680	0.0747	0.0161	0.0217	0.0249	0.0304
0.5	Mean bias	0.0006	0.0106	0.0153	0.0189	-0.0055	0.0036	0.0046	0.0117
	Variance	0.0114	0.0143	0.0176	0.0189	0.0029	0.0053	0.0050	0.0064
Fréchet									
1.5	Mean bias	-0.0765	-0.0296	0.0061	0.0105	-0.0762	-0.0286	-0.0253	-0.0020
	Variance	0.0944	0.1251	0.1482	0.1699	0.0263	0.0490	0.0395	0.0609
1	Mean bias	-0.0480	-0.0207	0.0044	0.0062	-0.0490	-0.0187	-0.0147	0.0019
	Variance	0.0433	0.0552	0.0671	0.0767	0.0112	0.0219	0.0176	0.0283
0.5	Mean bias	-0.0250	-0.0113	0.0019	0.0026	-0.0245	-0.0091	-0.0077	-0.0003
	Variance	0.0101	0.0128	0.0157	0.0175	0.0028	0.0053	0.0044	0.0067

Note:
The table reports the comparison of "dAMSE" and "hall" with Hill's and OLS estimator for tail risk. The sample size is 100 for the first four columns and 500 for the last four columns. Two distributions (GPD, Fréchet) and three tail indices are used. The results for "dAMSE" and "hall" with the Hill's estimator are slightly different from the ones in Table B.1 and B.2 as we run another simulation of 10,000 times for this table.

Appendix C. Report delay: In-sample analysis

As shown in Figure 3, the data of Advisen contain multiple abnormal peaks due to inaccurate information. Therefore, to understand the true trend of cyber risk, it is necessary to deal with such abnormal data points. Traditionally, the literature tackles this issue by estimating the overall trend and replacing the abnormal points with estimated results (Wang et al. 2021). However, for our data, the problem is more related to the misallocation of cyber cases, which means that we cannot simply replace the extreme number with a lower and smoother one. To repair this anomaly, we assume the date of cyber events without accurate time follows a normal distribution and then replace the original date with a more accurate one. Based on this method, we can smooth the time trend of cyber risk in our dataset. In the following analysis, we will present results with both the original and adjusted data.

For the modeling of delay structure, we have three models available: GLM, GDM hazard, and GDM survivor. Therefore, it is useful to first test whether these models perform well for the in-sample forecast. Since Advisen began to collect data on cyber risk in 2007, we need to exclude all cases that occurred before 2007 to avoid inherent bias in the database. Therefore, we have 163 months from October 2007 to April 2021, and naturally, the longest possible delay period for training is 163 months. But in this case, we would have no data for in-sample forecast, hence it is necessary to select a period when we assume all cyber cases are counted.

As an example, we compare the cumulative proportion of cases reported for different maximum delay periods in Figure C.1 (the delay between the accident date and the first notice date). Although there is an increasing trend in each graph due to more missing values in recent times, we can still find the differences across different maximum delay periods. There is a trade-off between sample size and accuracy for the selection of the maximum delay period. For our case, we choose the period of 60 months since it includes at least 75% of all observable cases and also provides a sample of 104 months for in-sample analysis.

Given the maximum delay period of 60 months and the available sample of 104 months, we choose the 92nd month (so that we can forecast the following year) as the hypothetical present time, which means we only have observations up to this date. Then we censor the data accordingly, apply the models to this incomplete sample and compare their results with the actual number. Figure C.2 shows the results of the median estimated number for original and adjusted data, with 95% posterior predictive interval. Among the three models, GDM hazard has the most accurate confidence interval while GLM performs worst. Figure C.3 provides the sample estimates of $Cov[z_{t,d}, z'_{t,d}]$ by density plots of mean bias and the logarithm of the mean squared error between replicated and observed covariances. This further confirms that GDM hazard is the least biased and GDM survivor comes second for both samples. Therefore, for the out-of-sample analysis, we will focus on the GDM hazard framework.

