The changing landscape of cyber risk: An empirical analysis of frequency, severity, and tail dynamics

Martin Eling [1a], Rustam Ibragimov[b], Dingchen Ning[a]

[a] *Institute of Insurance Economics, University of St. Gallen, St. Gallen, Switzerland*
[b] *Imperial College Business School, London, UK*

Abstract

Cyber risk poses severe challenges to modern society and has become an important theme in operational research. Yet relatively little is known about its statistical features and how it evolves. This paper utilizes three databases to examine the properties of cyber risk. We first deal with report delays with an extended two-stage generalized Dirichlet-multinomial framework and then identify structural changes in the frequency and severity of different risk categories using change point detection methods. We document that for malicious events the frequency has grown exponentially in the past two decades and the financial loss distribution has shifted toward greater severity since 2018. The increasing trends for other categories are slower in frequency and less clear in severity. We also explore the tail dynamics by combining tail estimation and optimal threshold selection methods and find that the heavy-tailedness of cyber risk is persistent across all categories. Finally, we discuss the implications of documented empirical features and show theoretically that they lead to lower insurance demand and potentially higher risk levels for firms.

Keywords: Risk analysis, Cyber risk, Report delay, Change point detection, Heavy tails

[1] Corresponding author.
Email addresses: martin.eling@unisg.ch (M. Eling); i.rustam@imperial.ac.uk (R. Ibragimov); dingchen.ning@unisg.ch (D. Ning)

The changing landscape of cyber risk: An empirical analysis of frequency, severity, and tail dynamics

Martin Eling [1a], Rustam Ibragimov[b], Dingchen Ning[a]

[a] *Institute of Insurance Economics, University of St. Gallen, St. Gallen, Switzerland*
[b] *Imperial College Business School, London, UK*

Abstract

Cyber risk poses severe challenges to modern society and has become an important theme in operational research. Yet relatively little is known about its statistical features and how it evolves. This paper utilizes three databases to examine the properties of cyber risk. We first deal with report delays with an extended two-stage generalized Dirichlet-multinomial framework and then identify structural changes in the frequency and severity of different risk categories using change point detection methods. We document that for malicious events the frequency has grown exponentially in the past two decades and the financial loss distribution has shifted toward greater severity since 2018. The increasing trends for other categories are slower in frequency and less clear in severity. We also explore the tail dynamics by combining tail estimation and optimal threshold selection methods and find that the heavy-tailedness of cyber risk is persistent across all categories. Finally, we discuss the implications of documented empirical features and show theoretically that they lead to lower insurance demand and potentially higher risk levels for firms.

Keywords: Risk analysis, Cyber risk, Report delay, Change point detection, Heavy tails

1. Introduction

Cyber risk encompasses a broad range of risks to information and information systems, such as malware, ransomware, phishing, and system errors. It has recently become a major threat to the economy and society, resulting in substantial losses. In 2007, the American department store chain TJX discovered a data breach that exposed the credit card information of 94 million customers (ABC News, 2007). This was the largest recorded data breach at the time and although it was originally considered an outlier, such incidents soon became the new normal. Most notably, Yahoo experienced a breach in 2013 that compromised 3 billion user accounts, setting a new record for the scale of such incidents (Reuters, 2017). Not only are extreme events becoming more frequent, but various studies also indicate rapid increases in the overall frequency and severity. FBI (2022) reports a 127% increase in reported cybercrimes and a 281% increase in financial losses from 2018

[1] Corresponding author.
Email addresses: martin.eling@unisg.ch (M. Eling); i.rustam@imperial.ac.uk (R. Ibragimov); dingchen.ning@unisg.ch (D. Ning)

to 2022. The report by McAfee (2020) estimates the annual cost of global cybercrime at $1 trillion, a more than 50% increase from the 2018 estimation.

This anecdotal evidence underscores the increasing importance and dynamic nature of cyber risk, both of which have motivated the recent works in operational research. Eling and Wirfs (2019) and Dacorogna et al. (2023) use methods from extreme value theory to estimate cyber losses, Zhang et al. (2023) propose structural models for modeling cyber risk propagation process. There is also a strand of literature on cybersecurity, cyber defense, and deterrence from the operational research perspective (Khouzani et al., 2019; Cheung and Bell, 2021; Keith and Ahner, 2021; Welburn et al., 2023). However, there is still limited empirical knowledge about how cyber risk evolves and how it might affect optimal risk management. For example, the optimal security investment depends on the estimated threat level, such as the frequency and severity of security breaches (Gordon and Loeb, 2002). If the estimation is miscalculated due to the changing threat landscape, the optimal investment would be biased and lead to inefficient resource allocation. We provide a comprehensive analysis of cyber risk dynamics by using cyber incidents in the past two decades from three databases and discuss the implications of our empirical results. We focus on two research questions. First, what are the statistical properties of cyber risk and do they change over time? Second, what are the implications for cyber risk management given the evolving cyber threat landscape?

For the first question, we focus on three dimensions of cyber risk: frequency, severity, and tail risk. Before analyzing trends in frequency, we address the issue of report delay bias, which pertains to the time lag between the accident date and the observation date of an event. This subject is underexplored in the literature due to data limitations. Using the unique information in our data, we correct this bias by extending the generalized Dirichlet-multinomial model from Stoner and Economou (2020) to two delay stages. 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 these results, we study cyber risk frequency, especially to understand whether there have been fundamental shifts over time. We use a recent statistical method to detect the unknown number of change points in the time series data (Baranowski et al., 2019). The results show that malicious cyber risk has undergone exponential growth in the past two decades with no significant structural change. In contrast, other categories of cyber risk have grown more slowly.

We also analyze the dynamics of cyber risk severity. To ensure a comparable sample basis and rule out results driven by changes in sample composition over time, we start by matching victim firms each year. Then we proceed to the analysis of loss 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 and Müller, 2020), we consider the full distribution of cyber loss, thereby offering a more comprehensive understanding. The results indicate financial losses from malicious cyber incidents have shifted in a more severe direction after 2018. Given the extreme nature of cyber risks and manifold discussions around their insurability (Biener et al., 2015), the tail of the loss severity distribution requires a closer 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 specially selected optimal threshold selection method using simulations. Then we exploit a change point detection method for the tail index (Ibragimov and Müller, 2016) and find that the tail index is consistently around or below the threshold of 1, indicating extreme heavy-tailedness with no finite mean or variance in the cyber loss distribution.

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 and 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, consistent with the evidence from Swiss Re (2022) that over 90% of cyber losses are not covered by insurance. Furthermore, if report delays lead to severe underestimation of the risk level, the firm may invest less than optimally in security, resulting in higher risk.

Overall, this paper investigates the time dynamics of cyber risk to improve decision-making for this increasingly important risk type. It provides four contributions to the existing literature. First, we uncover the empirical properties of cyber risk and estimate the dynamic trends of frequency, severity, and tail risk using three different databases, advancing the understanding of cyber risk in addition to the previous works in operational research (Eling and Wirfs, 2019; Dacorogna et al., 2023) and applied statistics (Maillart and Sornette, 2010; Edwards et al., 2016). Second, we connect the empirical evidence of cyber risk with the theoretical work on information security (Gordon and Loeb, 2002; Böhme and Schwartz, 2010; Zhao et al., 2013) and show how the documented empirical properties influence the optimal investment in risk management. Third, methodologically we extend the framework of Stoner and Economou (2020) to two stages which better capture the delay structure of cyber loss data. Also, we provide a combined approach with optimal threshold selection, tail index estimation, and change point detection, which is specifically adjusted for the analysis of cyber risk. Lastly, some related studies (Maillart and Sornette, 2010; Farkas et al., 2021) have questioned the reliability of data, but there is still limited empirical evidence addressing this problem. The most relevant works such as Sangari et al. (2023) and Avanzi et al. (2023) identify the existence of report delays for cyber loss data, but we differ from their methods by including a two-stage delay correction model while jointly modeling the delay mechanism and the total count number, which provides higher accuracy for the corrected results.

The remainder of this paper is organized as follows. In the next section, we summarize the recent literature on the statistical analysis of cyber risk and the works on information security and insurance. Section 3 provides a description of our data and the categorization of cyber events. Section 4 discusses the methodology of report delay and change point detection methods. Section 5 presents the empirical results on time dynamics of cyber risk. Section 6 provides a theoretical

framework to discuss the implications of our empirical findings on cyber risk management. Lastly, Section 7 concludes.

2. Related literature

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 and 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 (Eling and Wirfs, 2019). Our research relates to two broad strands of literature: statistical analysis of cyber risk properties and the economics of information security.

2.1. Statistical analysis of cyber risk

The early stage of the empirical work focuses on the general statistical properties of cyber risk, including correlation structure (Böhme and Kataria, 2006; Wang and Kim, 2009a; and Wang and Kim, 2009b) and time trends (Maillart and Sornette, 2010; Wheatley et al., 2016; Edwards et al., 2016; and Romanosky, 2016). Starting from Eling and Loperfido (2017), more and more studies begin to study cyber risk frequency and severity by fitting existing statistical models (Eling and 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; and Dacorogna et al., 2023). These works have exploited the available databases 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. More recently, there are studies exploring the determinants of cyber risk (Aldasoro et al., 2022; Malavasi et al., 2022).

The study on the time dynamics of cyber risk has been scarce and results are inconsistent (Woods and Böhme, 2021). For example, using the data period from 2000 to 2008, Maillart and Sornette (2010) show a strong non-stationary growth culminating in July 2006, followed by a stable period. Edwards et al. (2016) find no evidence of an increasing trend for the size and frequency of data breaches 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 period, but only specific to hack-type events. Overall, the results appear to be somewhat inconsistent and the difference may largely stem from different datasets and methodologies. This inconsistency motivates us to reexamine the empirical properties over a long horizon with the comparison of three cyber databases. We also note that none of the above studies incorporates the bias problems inherent in all datasets.

[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.

2.2. The economics of information security and insurance

To deal with cyber risk, it is critical to invest in information security. Gordon and Loeb (2002) is among the first to consider an economic model for optimal investment in information security and shows that firms should prioritize 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 and 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 tool to model the interactions among participants(Laszka et al., 2014).

Among the possible mechanisms for improving security and transferring financial risk, 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 and Schwartz (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 and 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 and Liu, 2014), etc. [3]

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.

3. Data

3.1. Cyber loss data

Directly measuring cyber risk is challenging. In this paper, we analyze cyber risk using its realizations, that is, cyber incidents that have actually occurred. We look at three sources of data and all data focus on events that occur to legal entities (firms, public, and non-profit institutions) rather than to individuals. We examine two types of losses. The first type is the volume of information, measured by the number of affected records or accounts. The second type is the monetary losses arising from the incident. This includes first-party losses, such as the value of the lost records or the cost of business interruption, as well as third-party losses, which encompass payments to affected customers and fines in case of regulatory violations.

Our first data source is Advisen, whose database collects information from multiple publicly available sources such as government websites, official court documents, and other online resources.

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

This data has been increasingly analyzed in the academic context (Aldasoro et al., 2022; Malavasi et al., 2022). 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 countries. 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 from 2001 to 2021,[4] we have 5,789 incidents with known financial loss and 90,821 incidents with the known number of affected units.

The second one is SAS OpRisk Global data, the world's largest database on publicly reported operational losses. This database contains more than 35,000 operational events with losses above US$ 100,000 for various countries and industries. However, there is no classification for cyber incidents, so we cannot extract these directly from the data. We follow Eling and Wirfs (2019) and exploit text mining to extract cyber-related events. This resulted in 2,123 observations of cyber events and 18,287 observations of non-cyber operational events from 2001 to 2021.

The third source of data is the non-profit organization Privacy Rights Clearinghouse (PRC), a database that has been frequently used in the literature (Kamiya et al., 2021; Sun et al., 2021; Farkas et al., 2021). It collects information about breach events from government agencies and verifiable news sources in the U.S. starting from 2005. This data contains 6,733 records (with non-zero ID losses) up to the end of 2019.[5] The major difference from the previous two sources is that this database focuses only on data breaches and does not provide information about financial losses.

Although there are three databases, all of them gather information on cyber incidents. We study these databases separately as we aim to find the general pattern of cyber risk that is persistent across sources and categories. For example, the SAS dataset exhibits a bias toward incidents involving large losses; the PRC dataset is skewed toward data breaches. Among the three, the Advisen dataset is the most comprehensive but often includes incidents for which only limited information is available. A comparison could provide robust results on the dynamics of cyber risk, which contributes to the literature as all previous studies consider only one dataset to our best knowledge.

3.2. Categorization and descriptive statistics

To differentiate various categories of cyber incidents, we consider four categories, building upon and expanding the literature such as Edwards et al. (2016). The first category, "malicious", encompasses malicious cases, defined as cyber incidents initiated with harmful intent against victims, such as hackers infiltrating the system of the victim firms or internal employees illicitly profiting from stolen confidential information. The second category, "negligent", covers negligent cases, which are cyber incidents without malicious third parties, such as system errors or accidental disclosure of consumer information. The third category focuses on privacy violations, hereafter referred to as "privacy". This category includes incidents where a firm intentionally mishandles clients' information. While not causing direct damage to the firms, such incidents may lead to lawsuits and fines

[4]Since the database creates separate records for different kinds of losses arising from one incident such as direct damage and legal costs, we aggregate the original data for analysis. We restrict the sample to the period from 2001 given the scarcity of data from the 20th century. This also applies to other data sources.

[5]This data has a shorter end period due to maintenance issues.

for legal violations. The fourth category, "others", encompasses all other cyber incidents, including those of unknown cause or those triggered by external factors like natural events or physical damage. Further details are provided in Appendix A.

Table 1 summarizes key statistics of each category of cyber incident across three databases. In Advisen, fewer than 10% of incidents in each category have known financial loss information, while information on the affected count is available for around 80% of incidents. The financial loss of malicious cases is more severe than that from negligent cases, but the affected counts are higher for negligent cases than for malicious ones.

In SAS, the magnitude of losses is higher than in Advisen. This arises because SAS only includes cases with losses exceeding US$ 100,000. Negligent cases lead to higher loss severity than malicious cases, in contrast to Advisen data. This might be explained by the composition of the two samples, as the Advisen sample includes more incidents related to unintentional disclosure which results in lower financial losses.

In the PRC data, malicious cases have a higher number of records lost or affected than negligent cases. This is different from the Advisen data as the PRC data focuses on data breaches and thus has a higher loss related to the number of records.[6] Despite differences in size and types of losses across the three databases, the loss distributions are all heavily skewed as the median and mean values are significantly different.

4. Methodology

4.1. Report delay

Reliable data are crucial for analyzing cyber risk, yet existing databases may be biased. Consequently, empirical studies without bias correction could yield incomplete or even incorrect conclusions. We aim to extend recent methods from the field of statistics to identify and correct the potential bias in the data before conducting further analysis. One main problem is report delay[7]—the period before the total number of observable events becomes available. In the interim, we can only observe incomplete data, which may adversely affect our understanding of time dynamics and lead to misinterpretation of the actual number of events. In the realm of cyber risk, this problem is prevalent since many events are noticed and made public after a long time. Also, a delay may occur when the data provider cannot update the records in time due to limited maintenance resources.

