Communications for Statistical Applications and Methods
2023, Vol. 30, No. 1, 75–94

https://doi.org/10.29220/CSAM.2023.30.1.075
Print ISSN 2287-7843 / Online ISSN 2383-4757

Cyber risk measurement via loss distribution approach and GARCH model

Sanghee Kim [a], Seongjoo Song [1,a]

[a]Department of Statistics, Korea University, Korea

Abstract

The growing trend of cyber risk has put forward the importance of cyber risk management. Cyber risk is defined as an accidental or intentional risk related to information and technology assets. Although cyber risk is a subset of operational risk, it is reported to be handled differently from operational risk due to its different features of the loss distribution. In this study, we aim to detect the characteristics of cyber loss and find a suitable model by measuring value at risk (VaR). We use the loss distribution approach (LDA) and the time series model to describe cyber losses of financial and non-financial business sectors, provided in SAS® OpRisk Global Data. Peaks over threshold (POT) method is also incorporated to improve the risk measurement. For the financial sector, the LDA and GARCH model with POT perform better than those without POT, respectively. The same result is obtained for the non-financial sector, although the differences are not significant. We also build a two-dimensional model reflecting the dependence structure between financial and non-financial sectors through a bivariate copula and check the model adequacy through VaR.

Keywords: cyber risk, value-at-risk, loss distribution approach, GARCH, extreme value theory, copula

1. Introduction

Cyber risk, a loss or damage of a firm's assets related to information and technology, is gradually emerging as a significant problem. Industries attain confidential customer information, whether in the financial sphere or not, so its loss critically damages both companies and consumers. The recent Basel committee's newsletter (Basel Committee on Banking Supervision, 2021) indicated that more cyber risk management tools are required for worldwide companies to be resilient from cyber threats. Cyber risk is generally considered as a part of operational risk, which is defined as loss arising from unauthorized, deliberate, or accidental activities (Chernobai *et al*., 2008). According to Cebula and Young (2010), cyber risk includes losses related to information technology (IT) assets resulting from human activities, disasters, and other internal and external events. This extent illustrates that operational risk encompasses cyber risk. Thus, studies such as Biener *et al*. (2015) and Eling and Wirfs (2015) utilized the operational risk dataset and employed the methods generally used for analyzing

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MIST) (No. 2021R1F1A1048986) and by a Korea University Grant (K2207831).

This manuscript is based on the first author, Sanghee Kim's Master thesis from Korea University.

[1] Corresponding author: Department of Statistics, Korea University, 145 Anam-ro, Seongbuk-gu, Seoul 02841, Korea. E-mail: sjsong@korea.ac.kr

Published 31 January 2023 / journal homepage: http://csam.or.kr

operational risk to analyze cyber risk. However, as Eling and Wirfs (2015) indicated, cyber losses show different characteristics from operational losses, as also shown in Section 3 in this paper.

Many existing studies on cyber risk used the privacy rights clearinghouse (PRC) dataset, which started to report data-breach cases publicly since 2005. PRC data only contains information about data-breaches and does not include specific loss amounts in dolloars. Hence, it is not possible to model the severity of data-breaches with PRC data. To solve this problem, Edwards *et al*. (2015) transformed the number of records breached to loss amount by the equation offered by Jacobs (2014) and analyzed data-breach cases through the loss distribution approach (LDA). Even though the actual loss is unknown, Edwards *et al*. (2015) said that a lognormal distribution adequately explained the transformed loss amounts, and a negative binomial distribution explained the number of data-breach cases. Later, due to the limitation to get a comprehensive analysis of cyber risk with PRC data, Eling and Wirfs (2019) practiced the LDA with a larger operational risk database–SAS® OpRisk Global Data. According to SAS product description (SAS, Retrieved June 20, 2021), SAS® OpRisk Global Data is the world's largest and most accurate operational risk dataset. The database contains no missing values of loss amounts, provides a detailed description of the loss since it is regularly updated, and includes every type of cyber loss, including data-breaches. In this paper, we also utilize SAS® OpRisk Global Data with more recent data than Eling and Wirfs (2019) to offer a broader sense of cyber losses reflecting the current trend.