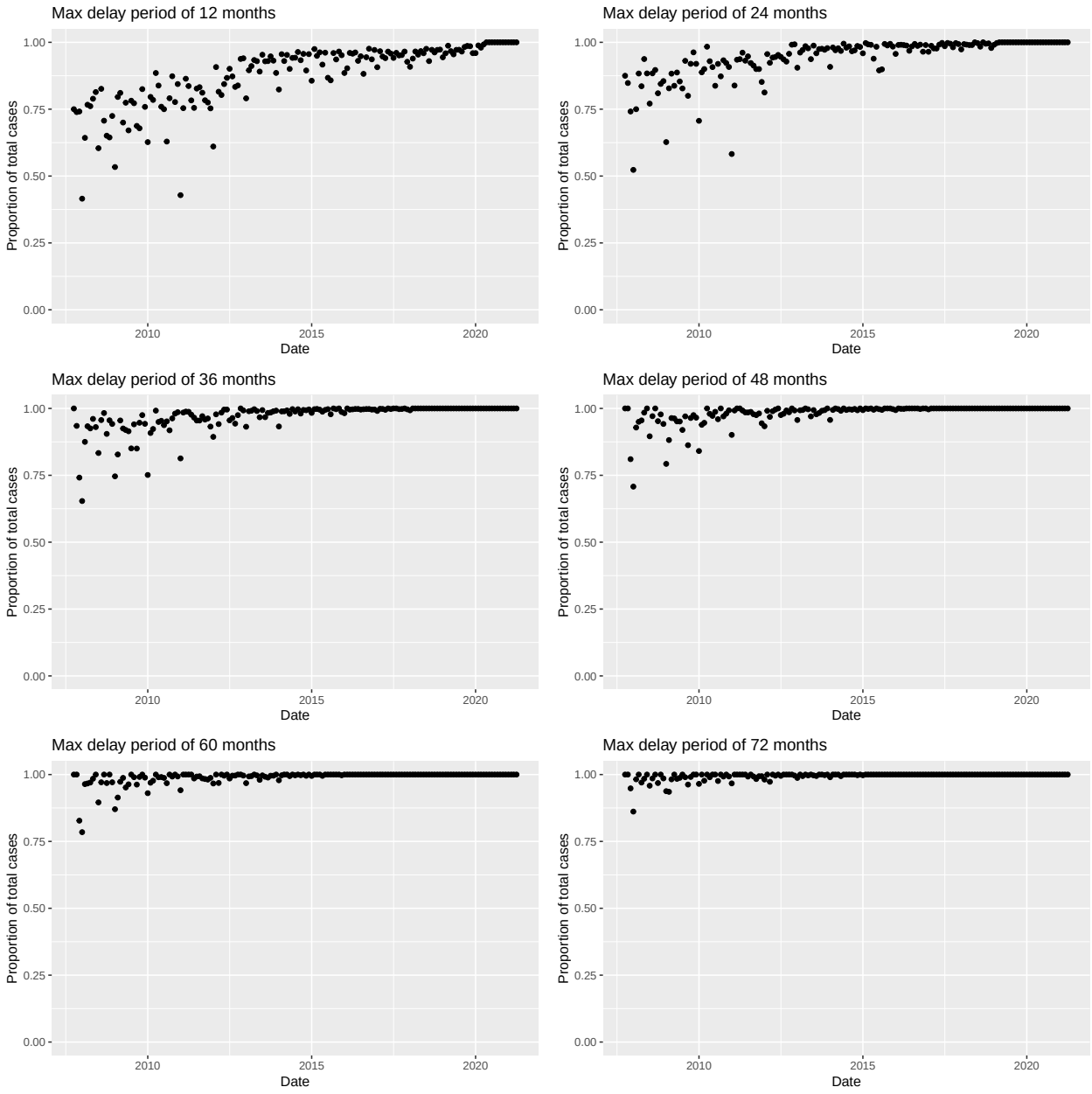

Figure C.1. Cumulative proportion reported

Notes: This figure plots the cumulative report percentage with different delay periods of 12 to 72 months. For each graph, every dot represents the percentage of cases reported in the delayed period out of the whole cases in the data for a specific month of the accident. Therefore, the increasing trend within each graph indicates the issue of report delay for recent periods. But the pattern across graphs shows how a longer period increases the percentage of reported cases.