To model report delay, we follow Stoner and Economou (2020) and extend their framework to include two stages that are unique in the Advisen dataset. The Advisen dataset is the focus of this section because it provides the detailed timeline of each incident, from the accident date to the date of first notice (first stage), until the date of creation in the database (second stage). This unique

[6]The terms "affected counts" and "number of records" are used separately in Advisen and PRC, but they are practically the same in our context.

[7]There may be other possible biases, but our data does not provide sufficient information for direct correction. For example, selection bias, where certain types of cyber losses are more likely to be reported, can be a concern for cyber severity analysis. We attempt to control for this indirectly in Section 5.3.

Table 1: Summary statistics of three databases

	Sample size	Number of cases with known losses	First quartile	Mean	Median	Third quartile	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
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. The dollar value is adjusted for inflation with 2021 as the reference year.

feature allows us to capture two delay mechanisms.[8] The reason we choose Stoner and Economou (2020) is that they provide high accuracy by jointly modeling the delay mechanism and the total count number. Traditionally, the task of correcting delayed reporting has been treated separately from the task of forecasting, overlooking 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 processes.

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 generalized Dirichlet-multinomial framework with hazard model (GDM hazard) is defined by:

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

where NB is a negative binomial distribution, λ_t is the expected rate of occurrences and θ allows

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

for overdispersion. 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, $\nu_{t,d}$ is the expected proportion of counts that will be reported at delay d out of those which are yet-to-be-reported and $\beta_{t,d}$ allows for temporal changes of delay mechanism. In addition, $\boldsymbol{\phi}$ controls for dispersion, ι and ψ_d are fixed effects. In this model, the delay mechanism is modeled through the difference of temporal structure in the proportion of reported cases across delay levels. A discussion of model performance is provided in Appendix B.

This model provides a flexible way of modeling delay structures for cyber risk, but how to connect two delay stages in our data remains a problem. Given that the available data represent the results after two delay stages, we can back-trace the original trend starting from the second stage and then proceeding to the first stage. In the second stage, assume that for the time of first notice t, the number of total cases is a_t but not available. Suppose after D months all the cases will be included in the database, but we have data only for $D' < D$ months. Therefore, after applying the GDM hazard method, 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: $\hat{b_{t,d}} = b_{t,d} * q_{t+d}$. After the adjustment, we apply the GDM hazard model to account for first-stage bias and obtain corrected results.

4.2. Time dynamics of loss frequency

We study loss frequency and in this context focus on the estimation of change points over the period since it is of interest to understand whether the cyber risk has undergone certain fundamental changes in the past two decades. 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. However, the literature mostly focuses on the problem under the assumption of piecewise-constant parameters but 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). [9] 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 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.

4.3. Time dynamics of loss severity

Dubey and 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 and Müller (2020) further provides an inference method

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

for the identification of change points in a sequence of distributions and we can use the bootstrap approximation to calculate critical values.

4.4. 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).[10] 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 finite if and only if $\zeta > 1$.[11] In this section, we provide a combined approach that incorporates multiple stages of tail estimation for cyber risk analysis.

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

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

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

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

$$log(i - 1/2) = a - \zeta log(x(i)).$$

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. As there is no consensus on the choice of the threshold selection method, we conduct a simulation analysis based on the stylized properties of cyber losses to compare the in-sample performance of various methods. We consider the R package "tea" from Ossberger (2020), which includes 12 approaches. The simulation results show that two methods ("dAMSE" and "hall") provide better performance than the others in the package. However, there is a downward bias in these two methods and we find using the OLS estimator significantly reduces that bias (details in Appendix D). Therefore, we use these two methods in combination with the OLS estimator for the estimation of the tail index.

[10] 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.

[11] 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.

Change point detection: To analyze the trend or potential change points in the extreme value index, we exploit the method from Ibragimov and Müller (2016). The empirical strategy is to partition the sample into two periods: before and after a possible break point. Then we divide each period into q groups chronologically (in this paper we use four groups), and compute the Behrens-Fisher (BF) statistic:

$$BF = \frac{\hat{\zeta}_1 - \hat{\zeta}_2}{\sqrt{\frac{(s_1)^2}{q_1} + \frac{(s_2)^2}{q_2}}},$$

where $\hat{\zeta}_i = q_i^{-1} \sum_{j=1}^{q_i} \zeta_{i,j}$, $(s_i)^2 = (q_i - 1)^{-1} \sum_{j=1}^{q_i} (\zeta_{i,j} - \hat{\zeta}_i)^2$, $\zeta_{i,j}$ is the tail estimator for period i and group j with $i = 1, 2$ and $j = 1, 2, 3, 4$. 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.

5. Empirical results

5.1. Report delay

To understand the problem of report delay, we first briefly compare our three databases. To ensure their comparability, we have restricted the period to 2005 and onwards and focus only on the malicious category in the U.S. (the PRC data starts from 2005 and covers only U.S. data). Various sources and reports (Accenture, 2021; Allianz, 2021) indicate that cyber risk has been increasing quickly over the years, but as shown in Appendix Figure 1, this trend is not as significant 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 pattern in recent years. The sudden drops in the number of cases in 2018 for PRC and in 2020 for Advisen indicate that the problem of report delay may be one reason. To look more closely at the problem of report delay, we use the date of creation in Advisen to show how the trend has evolved over the years in Appendix Figure 2. We plot the evolution of cyber risk based on four creation dates (every four years from 2009 to 2021) so that each graph shows only cyber events before the creation date. 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.

We apply the two-stage method with the GDM hazard framework introduced in the methodology section to the whole sample period. The result is shown in Figure 1. The increasing trend for the malicious, negligent, and "others" categories 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 presented 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. We also apply the same procedure to the SAS data and find similar results (see Appendix B).

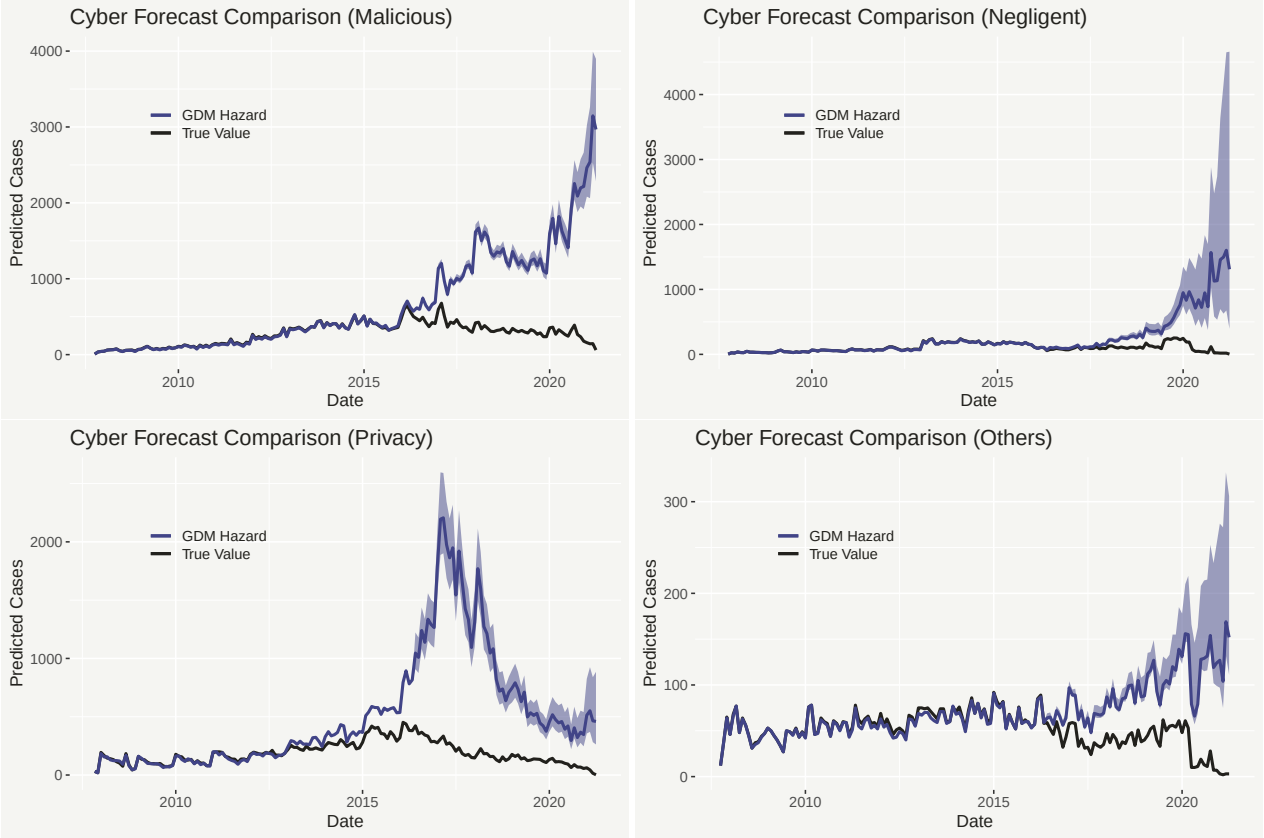

Figure 1: Bias correction for the Advisen data

Note: 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.

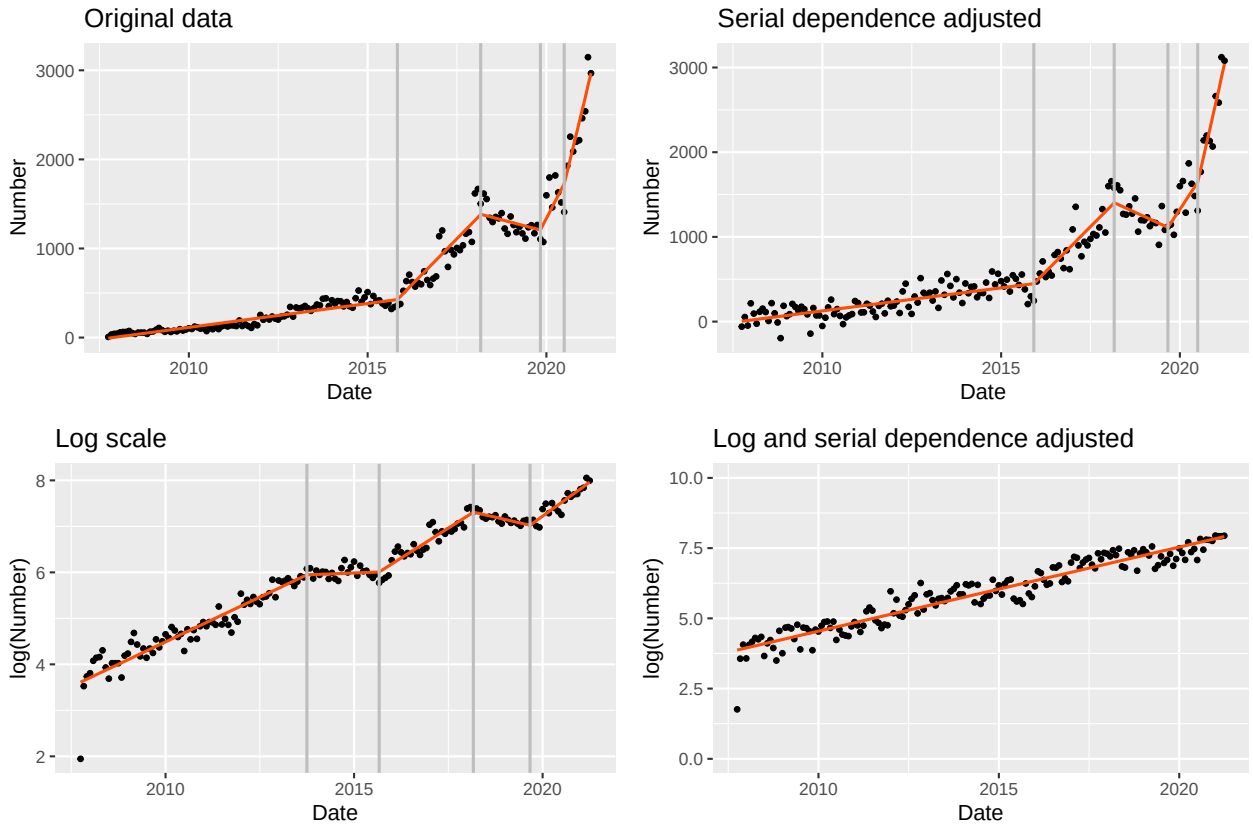

Figure 2: Change points for loss frequency (Malicious)

Note: 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. The red lines are the predicted linear trends for different periods.

5.2. Time dynamics of loss frequency

To better understand the dynamics of cyber loss frequency, we apply the narrowest-over-threshold method to the bias-adjusted time series data in Advisen. The top left graph in Figure 2 plots the dynamics of cyber incidents of the malicious category with change points as the grey vertical lines. Given that we are working with time series data, serial dependence can be a 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 as the noise generated in this manner is sufficient to remove the serial dependence of the residuals. The top right graph of Figure 2 plots the result after adjusting serial dependence and the overall pattern is consistent with the original results.

However, the current results assume a linear pattern, but the fact that there are multiple phases with increasing slopes in 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. To control for the potential bias from serial dependence, we add additional IID Gaussian noise with zero mean as in the previous case. Since the noise is random, there might be different results depending on the actual values of the Gaussian noise in simulations. We consider the most conservative results with the fewest change points among different possibilities. This is shown in the bottom right graph of Figure 2, providing further evidence that the linear pattern assumption is not likely to hold and that cyber risk of the malicious category has in fact undergone exponential growth in the past two decades. There might be certain disturbances such as from March 2018 to December 2019 as shown in the bottom left graph, but these disturbances do not significantly alter the overall pattern.

To quantify the growth rate of malicious cyber events, we estimate the slope of the fitted line and find the frequency is increasing following the exponential form $e^{0.025t+3.85}$, where t is the month of concern since October 2007. For example, the monthly malicious events by January 2025 would be 8,518 cases, up from 2,565 in January 2021 if the pattern continues. This rapid growth is related to the exponential increase in the usage of IT technologies and the exploitation of their vulnerabilities.

For other categories of cyber risk in Advisen, we present the results in Figure 3 with log transformation and serial dependence adjustment. The top left graph is the result for the malicious category. The top right graph presents the possible change points for the negligent category. After trying different simulations of Gaussian noise to the data, the decreasing phase between August 2014 and January 2017 is still present. Therefore, for the negligent category, the exponential growth is not continuous but subject to certain disruptions. In the recent period, the monthly frequency of negligent cases follows the pattern $e^{0.054t+4.45}$, where t starts from January 2017.

For the privacy category, the time pattern is significantly different. 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 that has continued to date. 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 their permission. The declining trend after 2017 might be driven by the possibility that more and more firms comply with the acts after a significant number of lawsuits are filed by consumers. Furthermore, the European Union adopted the General Data Protection Regulation in April 2016, which enhances individuals' control and rights over their personal data. This also might contribute to the decline in the number 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 rate is much lower than for the malicious and negligent categories. Therefore, we will not focus on this category in the following analysis. The comparison of different categories indicates that the malicious category is the biggest threat as the number of malicious incidents is growing exponentially and shows no sign of slowing down in the observed period.