The Basel Committee on Banking Supervision (2006) recommended that banks and financial institutions review potential operational losses through an internal model, which is fitted only with the firm's own data. However, due to the lack of extreme cases in the internal database, the external data, which includes the losses from other firms, could be used. If we form a dataset by extracting cyber loss cases from SAS® OpRisk Global Data, we could consider the dataset as the external data. In this context, what we aim in this paper is to provide a suitable analysis of cyber losses for firms that would utilize external data.

Our study focuses on modeling the monthly cyber loss of financial and non-financial business sectors. Cyber losses of firms in the financial category tend to be larger than those in the non-financial category. This feature is mentioned in Eling and Wirfs (2015) and can also be seen in Section 3. We first use the loss distribution approach (LDA) described in Aue *et al*. (2007) and Eling and Wirfs (2015) to model monthly cyber losses, extracted from SAS® OpRisk Global Data. In addition to LDA which regards each loss as independent, we also consider generalized autoregressive conditional heteroskedasticity (GARCH) model for reflecting the time-dependence of monthly cyber losses. We follow the two-step method suggested in McNeil and Frey (2000), as described in Section 2. We fit four models to financial and non-financial business sectors–LDA without POT, LDA with POT (LDA-POT), GARCH without POT, and GARCH with POT (GARCH-POT). Then, we measure cyber risk through value at risk (VaR) from each model. VaR is one of the most common measures for the operational risk and through VaR, we can suggest the minimum capital amount that each firm should hold to forestall future losses. By comparing the violation rates of VaR from models with or without POT, we can see the effectiveness of POT in measuring cyber risk. And we determine a suitable model for modeling cyber losses in financial and non-financial sectors using the violation rates of VaR. Note that we use the skewed-t distribution as the innovation distribution when practicing GARCH because it would describe tail heaviness more accurately than the ordinary t-distribution.

Besides, we model the joint distribution of cyber losses from financial and non-financial sectors using copulas upon LDA. When we want to figure out the level of cyber risk in a portfolio consisting of firms from both of financial and non-financial sectors, joint distributions incorporating the correlation between sectors would be advantageous. We fit a bivariate copula model to monthly cyber losses

based on LDA approach, using t-copula and Gaussian copula. These copulas are widely used ones in financial contexts, as mentioned in, for instance, Di Clemente *et al*. (2004).

The remainder of the paper is organized as follows. Section 2 introduces methods and models we use. Section 3 describes data used and procedures of model selections. VaR and backtesting results are presented in Section 4. Then, we conclude in Section 5.

2. Methods and models

2.1. Value at risk

Value at risk is the maximum amount of loss that can occur in a specified time period with a given confidence level. It is one of the common methods to quantify financial risks, providing a sense of the worst scenario. In the operational risk context, the Basel committee suggests 99.9% and the period of one year as a suitable confidence level and time horizon. For a given confidence level of α, VaR for a specified time horizon is defined as

$$\mathrm{VaR}_\alpha = \inf\{L : P(\mathrm{Loss} > L) \leq 1 - \alpha\}. \tag{2.1}$$

Detailed steps of calculating VaR will be presented in Section 4.

2.2. Peaks over threshold (POT)

POT is a widely used technique in financial modeling because it nicely captures the heavy-tailed characteristic. It is based on Balkema and de Haan (1974) and proposed by Pickands (1975), suggesting that the upper tail of a distribution be modeled through generalized Pareto distribution (GPD). Note that when we apply POT to risk analysis, we denote losses that exceed a sufficiently high threshold to "tail" and other losses to "body" of the distribution.