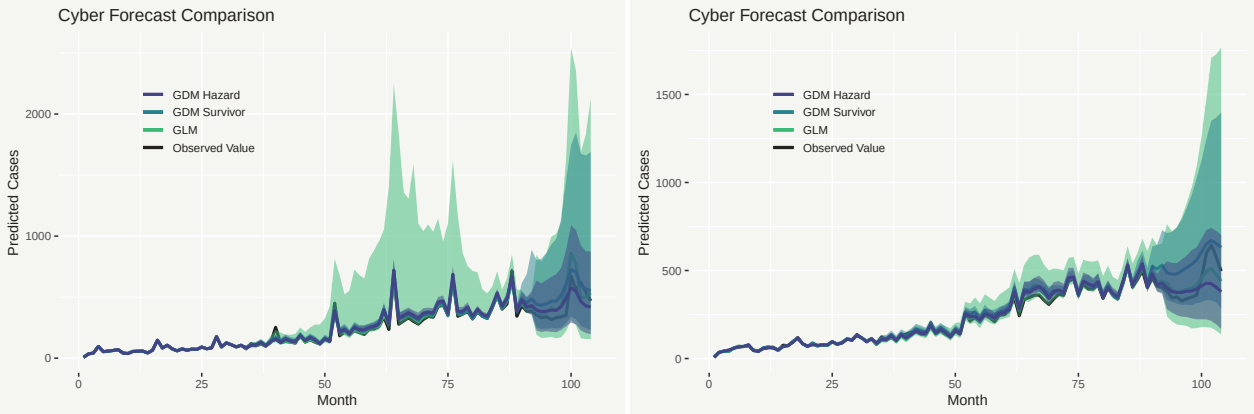

Figure C.2. In-sample cyber forecast comparison

Notes: This figure presents the forecast results of three methods: GDM Hazard, GDM Survivor, and GLM. The adjusted data are the original data after smoothing the abnormal peaks due to unknown dates.