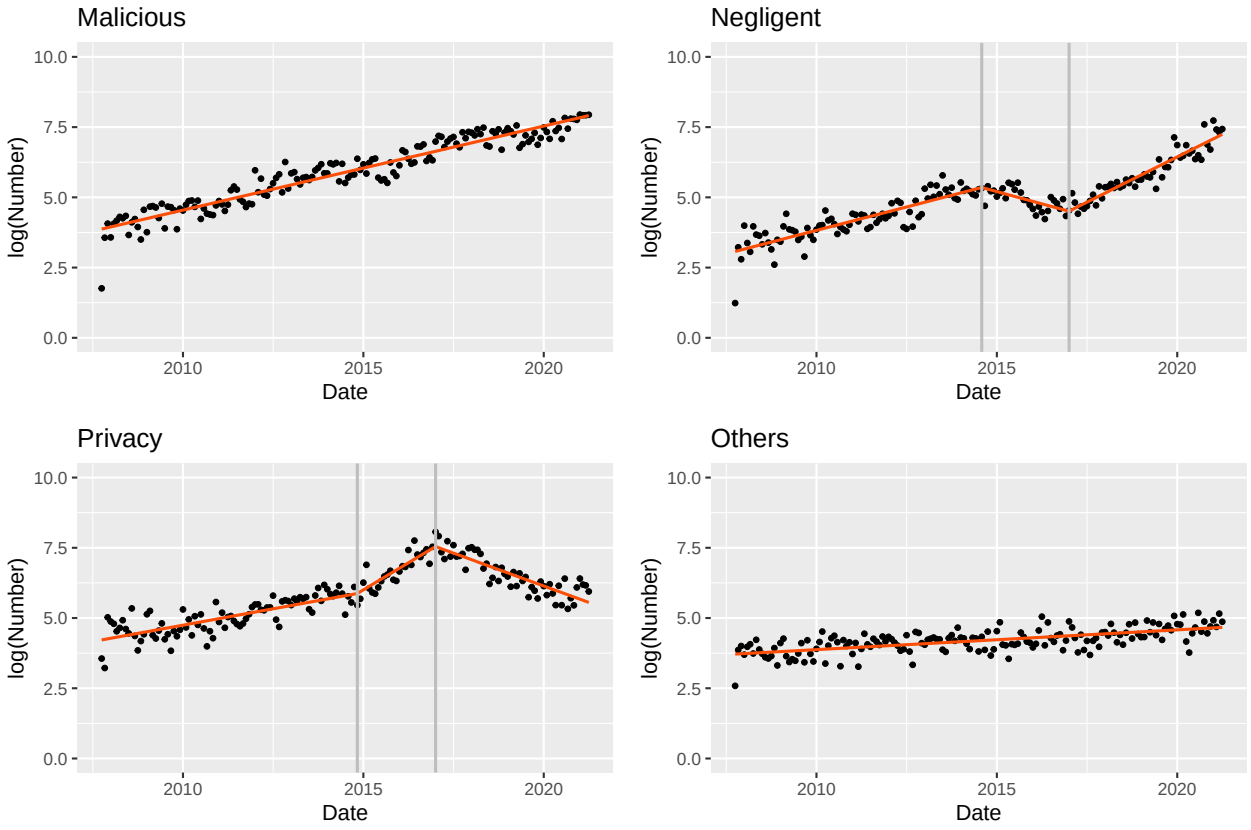

Figure 3: Change points for loss frequency by category

Note: 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. The red lines are the predicted linear trends for different periods.

5.3. Time dynamics of loss severity

In addition to cyber loss frequency, the severity of cyber losses is another key dimension for measuring cyber risk. However, when we compare the loss distributions across different years, it might be difficult to draw conclusions due to potential selection bias. For example, cyber incidents affecting public firms might be more likely to be reported prior to mandatory disclosure, and thus a perceived declining trend of severity could occur when more losses from private firms become reported. To control for such selection bias, we first estimate the propensity score of firms for each year based on characteristics such as revenue, number of employees, and industry category. Then we exploit inverse probability weighting to adjust the weights of different types of firms so that we have comparable samples across different years.

Figure 4 presents the annual density plot for the weight-adjusted cyber losses sample. As the adjustment method requires additional information for each record, we focus on the Advisen and SAS data from 2006 due to limited data.[12] The upper panel in Figure 4 compares the financial losses from malicious cases in Advisen and SAS and there is a common pattern in which the distribution is shifting to the right in the recent period, meaning the severity of cyber incidents is increasing over time. The lower panel depicts the distribution of the number of affected records or accounts per event. There are no easily discernible trends for malicious or negligent cases.

To identify the structural change in loss distributions, we apply the method in Section 4.3. Figure 5 shows the loss distribution before and after the identified change point. For financial losses of malicious cyber incidents, the structural change occurred in 2018. After this change, the distribution shifted to the right, which confirms the observation in Figure 4. Quantitatively, the empirical average loss per event increased from \$162,832 to \$1,250,783 in Advisen, while the average loss rose from \$1,617,173 to \$2,947,771 in SAS after transforming the log scale back to the linear scale. The loss numbers are much lower than the mean in descriptive statistics (Table 1) because the weights are adjusted in our sample and very likely large firms with large losses are overrepresented in the raw data. Still, this change indicates that malicious cyber losses are becoming more severe over time for the same set of firms.

The lower panel in Figure 5 shows that the distribution of affected records or accounts has a slight shift after the change point for both malicious and negligent cases. More specifically, the empirical average decreased from 976 (malicious) and 572 (negligent) to 458 (malicious) and 232 (negligent) affected records per event after 2007. The numbers are again lower than those presented in Table 1, a difference likely attributed to weight adjustments. Taking the results together, the general pattern is that the financial losses are increasing for malicious cyber incidents but not non-financial losses. Still, for extreme cases, a closer examination of the tail of the distribution is necessary.

[12] In Appendix E, we provide analysis without adjusting the weights so that the trends of all categories of cyber losses are presented. The results for malicious cases are consistent with the adjusted results. For negligent cases, the results suggest a decrease in severity, but it is difficult to conclude without matching.

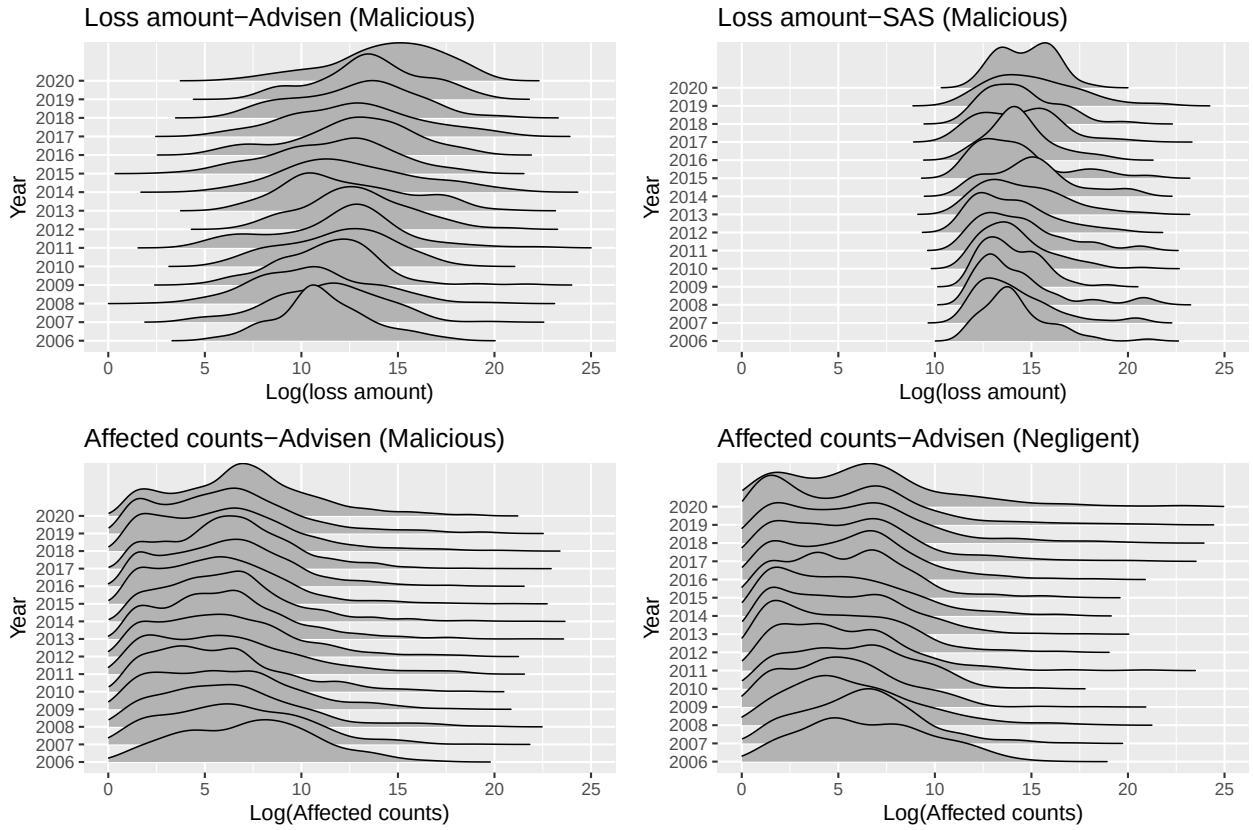

Figure 4: Comparison of loss distributions

Note: This figure presents the comparison of distributions for the financial losses and affected counts (log scale) in Advisen and SAS. The density plot is based on the weight-adjusted sample. As the events in SAS have losses of at least $100,000, the distribution is located to the right compared to the distribution in Advisen.

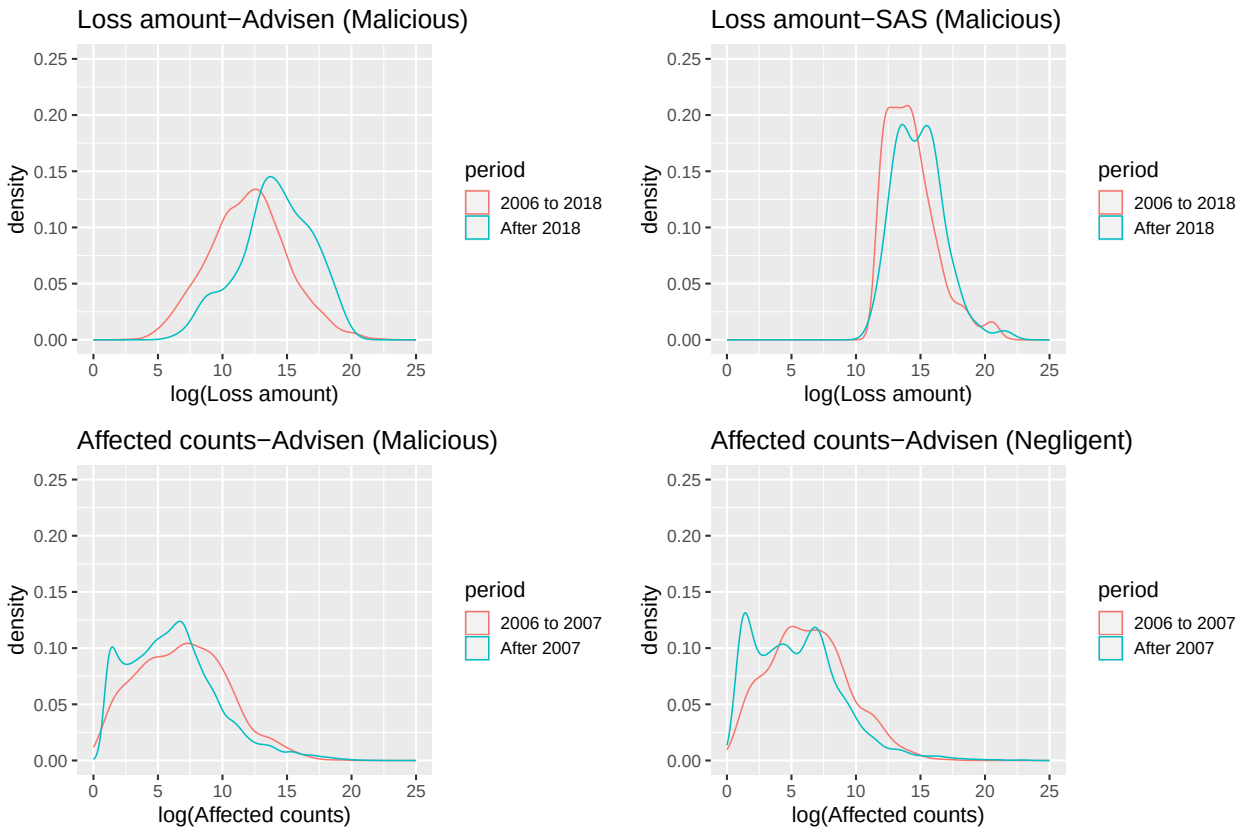

Figure 5: Change points of loss distributions

Note: This figure presents the comparison of distributions before and after the identified change point for different categories of cyber incidents. The red line shows the density plot before the change point and the blue line shows the density plot after the change point.

5.4. Time dynamics of tail risk

5.4.1. Estimation of cyber tail index

We first provide a detailed comparison of the tail index in Table 2 using two methods ("dAMSE" and "hall") mentioned in the previous section 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. [13]

The estimation in Table 2 shows that different kinds of cyber risk all yield severe heavy-tailedness, with a tail index lower than 1 for most cases. This is a serious concern as there is no finite mean for cyber loss distributions. This also confirms the results from the literature (Eling and Wirfs, 2019; Farkas et al., 2021) that cyber loss distribution should be modeled by heavy-tailed distributions. Furthermore, the monetary losses exhibit a less heavy tail compared to the non-monetary losses such as the number of records breached. This is reasonable given the rapid increase in data storage capacity in recent years, and thus the number of records affected by a single cyber incident can be extremely high. Consequently, the tail for non-monetary losses can be heavier than that for monetary losses.

5.4.2. 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 6 plots the rolling window estimation of the tail index for financial loss of cyber incidents and shows the most probable change point in the time period using the aforementioned approach. 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. However, for the negligent cases, the general trend is relatively stable. Figure 7 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, which is likely related to the differences between the two data sources. But the tail index remains below 1 over time. For the negligent cases, there is a clear and consistent declining trend after 2016, despite the differences between the two databases. This leads to an even heavier tail for negligent cyber risk. Appendix F provides more details of the comparison before and after the change point.

[13] For most of the cases, the results using two threshold selection methods ("dAMSE" and "hall") are reasonably close, but there is a sizeable 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 "eye" and "mindist", respectively. Hence, the result from "dAMSE" is more reliable. In the following analysis, we will 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).