Suppose that $X_1, X_2, \ldots, X_n$ are independent and identically distributed random variables with the cumulative distribution function (CDF) F. The CDF of the excess loss $X - u \mid X > u$ that exceeds the threshold u can be written as

$$\begin{aligned} F_u(y) &= P(X - u \leq y \mid X > u) \\ &= \frac{F(y+u) - F(u)}{1 - F(u)}. \end{aligned}$$

From the CDF of the excess loss, we can obtain the CDF of X with $x = y + u$ as

$$F(x) = F_u(y)[1 - F(u)] + F(u). \tag{2.2}$$

Based on the theorem Balkema and de Haan (1974) and Pickands (1975) suggested, Embrechts *et al*. (2013) explained that GPD is an appropriate approximation of F_u for large u since we can find some positive function β for a large value u. The following Equation (2.3) shows this relationship.

$$\lim_{u \to x_F} \sup_{0 < x < x_F - u} \left| F_u(x) - G_{\xi,\beta(u)}(x) \right| = 0. \tag{2.3}$$

The CDF of GPD is defined as

$$G_{\xi,\beta}(y) = \begin{cases} 1 - \left(1 + \dfrac{\xi y}{\beta}\right)^{-\frac{1}{\xi}} & \text{if } \xi \neq 0, \\ 1 - \exp\left(-\frac{y}{\beta}\right) & \text{if } \xi = 0, \end{cases} \tag{2.4}$$

$\beta > 0$, $y > 0$ when $\xi \geq 0$, and $0 \geq y \leq -\beta/\xi$ when $\xi < 0$ where ξ is the shape parameter and β is the scale parameter of GPD. We say that the distribution has a heavy tail when $\xi > 0$. For more details on the definition and calculation of POT, refer to Embrechts *et al.* (2013), Beirlant *et al.* (2004) and Gilli (2006).

Practically, risk quantification depends on the choice of the threshold with POT. According to Rydman (2018), a commonly selected threshold is the 90^{th} percentile when the whole data is large enough. The rule of thumbs for the number of exceedances k are $k = \sqrt{N}$ or $k = N^{2/3}/\log(\log(N))$ where N is the total number of observations. According to McNeil and Frey (2000), GPD estimators would have low variances when a sufficiently large number of exceedances exist over the threshold k. Through the simulation study of choosing a threshold value, McNeil and Frey (2000) concluded that setting k as 100 would result in accurate tail estimation in the present application.

We can also use a graphical tool for choosing the threshold. The mean excess plot is a commonly used for this purpose, as in Rydman (2018). The mean excess function of a random variable $X \sim \text{GPD}_{\xi,\beta}$ with threshold u is the conditional expectation of the excess loss as

$$M(u) = E(X - u \mid X > u). \tag{2.5}$$

Since $E(X) = \beta/(1-\xi)$, $M(u)$ can be expressed as

$$M(u) = \frac{\beta}{1-\xi} + \frac{\xi}{1-\xi}u, \tag{2.6}$$

when $0 \leq u < \infty$ for $0 \leq \xi < 1$, and $0 \leq u \leq -\beta/\xi$ for $\xi < 0$. It is easy to see that $M(u)$ must be linear in u for the parts that follow GPD. Therefore, the starting point of the increasing linear slope in the mean excess plot would be a reasonable choice for the threshold u. The details of the definition and computation of mean excess function is presented in the introduction of Ghosh *et al.* (2010) and in Section 4.3 of Coles *et al.* (2001).