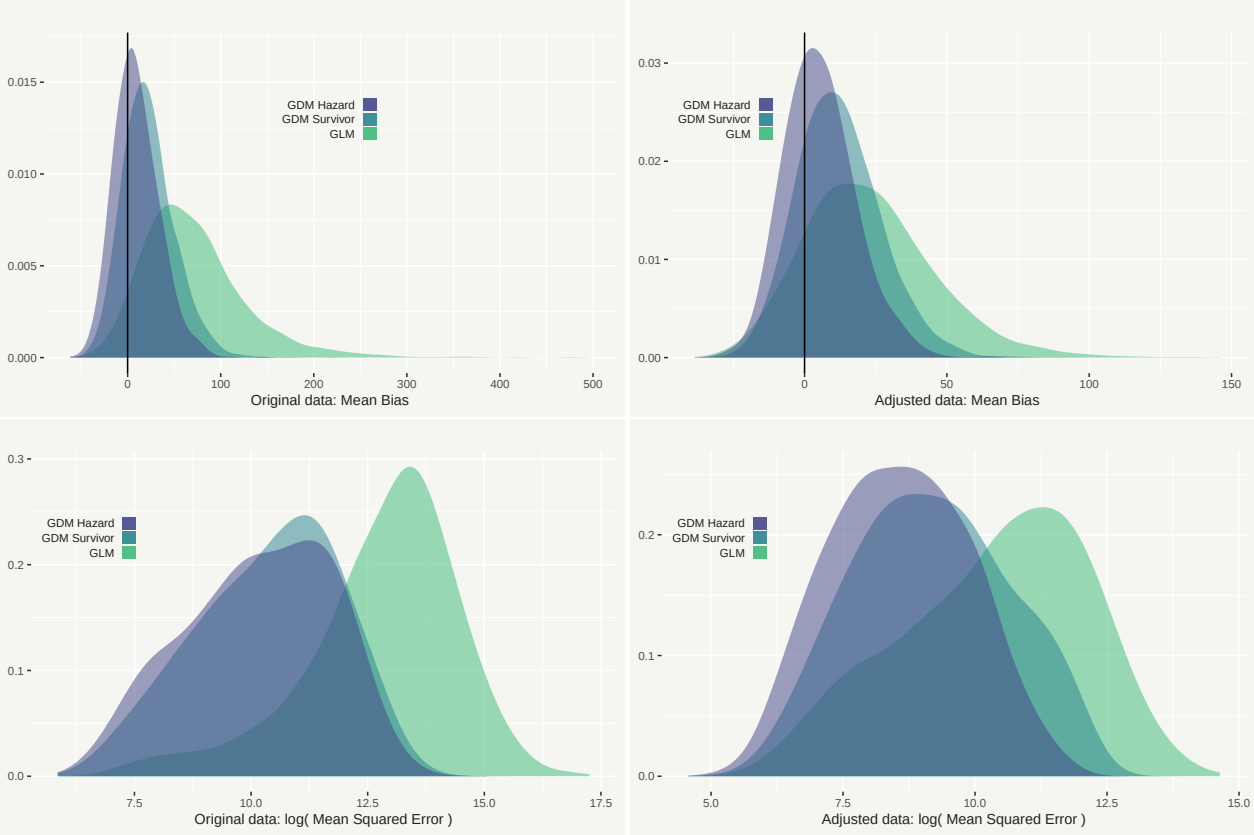

Figure C.3. Covariance of Z

Notes: This figure compares the sample estimates of $Cov[z_{t,d}, z'_{t,d}]$ from three methods by density plots of mean bias and the logarithm of the mean squared error between replicated and observed covariances.

Appendix D. Time dynamics of cyber frequency: Comparison of methods

In this appendix, we compare the results of the method from Baranowski et al. (2019) with other alternatives to show the robustness and advantage of the method we choose. As the increasing pattern is clear in our data, some methods that deal with constant mean are not suitable. Therefore, we consider two alternatives that we can find in the literature. The first one is the methodology proposed by Bai & Perron (2003), which was implemented in R by Zeileis et al. (2022) (denoted as B&P). The change points are estimated by minimizing the residual sum of squares using the dynamic programming approach. The second approach is from Kim et al. (2009), implemented in R by Arnold et al. (2020) (denoted as TF). This method is not designed for the detection of change points but rather for the performance of trend filtering. Still, the results are comparable when we consider piecewise linear signals.

Figure D.1 shows the results of these methods, where TF1 and TF2 are based on different thresholds for change points. The main method and B&P detect similar breakpoints, but the TF method identifies more change points. This is also consistent with the simulation results from Baranowski et al. (2019), that B&P provide similar results while the TF approach is more sensitive and detects more change points.[13] Overall, all the results are consistent with respect to the location of change points but different for the number of change points. This provides further validation for the method we use in the main analysis and the conclusions we draw from the results.

[13]In addition, the main method is much faster than other approaches, more details can be found in Baranowski et al. (2019).