Table 2: 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	85	0.87	0.69	1.06	0.90	0.63	1.18
dAMSE	156	0.76	0.64	0.88	0.85	0.66	1.04
Advisen (loss amount)–Negligent							
hall	36	0.55	0.37	0.73	0.63	0.34	0.92
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	48	1.11	0.79	1.42	1.14	0.68	1.59
dAMSE	27	1.06	0.66	1.46	1.18	0.55	1.80
Advisen (affected counts)–Malicious							
hall	310	0.64	0.57	0.71	0.85	0.71	0.98
dAMSE	687	0.52	0.48	0.56	0.64	0.57	0.71
Advisen (affected counts)–Negligent							
hall	2054	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	535	0.43	0.39	0.46	0.43	0.38	0.49
dAMSE	166	0.43	0.36	0.49	0.48	0.38	0.59
PRC (number of records)–Negligent							
hall	1217	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: This table presents the tail index estimation of different cyber categories from three databases. 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. The estimated tail index and 95% confidence interval (CI) are provided for each estimation method.

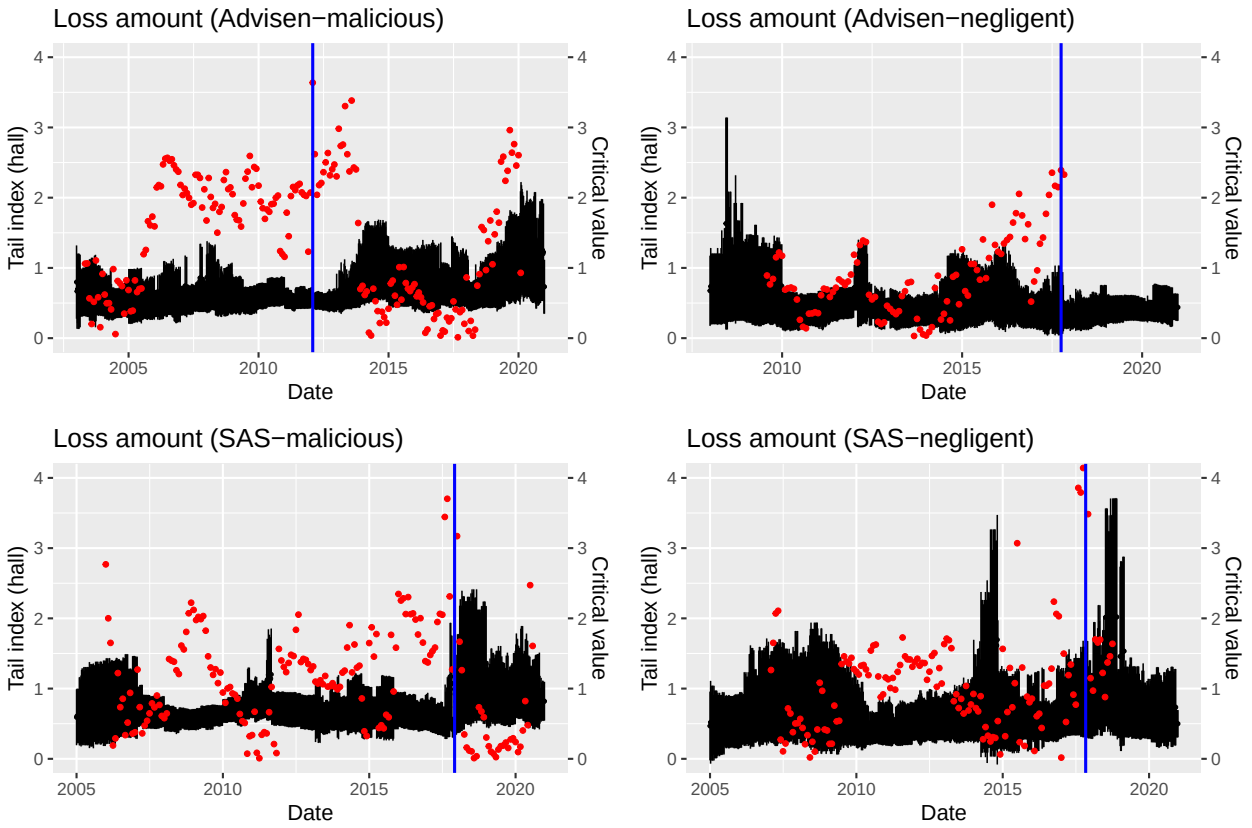

Figure 6: Change points for tail index (loss amount)

Note: 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 and Müller (2016). The blue line is the point with the highest BF value, which indicates the most possible change point in the whole period.

Overall, the results across different databases indicate that the cyber tail index has remained relatively stable, consistently around or below the threshold of 1, even with changes over the past two decades. Therefore, tail risk will remain a critical aspect of cyber risk management for the foreseeable future.

6. Implications for risk management

One distinct feature of cyber risk is its dynamic nature, which poses a serious challenge for risk management. 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, the perception and estimation of cyber risk can vary substantially among different parties. In a basic setting of a firm and an insurer, the insurer is likely to hold an information advantage since it specializes in risk management. Furthermore, insurers routinely deal with report delays as part of their standard procedures for calculating claim reserves. However, the decision-makers in firms may have more information about their own vulnerability related to cyber

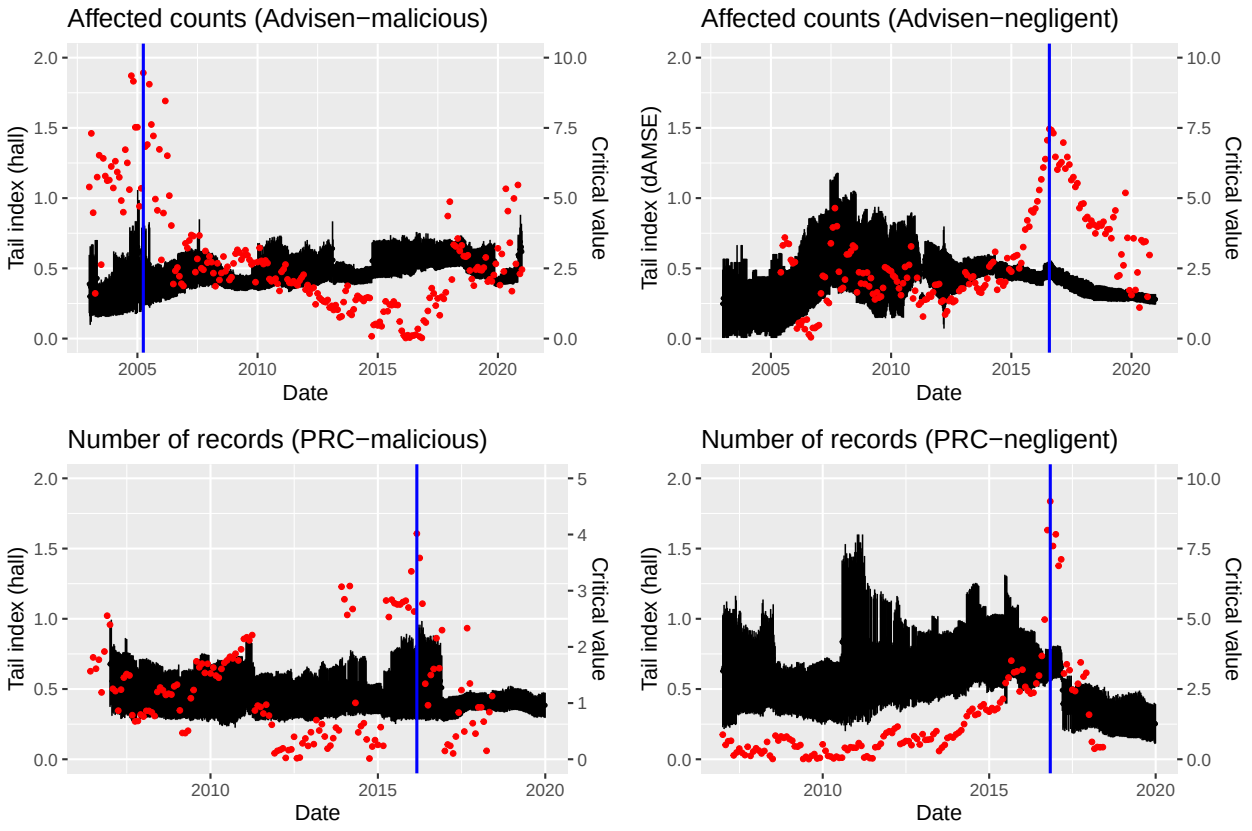

Figure 7: Change points for tail index (number of records)

Note: 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 "dAMSE" method). The red dots are the BF values from Ibragimov and Müller (2016). The blue line is the point with the highest BF value, which indicates the most possible change point in the whole period.

risk but lack expertise in analyzing the trend of cyber risk. Thus, there can be an information gap between the firm and the insurer, which leads to diverging estimates of cyber risk probabilities.

In addition to the problem of delayed information, we provide evidence for the extreme heavy-tailedness of cyber risk. As mentioned earlier, Ibragimov et al. (2009) have shown this feature might induce the non-diversification trap, resulting in no market for cyber risk insurance. In practice, the insurance market exists and has been increasing, but insurers tend to offer contracts with low coverages to avoid extreme scenarios that might severely undermine their financial stability. Although this strategy can be useful for protecting insurers from extreme tail risk, this level of coverage falls short in supporting businesses with increasingly high exposure to cyber risk and consequently limits the value of insurance as a risk management tool.

To understand how the feature of delayed information and tail risk affect the optimal decision of cyber risk management, we adopt the widely used framework from Ehrlich and Becker (1972). 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. The firm assumes the probability of cyber loss is $p_0 < 1$, while the true probability is $p_1 > p_0$. We introduce tail risk by considering a small probability ϵp of bankruptcy for the firm. Therefore, in the event of cyber loss, there is a probability ϵp that the loss exceeds the initial wealth, $L_m >> w_0$, and a probability $(1-\epsilon)p$ that the loss L is medium such that $0 < L < w_0$. 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 (Froot and Stein, 1998).

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 $\alpha\pi(p)L$, where $0 \leq \alpha \leq 1$. The insurance coverage is 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. The firm could reduce the probability of cyber risk by implementing 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 $\lim_{x\to\infty} p(x) = 0$. $p(p_1, x)$ and $p(p_0, x)$ denote the true probability and the perceived probability by the firm after self-protection. Hence, the maximization problem of the firm is

$$\begin{aligned} EU = (1 - p(p_0, x))u(w_0 - x - \alpha\pi(p(p_1, x))L) \\ + (1-\epsilon)p(p_0, x)u(w_0 - x - L + \alpha L - \alpha\pi(p(p_1, x))L) + \epsilon p u(0), \end{aligned}$$

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

Therefore, the first-order condition (FOC) with respect to the optimal coverage α is

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

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

Rearranging the equation, we have

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

The right-hand side is smaller than 1 as $p(p_0,x) < p(p_1,x)$. Therefore, to ensure the equality of FOC, $u_1' < u_2'$ holds. In the standard model without tail risk and delayed information, we have $u_1' = u_2'$, which means the state with insurance and without insurance is equal and thus full insurance is optimal. But with additional features, the results indicate the firm opts for only partial insurance.

For optimal self-protection, FOC is shown as

$$\begin{aligned}\frac{\partial EU}{\partial x} = &-p'(p_0,x)u_1(\cdot) + (1-p(p_0,x))u_1'(\cdot)(-1-\alpha p'(p_1,x)L) \\ &+ (1-\epsilon)p'(p_0,x)u_2(\cdot) + (1-\epsilon)p(p_0,x)u_2'(-1-\alpha p'(p_1,x)L) = 0.\end{aligned}$$

Simplifying the equation yields

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

To understand the result, we first consider the case where $p(p_0,x) = p(p_1,x)$. In this case, the only difference is the term $1-\epsilon$ when compared with the standard result without additional features. The right-hand side is larger than the standard result 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. Next, we consider the change in the probability that $p(p_1,x) > p(p_0,x)$. This change does not affect the left-hand side but the right-hand side. Due to the convexity of $p(x)$, we have $-p'(p_0,x) < -p'(p_1,x)$. The denominator is also lower for $p'(p_0,x)$ as $u_1' < u_2'$ and ϵ is sufficiently small. Therefore, when delayed information is severe, the change in the nominator dominates the change in the denominator and the right-hand side is lower than before. This means that the optimal protection x becomes lower. Overall, there are different effects from delayed information and tail risk for the optimal self-protection level. The reason is that delayed information leads to a lower estimation of risk by the firm and thus invests less than optimal for self-protection, while tail risk reduces the value of insurance and incentivizes the firm to increase self-protection.

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 Swiss Re (2022) that more than 90% of the cyber loss is not covered by insurance. Previous work such as Böhme and Schwartz (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.

7. Discussion and conclusion

This paper leverages data from three databases to probe the fundamental empirical properties of cyber risk, an emergent and critical risk category. 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 malicious cyber risk has been growing exponentially in the past two decades and its financial loss distribution has shifted to the right, resulting in higher severity per event. Other cyber risk categories such as negligent incidents have lower growth rates in frequency. Moreover, we explore the dynamics of tail risk and find that all categories of cyber losses exhibit heavy-tailedness and this pattern remains consistent over time. Based on these results, we incorporate two empirical features (delayed information and tail risk) into the risk management framework with self-protection and insurance. The results indicate that these additional features lead to significantly lower insurance demand and potentially lower self-protection, thus increasing the risk level of the firm.

Our results have several implications for cyber risk management. First, we show that the cyber risk landscape is in constant flux and the time trends differ across various categories of cyber risk. It is of paramount importance to diligently monitor these changes to ensure updated and timely responses. As shown in our model, delayed information can induce the underestimation of threat level and thus suboptimal investments in security. A centralized cyber research center could facilitate real-time threat analysis and disseminate this information to organizations, thereby incentivizing better security investments (Zhuang et al., 2020). Second, the analysis indicates the heavy-tailedness of cyber losses, implying a greater probability of extreme losses compared to normal distributions. More broadly, heavy-tailed distributions are prevalent in economics, finance, and other areas (Ibragimov et al., 2015). This emphasizes the importance of studying the tail properties of cyber losses and tailoring the security investments accordingly to reduce the impacts of extreme scenarios. Third, our model examines the option of cyber insurance and finds that insufficient coverage reduces its demand. This implies that firms are susceptible to potential financial consequences arising from cyber incidents, particularly those of a severe nature. However, the insurers are reluctant to provide higher coverages due to the non-diversification trap from heavy-tailedness. Government intervention by providing backup to insurers, akin to a lender-of-last-resort in the banking sector, could alleviate this challenge (Drechsler et al., 2016).

Our work has several limitations, which might open avenues for future research. First, although we apply different statistical methods for correcting data biases, there is potential for other forms of bias to affect our findings. The bias arises partly from organizations' reluctance to disclose information about cyber incidents, and partly from the challenging detection and quantification of cyberattacks. Future studies would benefit from increased data sharing and standardization in cyber incident reporting. Second, although we connect our empirical findings with the theoretical framework, the link depends on several assumptions. For example, we hypothesize report delays could create an information gap between firms and insurers. While this assumption is plausible, it is not directly confirmed in our data. Future studies could empirically validate this link.