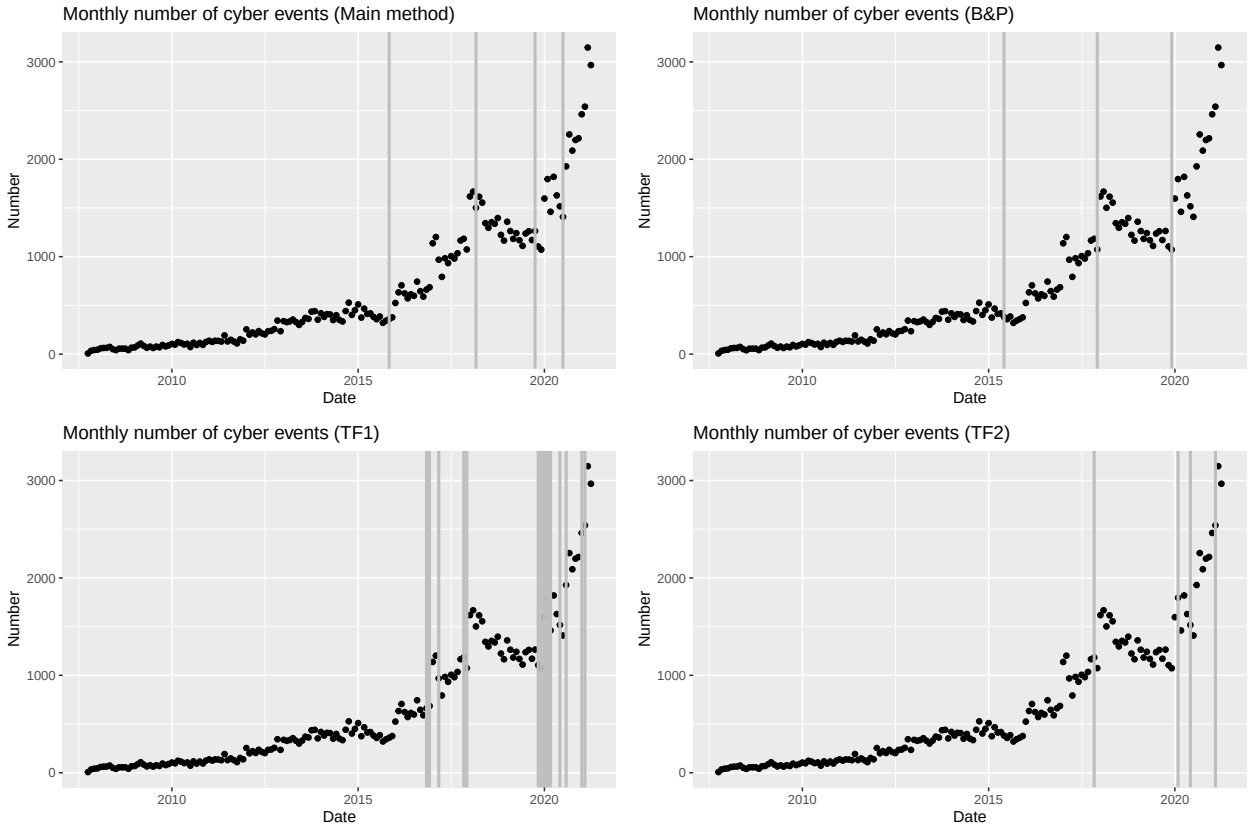

Figure D.1. Comparison of methods for change points of frequency

Notes: This figure compares the results of different change point detection methods for cyber risk frequency after bias correction.

Appendix E. Categorization of cyber risk for three databases

Appendix E.1. Risk type in Advisen

In the Advisen database, there is already a more granular level of categorization and we can rely on this information for our purpose. More specifically, the malicious category includes the risk types such as "Data - Malicious Breach", "Phishing, Spoofing, Social Engineering", "Skimming, Physical Tampering", "Cyber Extortion", and "Identity - Fraudulent Use/Account Access"; the negligent category includes "Data - Unintentional Disclosure"; the privacy category includes "Privacy - Unauthorized Contact or Disclosure" and "Privacy - Unauthorized Data Collection". In addition, there are several types that are not easily distinguishable such as "Industrial Controls & Operations", "Network/Website Disruption", "IT - Configuration/Implementation Errors", and "IT - Processing Errors". These types either belong to the malicious or negligent category, depending on whether there is any malicious party involved. To differentiate these two types, we consider a list of keywords and use this to locate the malicious cases. [14] Finally, the incidents that do not belong to the above categories are classified as "others".

Appendix E.2. Risk type in PRC

In the PRC data, we rely on the variable "type of breach". The type of breach of each incident is indicated with a four-letter abbreviation. In this subsection, we first provide the definition of each four-letter abbreviation:

- CARD: Fraud involving debit and credit cards not via hacking (skimming devices at point-of-service terminals, etc.)