There are also other promising directions for future research. For example, one statistical aspect of cyber risk that is not emphasized in our paper is the correlation structure of cyber events, although there are several studies on this topic (Böhme and Kataria, 2006; Zhang et al., 2023), the dynamic trend of correlation structure is still under-explored empirically. Furthermore, given the evolving nature of cyber threats, future work could develop new frameworks for optimal security management, incorporating factors like information delay and risk uncertainty.

Acknowledgement

This work was supported by the Swiss National Science Foundation 100013_204381.

References

ABC News (2007). TJX data breach may involve 94 million credit cards. https://abcnews.go.com/Technology/story?id=3773782. Accessed January 23, 2024.

Accenture (2021). Cyber threat intelligence report. https://www.accenture.com/lu-en/insights/security/cyber-threat-intelligence-report-2021. Accessed January 23, 2024.

Aldasoro, I., Gambacorta, L., Giudici, P., and Leach, T. (2022). The drivers of cyber risk. *Journal of Financial Stability*, 60:100989.

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 January 23, 2024.

Anderson, R. and Moore, T. (2006). The economics of information security. *Science*, 314(5799):610–613.

Avanzi, B., Tan, X., Taylor, G., and Wong, B. (2023). Cyber insurance risk: Reporting delays, third-party cyber events, and changes in reporting propensity–an analysis using data breaches published by us state attorneys general. *arXiv preprint arXiv:2310.04786*.

Baranowski, R., Chen, Y., and 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., and Hillairet, C. (2021). Multivariate hawkes process for cyber insurance. *Annals of Actuarial Science*, 15(1):14–39.

Biener, C., Eling, M., and 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. and Kataria, G. (2006). Models and measures for correlation in cyber-insurance. *Workshop on the Economics of Information Security*, 2:3.

Böhme, R. and Schwartz, G. (2010). Modeling cyber-insurance: towards a unifying framework. *Workshop on the Economics of Information Security*, 1:3.

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

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

Cheung, K.-F. and Bell, M. G. (2021). Attacker–defender model against quantal response adversaries for cyber security in logistics management: An introductory study. *European Journal of Operational Research*, 291(2):471–481.

Dacorogna, M., Debbabi, N., and Kratz, M. (2023). Building up cyber resilience by better grasping cyber risk via a new algorithm for modelling heavy-tailed data. *European Journal of Operational Research*, 311(2):708–729.

Drechsler, I., Drechsel, T., Marques-Ibanez, D., and Schnabl, P. (2016). Who borrows from the lender of last resort? *Journal of Finance*, 71(5):1933–1974.

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

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

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

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

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

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

FBI (2022). Internet crime report. https://www.ic3.gov/. Accessed January 23, 2024.

Froot, K. A. and Stein, J. C. (1998). Risk management, capital budgeting, and capital structure policy for financial institutions: an integrated approach. *Journal of Financial Economics*, 47(1):55–82.

Gabaix, X. and 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. and 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., and Sohail, T. (2003). A framework for using insurance for cyber-risk management. *Communications of the ACM*, 46(3):81–85.

Hill, B. M. (1975). A simple general approach to inference about the tail of a distribution. *The Annals of Statistics*, 3(5):1163–1174.

Ibragimov, M., Ibragimov, R., and Walden, J. (2015). *Heavy-tailed Distributions and Robustness in Economics and Finance*, volume 214. Springer.

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

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

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

Kamiya, S., Kang, J.-K., Kim, J., Milidonis, A., and 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.

Keith, A. and Ahner, D. (2021). Counterfactual regret minimization for integrated cyber and air defense resource allocation. *European Journal of Operational Research*, 292(1):95–107.

Khouzani, M., Liu, Z., and Malacaria, P. (2019). Scalable min-max multi-objective cybersecurity optimisation over probabilistic attack graphs. *European Journal of Operational Research*, 278(3):894–903.

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

Maillart, T. and Sornette, D. (2010). Heavy-tailed distribution of cyber-risks. *The European Physical Journal B*, 75(3):357–364.

Malavasi, M., Peters, G. W., Shevchenko, P. V., Trück, S., Jang, J., and Sofronov, G. (2022). Cyber risk frequency, severity and insurance viability. *Insurance: Mathematics and Economics*, 106:90–114.

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

McAfee (2020). The hidden costs of cybercrime. https://www.mcafee.com/blogs/other-blogs/executive-perspectives/the-hidden-costs-of-cybercrime-on-government/. Accessed January 23, 2024.

Naghizadeh, P. and Liu, M. (2014). Voluntary participation in cyber-insurance markets. *Workshop on the Economics of Information Security*, 1:30.

Niu, Y. S., Hao, N., and Zhang, H. (2016). Multiple change-point detection: a selective overview. *Statistical Science*, 31(4):611–623.

Ossberger, J. (2020). Package 'tea'. https://cran.r-project.org/web/packages/ tea/index.html. Accessed January 23, 2024.

Reuters (2017). Yahoo says all three billion accounts hacked in 2013 data theft. https://www.reuters.com/article/us-yahoo-cyber-idUSKCN1C82O1. Accessed January 23, 2024.

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

Sangari, S., Dallal, E., and Whitman, M. (2023). Modeling reporting delays in cyber incidents: an industry-level comparison. *International Journal of Information Security*, 22(1):63–76.

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

Stoner, O. and Economou, T. (2020). Multivariate hierarchical frameworks for modeling delayed reporting in count data. *Biometrics*, 76(3):789–798.

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

Swiss Re (2022). Cyber insurance: strengthening resilience for the digital transformation. https://www.swissre.com/institute/research/topics-and-risk-dialogues/ digital-business-model-and-cyber-risk/cyber-insurance-strengthening-resilience.html. Accessed January 23, 2024.

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

Wang, Q.-H. and Kim, S. H. (2009a). Cyber attacks: Cross-country interdependence and enforcement. *Workshop on the Economics of Information Security*, 1:1–16.

Wang, Q.-H. and Kim, S. H. (2009b). Cyberattacks: does physical boundry matter? *ICIS 2009 Proceedings*, page 48.

Welburn, J., Grana, J., and Schwindt, K. (2023). Cyber deterrence with imperfect attribution and unverifiable signaling. *European Journal of Operational Research*, 306(3):1399–1416.

Wheatley, S., Hofmann, A., and 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., and 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. and Böhme, R. (2021). Sok: Quantifying cyber risk. In *2021 IEEE Symposium on Security and Privacy (SP)*, pages 211–228. IEEE.

Woods, D. W., Moore, T., and Simpson, A. C. (2021). The county fair cyber loss distribution: drawing inferences from insurance prices. *Digital Threats: Research and Practice*, 2(2):1–21.

Zhang, X., Xu, M., Su, J., and Zhao, P. (2023). Structural models for fog computing based internet of things architectures with insurance and risk management applications. *European Journal of Operational Research*, 305(3):1273–1291.

Zhao, X., Xue, L., and 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.

Zhuang, Y., Choi, Y., He, S., Leung, A. C. M., Lee, G. M., and Whinston, A. (2020). Understanding security vulnerability awareness, firm incentives, and ict development in pan-asia. *Journal of Management Information Systems*, 37(3):668–693.

Appendix to
The changing landscape of cyber risk: An empirical analysis of frequency, severity, and tail dynamics

Martin Eling[1a], Dingchen Ning[a], Rustam Ibragimov[b]

[a] *Institute of Insurance Economics, University of St. Gallen, St. Gallen, Switzerland*
[b] *Imperial College Business School, London, UK*

A. Categorization of cyber incidents for three databases

A.1. Risk type in Advisen

In the paper, we classify cyber events into four categories: malicious, negligent, privacy, and others. The risk types in Advisen are more granular and thus we aggregate them into these four categories. 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.[2] Finally, the incidents that do not belong to the categories above are classified as "others".

A.2. Risk type in PRC

In the PRC data, we rely on the variable "type of breach" for classification. The type of breach of each incident is indicated with a 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.)

[1]Corresponding author.
Email addresses: martin.eling@unisg.ch (M. Eling); i.rustam@imperial.ac.uk (R. Ibragimov); dingchen.ning@unisg.ch (D. Ning)

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

- 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 Advisen data, we allocate the incidents that belong to "HACK", "INSD", and "CARD" to the malicious category, and the incidents that belong to "DISC" to the negligent category, and the rest are categorized into "others". There is no "privacy" category in this database.

A.3. Risk type in SAS

The categorization of operational risk in SAS is based on the Basel categorization table from the Bank of International Settlement (BIS, 2001). There are three levels in this categorization: event risk type (level 1), sub-risk type (level 2), and activity (level 3). Level 1 and 2 in this table are at a high level and not suitable for the categorization of cyber risk. Therefore, in the first step, we allocate the level 3 activities to our risk categories as follows.

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

The "others" category: "Check kiting", "Non-physical damage abuse", "Theft", "Fire", "Natural catastrophe", "Violence against person", "Violence against property".[3]

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 category includes only "Internal Fraud" and "External Fraud", and the rest are allocated into the negligent category.

B. Report delay

B.1. Figures for time trend (Figure 1 and 2)

B.2. Alternative models for report delay

To show the method we use in the paper has the best performance for our case, we compare the in-sample performance of our method with two other models: a generalized linear model (GLM) (Salmon et al., 2015), and a generalized Dirichlet-multinomial survivor model (GDM survivor) (Stoner and Economou, 2020).[4] In this section, we first briefly introduce two other models.

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:

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

[3]The risk types above are not the complete list in the activity level from the Basel categorization table since not every type has cyber incidents.

[4]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 Poisson model (Renshaw and 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 (Mack, 1993). 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 scarcer than the information on the number of events, thus we only focus on the report delay issue for the frequency data.

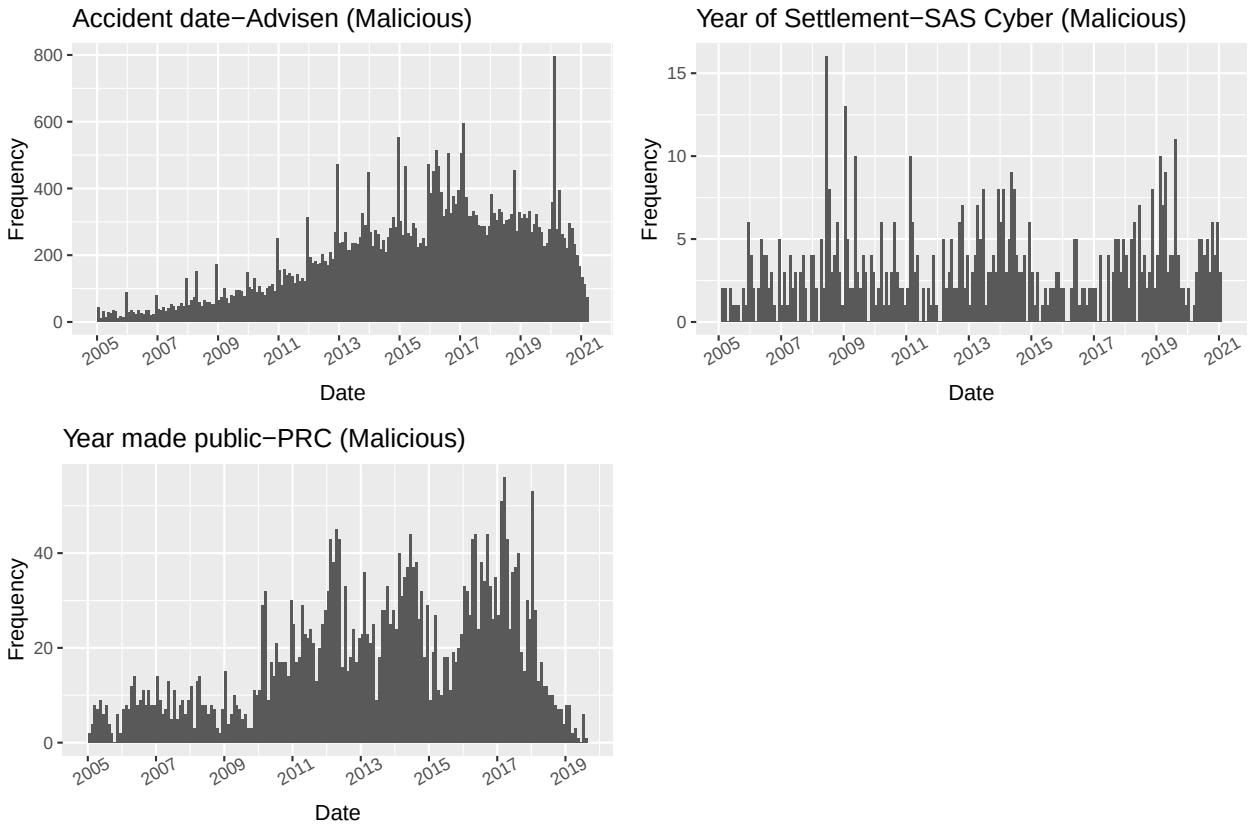

Figure 1: Different datasets of cyber risk

Note: 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.

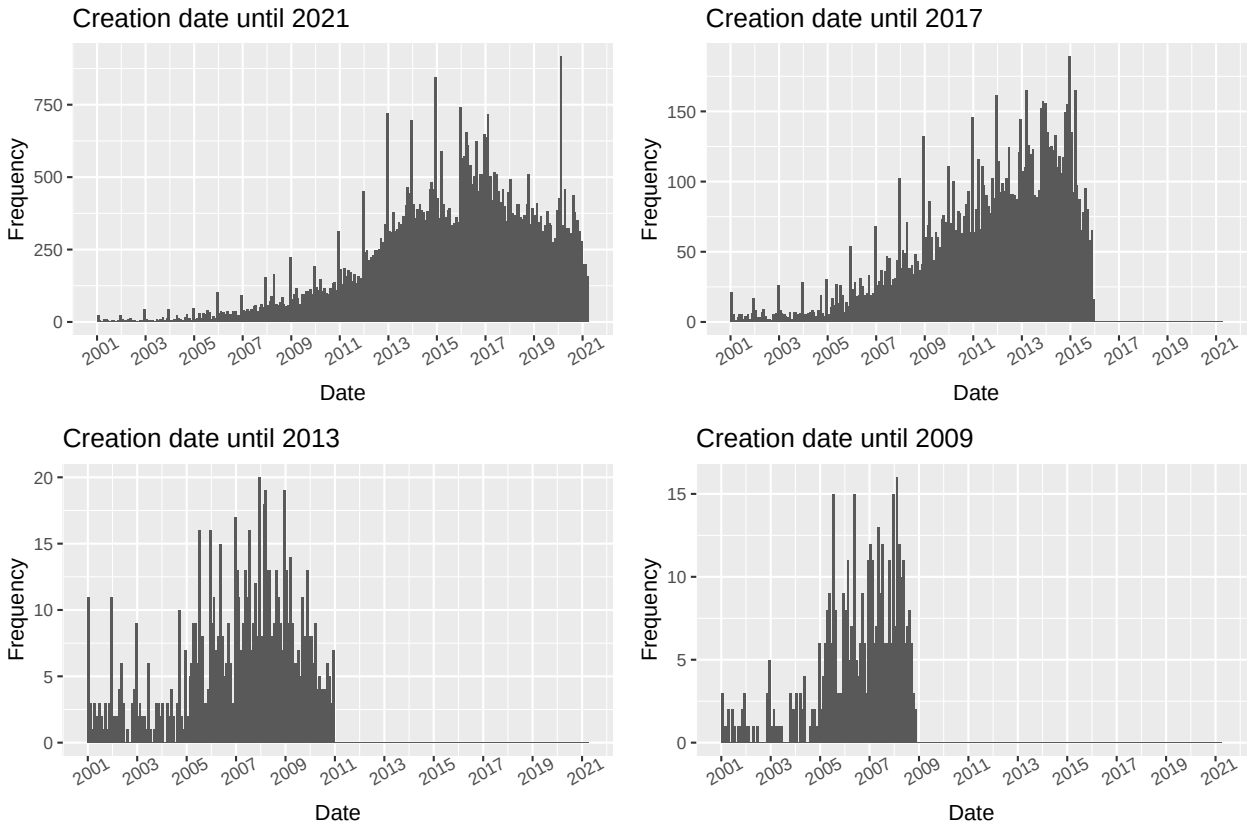

Figure 2: Different dates in Advisen (Malicious)

Note: This figure reports the monthly frequency of malicious cyber events in Advisen. The creation date is the date when the record of the cyber event was created in the database. For example, the creation date until 2021 means all the events that were recorded in the database before 2021 are included.

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.

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

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

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.

B.3. In-sample comparison

The data of Advisen contains multiple abnormal peaks due to inaccurate information. 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 correct this anomaly, we assume the date of cyber events without accurate time follows a normal distribution and 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 test which model performs the best 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 3 (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.

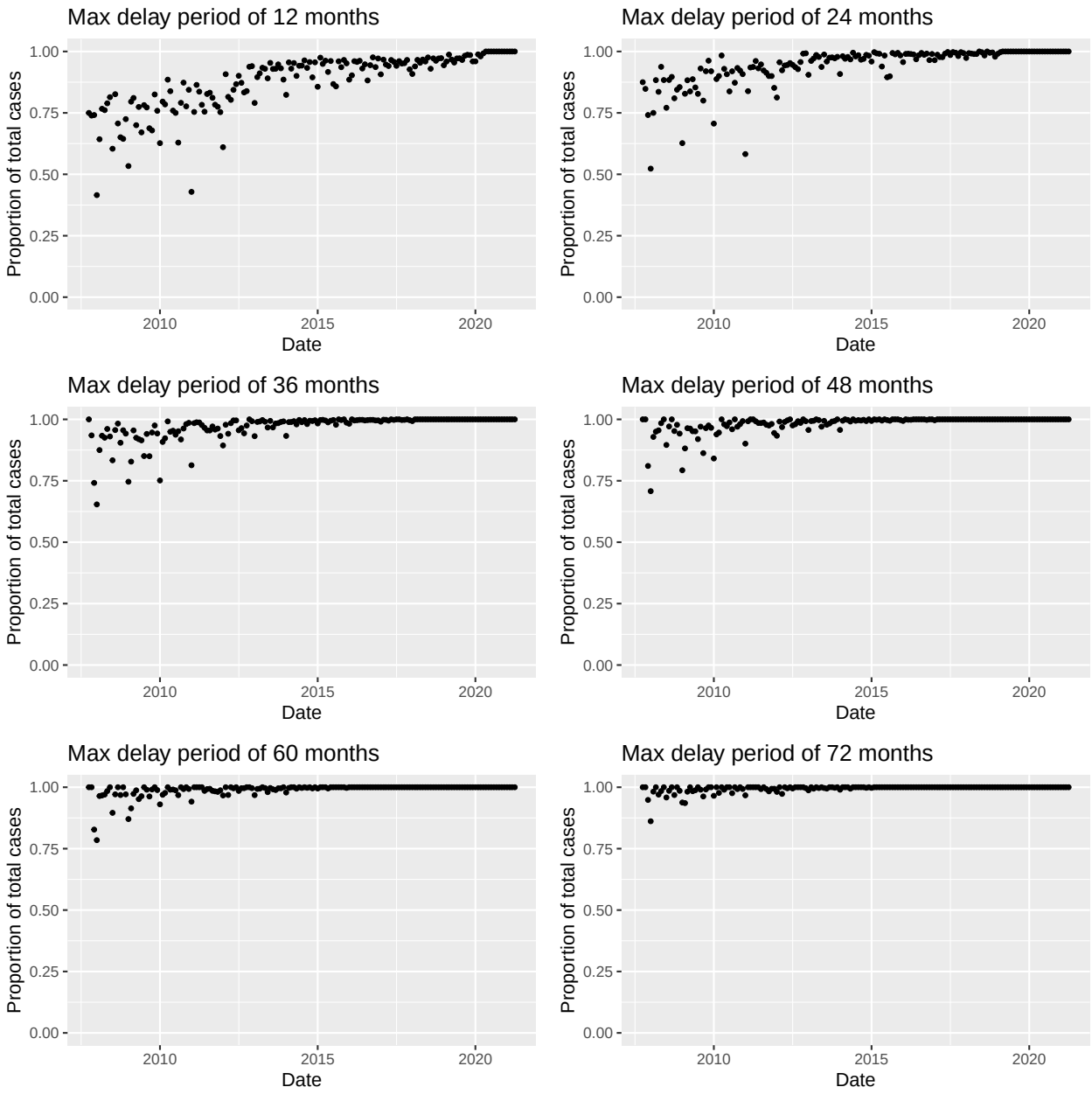

Figure 3: Cumulative proportion reported

Note: 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. However, the pattern across graphs shows how a longer period increases the percentage of reported cases.

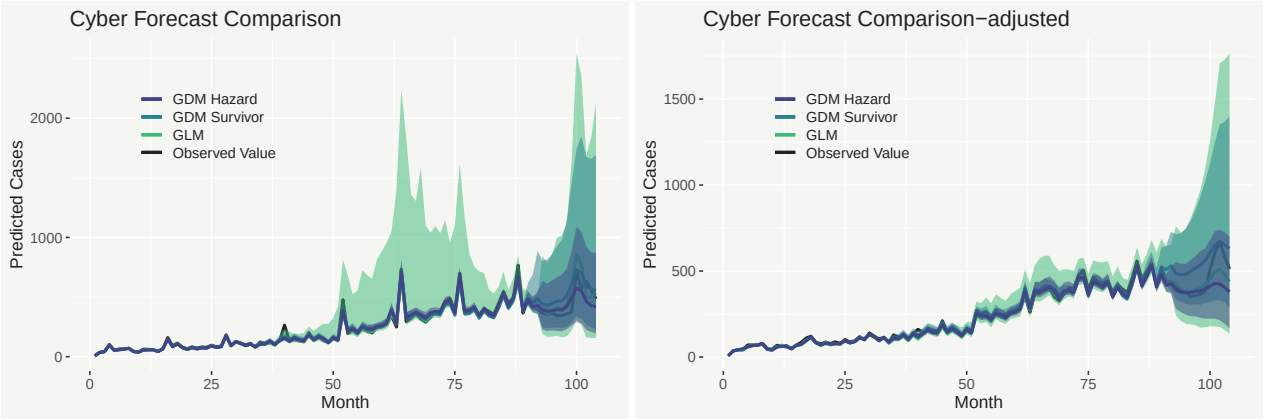

Figure 4: In-sample cyber forecast comparison

Note: 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.

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 4 shows the results of the median estimated number for original and adjusted data of malicious events, with 95% posterior predictive interval. Among the three models, GDM hazard has the most accurate confidence interval while GLM performs worst. Figure 5 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.

B.4. Two stages of data reporting

In general, the process of collecting data related to cyber risk can be divided into two stages. The first stage is from the event date to the date of the first notice. This period can be short for some categories of events, such as cyber extortion or malfunction of devices, which the victims would notice almost immediately. But for other categories 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 delay in this stage is primarily attributable to the data provider's efficiency. While in some instances the staff may update the data immediately, it is more common to encounter moderate delays. 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 began to collect data in 2007, the majority of their events were created in recent years, especially during 2016-2018.

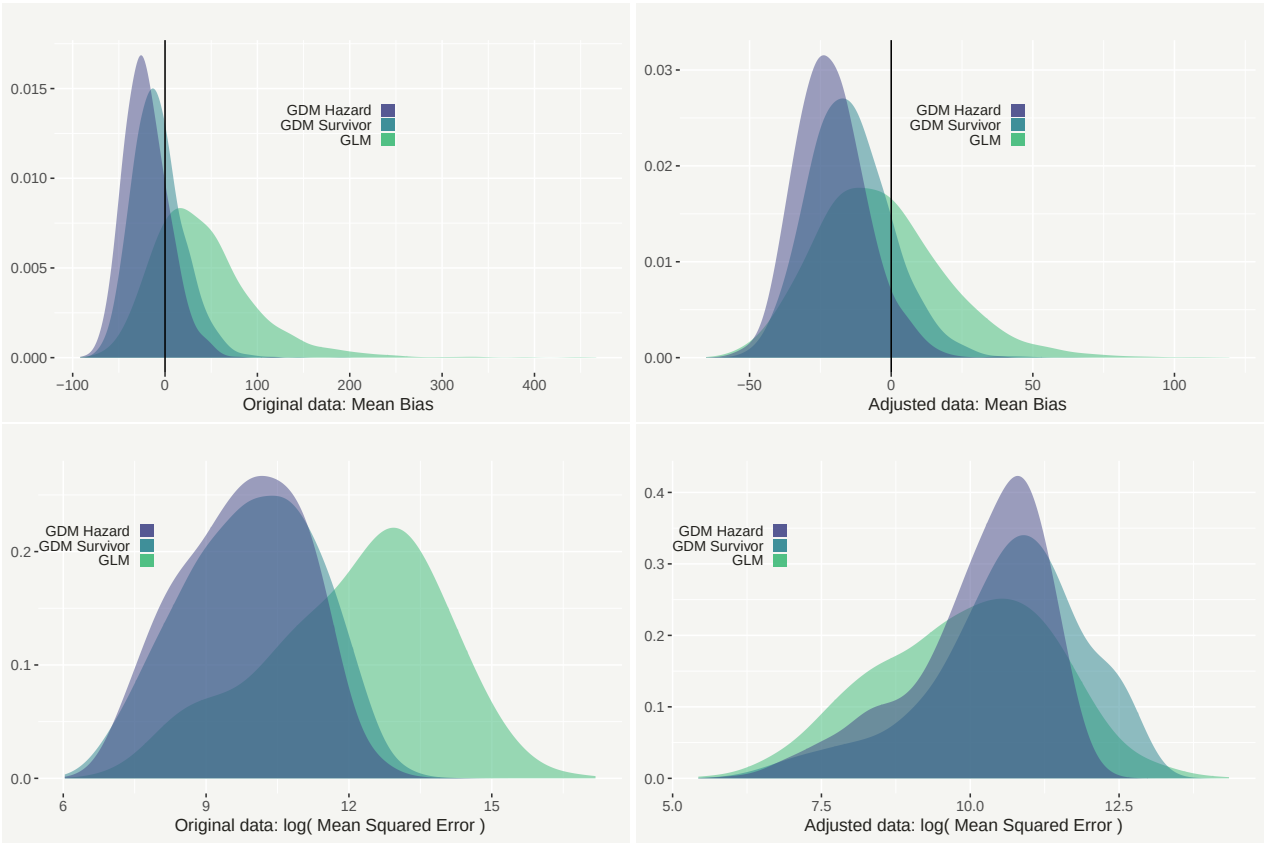

Figure 5: Covariance of Z

Note: 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.

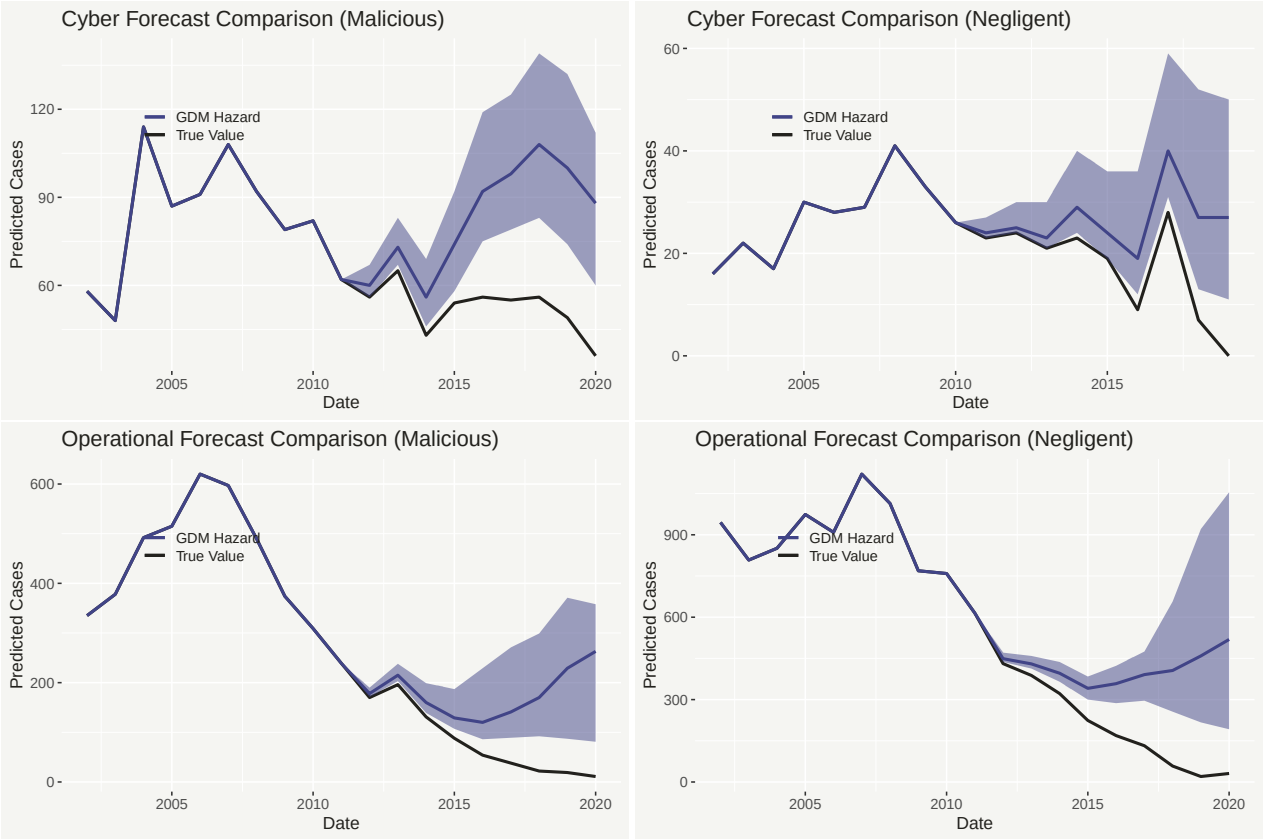

Figure 6: Bias correction for the SAS data

Note: 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.