- Hack: Hacked by an outside party or infected by malware

- INSD: Insider (employee, contractor, or customer)

- PHYS: Physical (paper documents that are lost, discarded, or stolen)

- PORT: Portable device (lost, discarded, or stolen laptop, PDA, smartphone, memory stick, CDs, hard drive, data tape, etc.)

- STAT: Stationary computer loss (lost, inappropriately accessed, discarded, or stolen computer or server not designed for mobility)

- DISC: Unintended disclosure not involving hacking, intentional breach, or physical loss (sensitive information posted publicly, mishandled or sent to the wrong party via publishing online, sending in an email, sending in a mailing or sending via fax)

- UNKN: Unknown (not enough information about the breach to know how exactly the information was exposed)

Based on this categorization and the procedure in the case of Advisen data, we define the malicious type as the incidents that belong to "Hack", "INSD", "CARD", and the negligent type as the incidents that belong to "DISC", and the rest are categorized into "others". There is no "privacy" category in this database.

[14] The keyword includes: attack, malware, infect, infiltrate, hack, phish, spam, virus, worm, breach.

Appendix E.3. Risk type in SAS

The categorization of operational risk in SAS is based on the Basel categorization table. [15] There are three levels in this categorization: event risk category (level 1), sub-risk category (level 2), and activity (level 3). Level 1 and 2 in this table are at a very high level and not suitable for the categorization of cyber risk. Therefore, we focus on the activity level and manually check the cyber incidents in each type.

After the manual check, the malicious type includes the following categories: "Account Takeover", "Credit fraud", "Embezzlement", "Extortion", "Fraud", "Insider trading", "Making worthless deposit", "Misappropriation of asset", "Money laundering", "Computer-related fraud", "Hacking damage", "Conducting unauthorized transaction", "Theft of information (w/monetary loss)", "Transaction fraud", "Provision of unapproved access to account", "Insurance fraud", "Forgery", "Hacking damage (if not physical damage)".

The negligent type includes: "Hardware failure", "Software failure", "Telecommunications failure", "Utility outage/disruption", "Failure in obligation to client", "Improper trade/market practice", "Exceeding client exposure limit", "Market manipulation", "Overcharging", "Sale of faulty product", "Anti-competitive action (non-antitrust)", "Commercial right infringement", "Failure in duty to shareholders", "False or incomplete reporting", "Illegal trade", "Improper accounting practice", "libel", "Obstruction of investigation", "Poaching", "Regulation breach/avoidance (non-antitrust)", "Theft of trade secret", "Model error", "Product defect", "Service error", "Accounting error/entity attribution error", "Billing error", "Data entry, maintenance or loading error", "Delivery failure", "Miscommunication", "Missed deadline or responsibility", "Reference data maintenance failure", "Task misperformance", "Failure in mandatory reporting obligation", "Delivery of inaccurate external report", "Recording of incorrect client record", "Damaging of client asset", "Data security failure", "Loss of client data", "Mismarking of position (intentional)".

The privacy type includes: "Breach of privacy", "Misuse of confidential client information", "Suitability/disclosure failure", "Legal document missing/incomplete".

The "others" category includes: "Check kiting", "Non-physical damage abuse", "Theft", "Fire", "Natural catastrophe", "Violence against person", "Violence against property".[16]

In the next step, to make sure we capture all the malicious cases, we use the keyword list above to search for incidents in the negligent, privacy, and "others" category. After this procedure, the categories for cyber incidents in SAS are comparable to the ones in Advisen and PRC.

In addition, the operational incidents in SAS (excluding cyber incidents) are good benchmarks for studying cyber risk, therefore we also categorize these incidents. To simplify the procedure, we broadly classify operational incidents into malicious and negligent cases. The malicious type includes only "Internal Fraud" and "External Fraud", and the rest are categorized into the negligent type.

[15] https://www.bis.org/bcbs/qisoprisknote.pdf

[16] The categories above are not the complete list of categories in the activity level from the Basel categorization table since not every category has cyber incidents.