B.5. Report delay in SAS

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 apply our method to the SAS dataset. However, since SAS data only have information at the yearly level about the accident date, we use this data as a robustness check only. As shown in Figure 6, 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 that there are more incidents of unintentional disclosure in Advisen which are on the rise and affect the overall trend. In addition, we find the trend of operational risk is decreasing even after the bias correction, which is in sharp contrast to the emerging cyber risk.

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 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.

C. Time dynamics of cyber frequency: Comparison of methods

In this section, 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 trend 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 and 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 dynamic programming. The second approach is the trend filtering from Kim et al. (2009), implemented in R by Arnold and Tibshirani (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 7 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 provides similar results while the TF approach is more sensitive and detects more change points.[5] Overall, all the results are consistent with respect to the location of change points but different for the number of change points. This provides validation for the method we use in the main analysis and the conclusions we draw from the results.

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

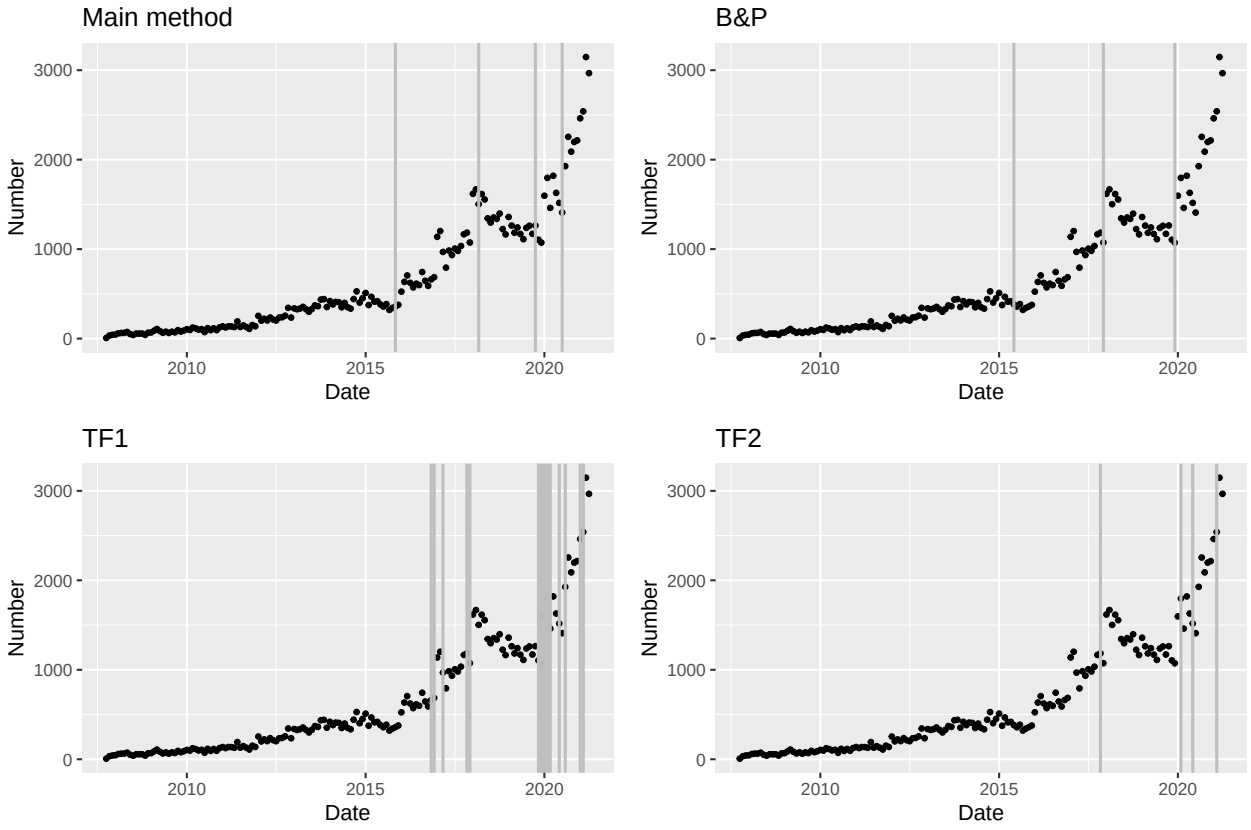

Figure 7: Comparison of methods for change points of frequency

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

D. Optimal threshold selection for tail index estimation

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 1 reports the results when the sample size equals 100, and Table 2 reports the results when $N = 500$. In both cases, two methods provide better results: "dAMSE" (Caeiro and Gomes, 2015) and "hall" (Hall, 1990). The first method is based on the concept of minimizing the average mean squared error (AMSE) 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 1 and 2. In other words, these two methods tend to estimate lower tail indices than 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 (Gabaix and Ibragimov, 2011) that is specially adjusted for small sample bias. Table 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.

E. Time dynamics of loss severity: additional results

In the main paper, we only report the comparison of loss distributions of part of our data due to the extra requirements of firm information when adjusting weights. To provide the full picture of our data, we present the results of all categories of cyber losses without adjusting the weights. Therefore, we need to be cautious when interpreting the results.

Table 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 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 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 1 and 2 as we run another simulation of 10,000 times for this table.

Figure 8 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 (consistent with the main results), while the distribution for the negligent cases is shifting to the left. In Figure 9, the distribution of the number of records over the years is presented. The malicious cases appear to lead to higher losses than the negligent cases.

Figure 10 compares the average distribution of financial loss before and after the change point. The left panel shows the distributions for malicious cyber risk and the results are consistent with the main results. The only difference is that the change point is different for the SAS data. 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 the change point. In general, the results in Figure 10 show the financial loss distributions from malicious cases and negligent cases are moving in the opposite direction. However, as we cannot match the firms in the negligent sample, it is difficult to estimate whether the shift is driven by selection bias or other factors.

Figure 11 presents the results for the distributions of the number of records or affected counts. For the Advisen data, we can find that the distribution is shifting to the left after the change point, which is consistent with the main results, although the exact shape of the distribution is different. For the PRC data, the detection method shows no significant change. The density plots in the figure use the most probable change point, with the distributions converging towards the center following this potential change. Again, this result might be driven by potential bias in the data.

F. Time dynamics of tail risk: additional results

To better understand the change in the tail index, we plot the distributions of the estimated tail index before and after the change point. As the tail index estimator follows normal distribution (Gabaix and Ibragimov, 2011), we use the estimated index and the standard deviation to simulate the whole distribution of the tail index at each time point. Then we combine all the simulation results before the change point to compute the empirical density of the sample. Similarly, we can get the density distribution for the sample after the change point.

Figure 12 presents the results for financial loss data. Consistent with the graphical illustration in the paper, the tail index for malicious cyber losses has a higher mean and variance after the change point. More specifically, the empirical average increases from 0.61 to 0.86 for Advisen data and from 0.70 to 1.09 for SAS data. However, the results for negligent cyber losses are mixed and the estimation from SAS yields a higher value than the one from Advisen. This might be related to the sample composition issue we discuss in the descriptive statistics section. As there are many more small losses in Advisen than in SAS, the discrepancy between small and large losses contributes to the pronounced skewness of the distribution and also drives up the heaviness of the tail. In addition, negligent cases have fewer observations than malicious cases in both datasets, which can lead to less accurate and divergent estimation. The comparison of the tail index for non-financial losses

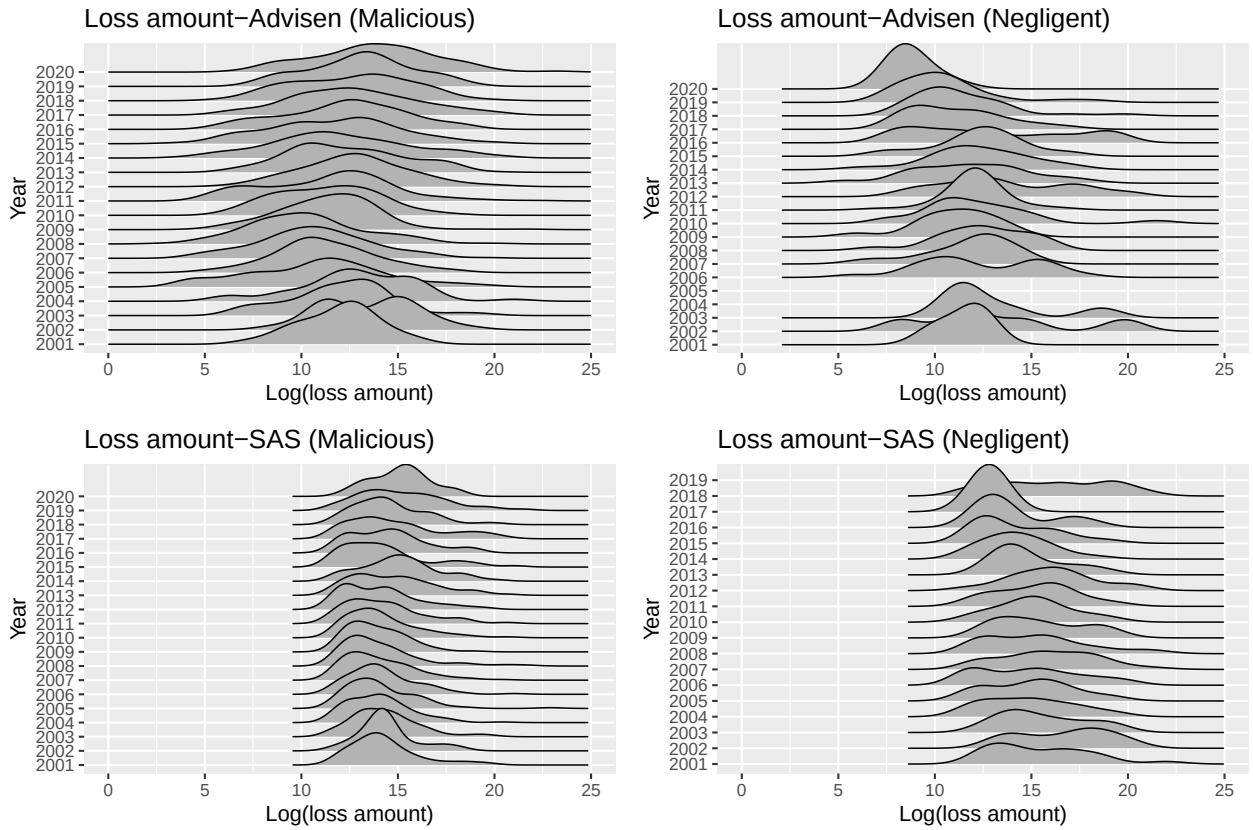

Figure 8: Dynamics of distributions (loss amount)

Note: 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. As the events in SAS have losses of at least $100,000, the distribution is located to the right compared to the distribution in Advisen.

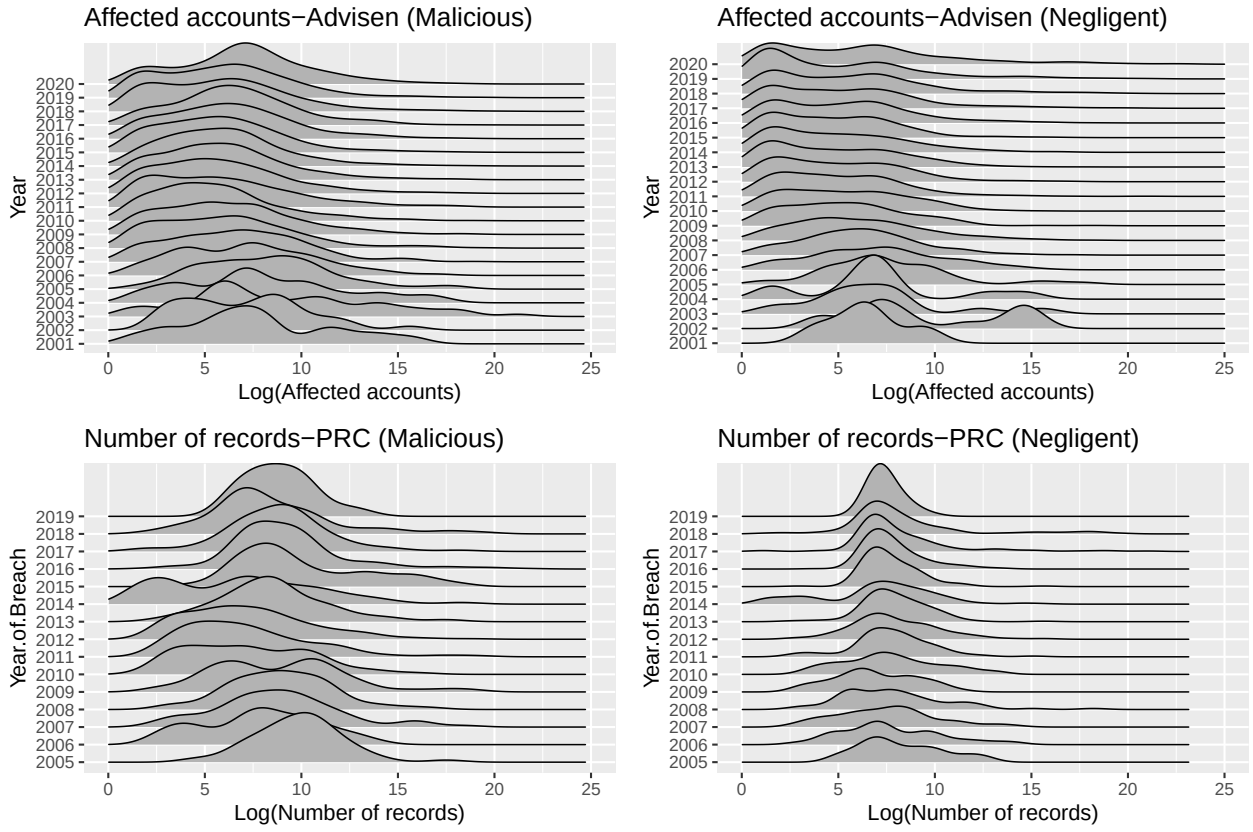

Figure 9: Dynamics of distributions (number of records)

Note: 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.

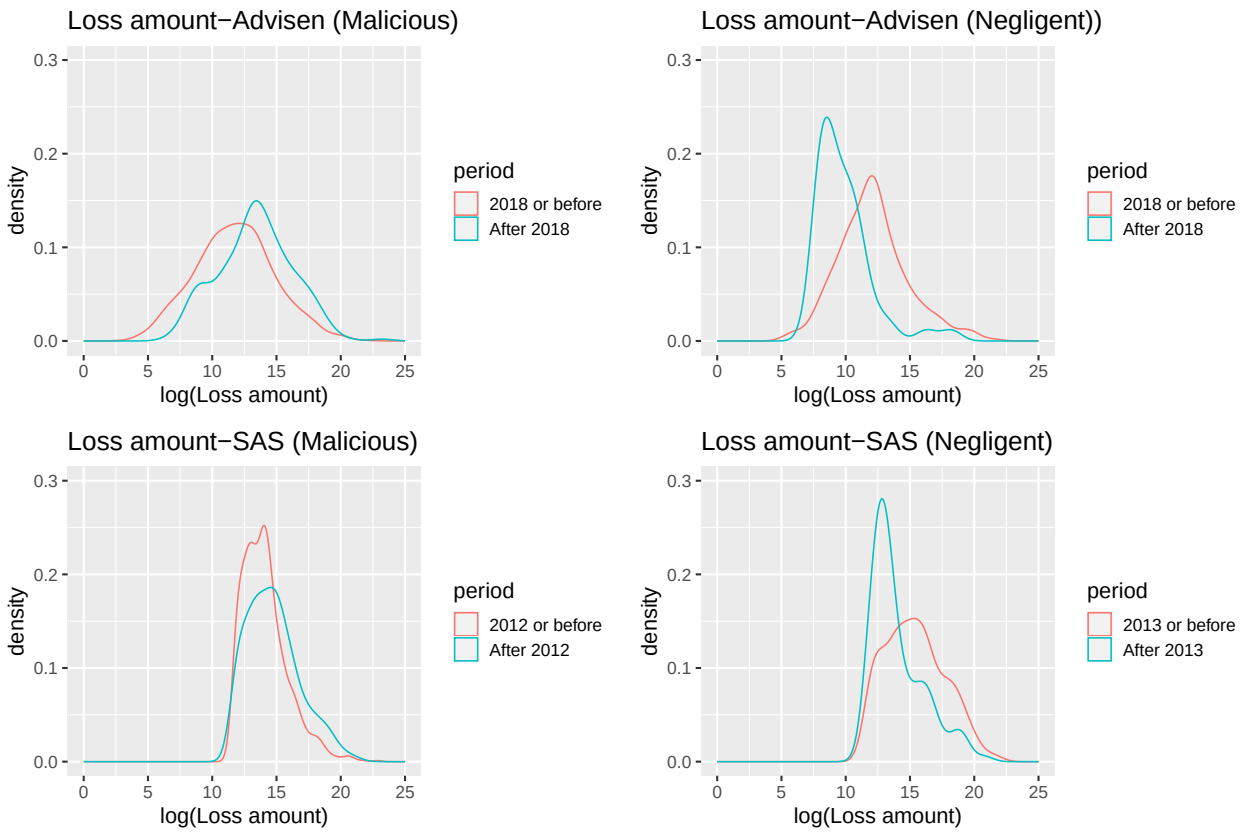

Figure 10: Change points of distributions (loss amount)

Note: 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.

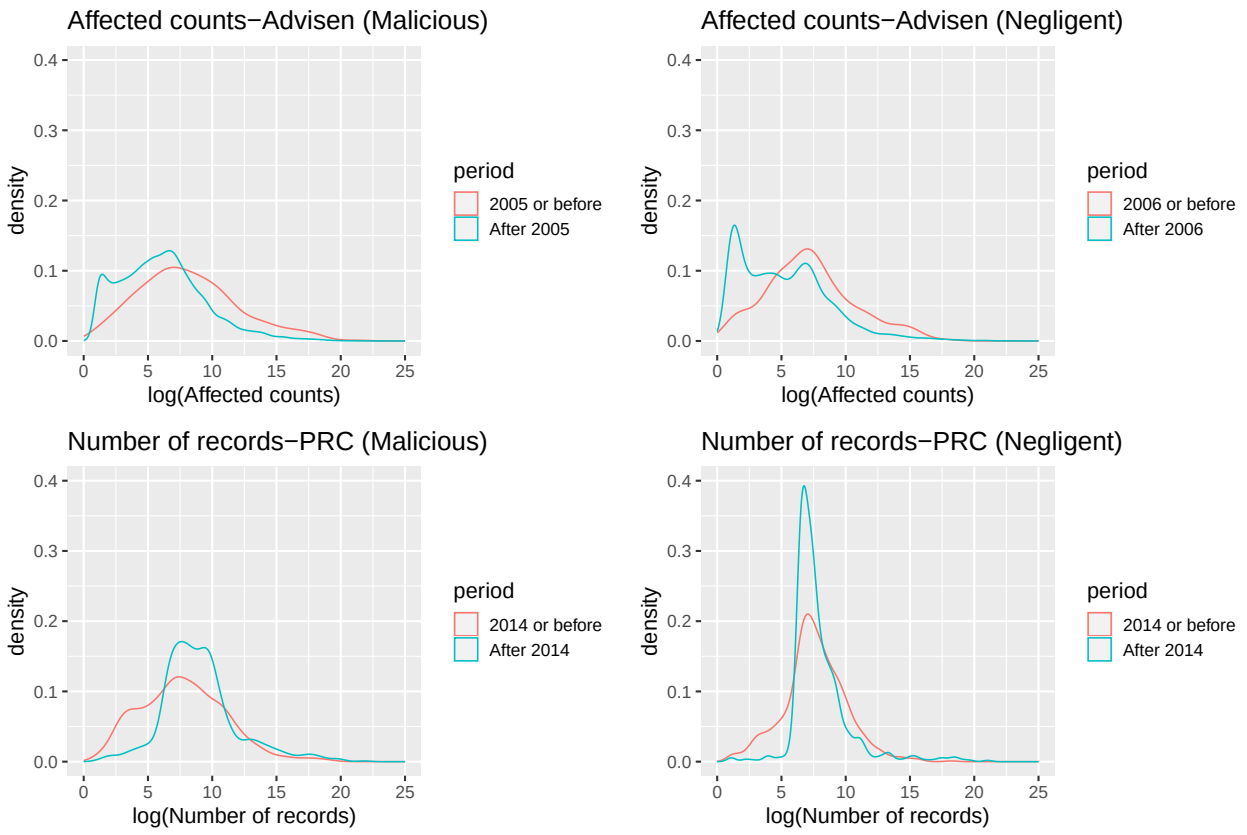

Figure 11: Change points of distributions (number of records)

Note: 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

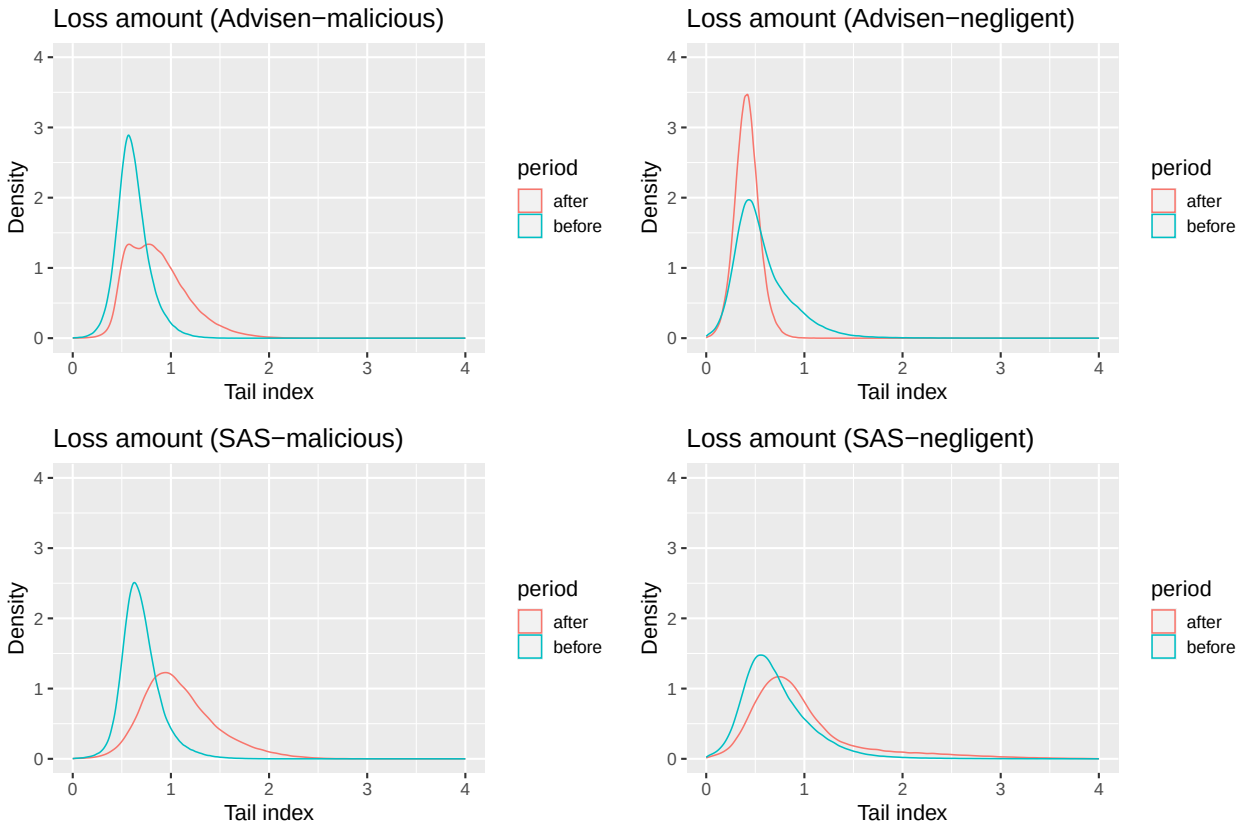

Figure 12: Comparison of tail index distribution (loss amount)

Note: This figure presents the distribution of the tail index before and after the detected change point for the estimation of monetary losses in Advisen and SAS.

is shown in Figure 13. The results for malicious cases are mixed from Advisen and PRC, but the estimation predominantly falls in the range of 0.4 to 0.5. The pattern for negligent cases is clearer, as there is a common trend of declining tail index after the change point. The empirical average drops to 0.34 in Advisen and 0.37 in PRC. The results also indicate that negligent losses are more heavy-tailed than malicious losses. The reason that negligent incidents lead to more extreme losses might 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. This is consistent with the extensive literature that considers the employees' misbehaviors as a key challenge to information security (Bulgurcu et al., 2010).

References

Arnold, T. and Tibshirani, R. (2020). Package 'genlasso'. *Statistics*, 39(3):1335–1371.

Bai, J. and Perron, P. (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics*, 18(1):1–22.

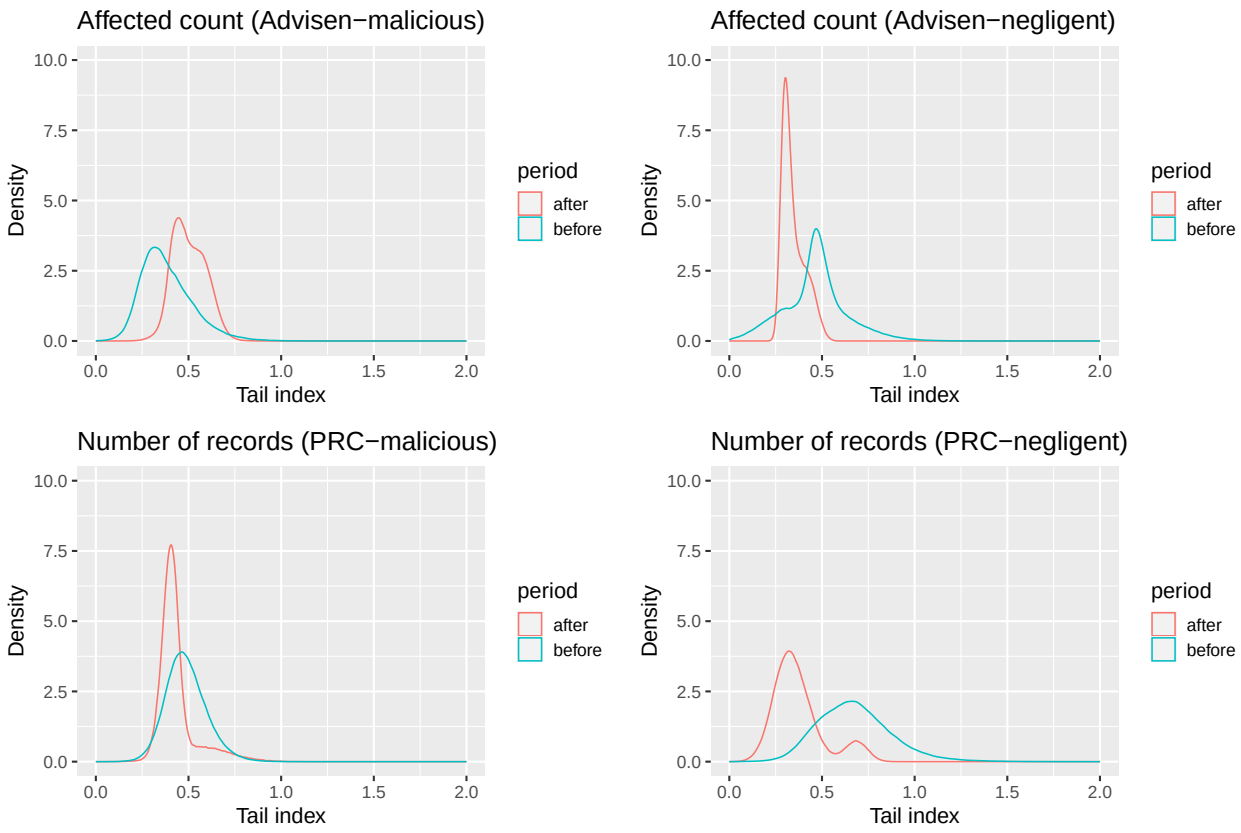

Figure 13: Comparison of tail index distribution (number of records)

Note: This figure presents the distribution of the tail index before and after the detected change point for the estimation of non-monetary losses in Advisen and PRC.

Baranowski, R., Chen, Y., and 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.

BIS (2001). Operational risk loss data. https://www.bis.org/bcbs/qisoprisknote.pdf. Accessed January 23, 2024.

Bulgurcu, B., Cavusoglu, H., and Benbasat, I. (2010). Information security policy compliance: an empirical study of rationality-based beliefs and information security awareness. *MIS Quarterly*, pages 523–548.

Caeiro, F. and Gomes, M. I. (2015). Threshold selection in extreme value analysis. *Extreme Value Modeling and Risk Analysis: Methods and Applications*, pages 69–82.

Gabaix, X. and 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.

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.

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

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.

Ossberger, J. (2020). Package 'tea'. https://cran.r-project.org/web/packages/ tea/index.html. Accessed January 23, 2024.

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

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

Stoner, O. and Economou, T. (2020). Multivariate hierarchical frameworks for modeling delayed reporting in count data. *Biometrics*, 76(3):789–798.

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

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

Zeileis, A., Leisch, F., Hornik, K., Kleiber, C., and Hansen, B. (2022). Package 'strucchange': Testing, monitoring, and dating structural changes. https://CRAN.R-project.org/ package=strucchange. Accessed January 23, 2024.