As in McNeil and Frey (2000), we estimate $F(u)$ by $(N - N_u)/N$ with N, the number of total observations and N_u, the number of exceedances over the threshold u. Then (2.2) can be written with the maximum likelihood estimators of GPD parameters ξ and β in (2.4) as

$$\begin{aligned}
\widehat{F(x)} &= G_{\hat{\xi},\hat{\beta}}(x-u)\left(1 - \frac{N-N_u}{N}\right) + \frac{N-N_u}{N} \\
&= \left[1 - \left(1 + \frac{\hat{\xi}(x-u)}{\hat{\beta}}\right)^{-\frac{1}{\hat{\xi}}}\right]\frac{N_u}{N} + \frac{N-N_u}{N} \\
&= 1 - \frac{N_u}{N}\left[1 + \frac{\hat{\xi}(x-u)}{\hat{\beta}}\right]^{-\frac{1}{\hat{\xi}}},
\end{aligned}$$

for $x > u$, assuming $\xi \neq 0$. Setting $\widehat{F(x)}$ to be q for $q > (N - N_u)/N$, we define the the tail estimator $\hat{x}_q$ as

$$\hat{x}_q = u + \frac{\hat{\beta}}{\hat{\xi}}\left(\left(\frac{1-q}{N_u/N}\right)^{-\hat{\xi}} - 1\right). \tag{2.7}$$

We use this tail estimator in calculating VaR for distributions with POT in Section 4.

2.3. Loss distribution approach (LDA)

LDA is one of the most common risk-quantification methods to model operational losses. LDA expresses the aggregated loss L for a given time period as

$$L = \sum_{i=1}^{N} X_i \,, \tag{2.8}$$

where N is the frequency of losses and $X_1, X_2, \ldots, X_N$ are independent individual losses. In the operational risk context, X_i's are positive continuous random variables that represent severity of each loss. Edwards *et al.* (2016) and Eling and Wirfs (2015) fitted negative binomial distribution and Poisson distribution to the frequency of cyber losses, respectively. They modeled the loss severity by various continuous distributions including lognormal and normal distributions. Following their studies, we also compare negative binomial and Poisson distributions for the loss frequency in Section 3. For the loss severity, we fit various continuous distributions that are commonly employed in actuarial contexts such as exponential, lognormal, normal, skewed-normal, skewed-t, logistic, GPD, and normal inverse Gaussian (NIG) distributions.

2.4. generalized autoregressive conditional heteroskedasticity model

Engle (1982) first proposed autoregressive conditional heteroskedasticity (ARCH) model, which explains time-varying volatility and volatility clustering in asset returns. ARCH(p) models the innovation term, a_t of a time series as follows.

$$a_t = \sigma_t \epsilon_t, \quad \sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i a_{t-i}^2 \,, \tag{2.9}$$

where $\alpha_0 > 0$, $\alpha_i \geq 0$, $i = 1, \ldots, p$ and ϵ_t is a white noise with zero mean and unit variance. As seen from the model, conditional variance σ_t^2 is positively correlated with the squared error terms. Hence, if the past innovations a_{t-i}^2 are large, then σ_t^2 becomes large, exhibiting volatility clustering. According to Cont (2007), in the financial context, volatility clustering means that large price change tends to cluster since the same movement continues for a while. See McNeil *et al.* (2015) for more details.

Although the ARCH(p) model explains the change of variance over time by including the past squared error terms, a high-order of p is required to describe the persisting high volatility. Bollerslev (1986) expanded the ARCH model and introduced the GARCH model to overcome this problem. The error term a_t of a time series is modeled by GARCH(p, q) as

$$a_t = \sigma_t \epsilon_t, \quad \sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i a_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 \,, \tag{2.10}$$

where again ϵ_t is a white noise with zero mean and unit variance and $\alpha_0 > 0$, $\alpha_i \geq 0$, $\beta_j \geq 0$, for $i = 1, \ldots, p$ and $j = 1, \ldots, q$. In practice, it is proven that the GARCH of a low order, such as GARCH(1, 1), parsimoniously describes the volatility clustering and is widely used to fit financial data, as shown in McNeil *et al.* (2015), Engle (2001), Bollerslev *et al.* (1992), and Bera and Higgins (1993). We also fit our cyber loss data to the GARCH(1, 1) when we consider a time series model. From (2.10), the conditional variance σ_t^2 under GARCH(1, 1) is

$$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \,, \tag{2.11}$$

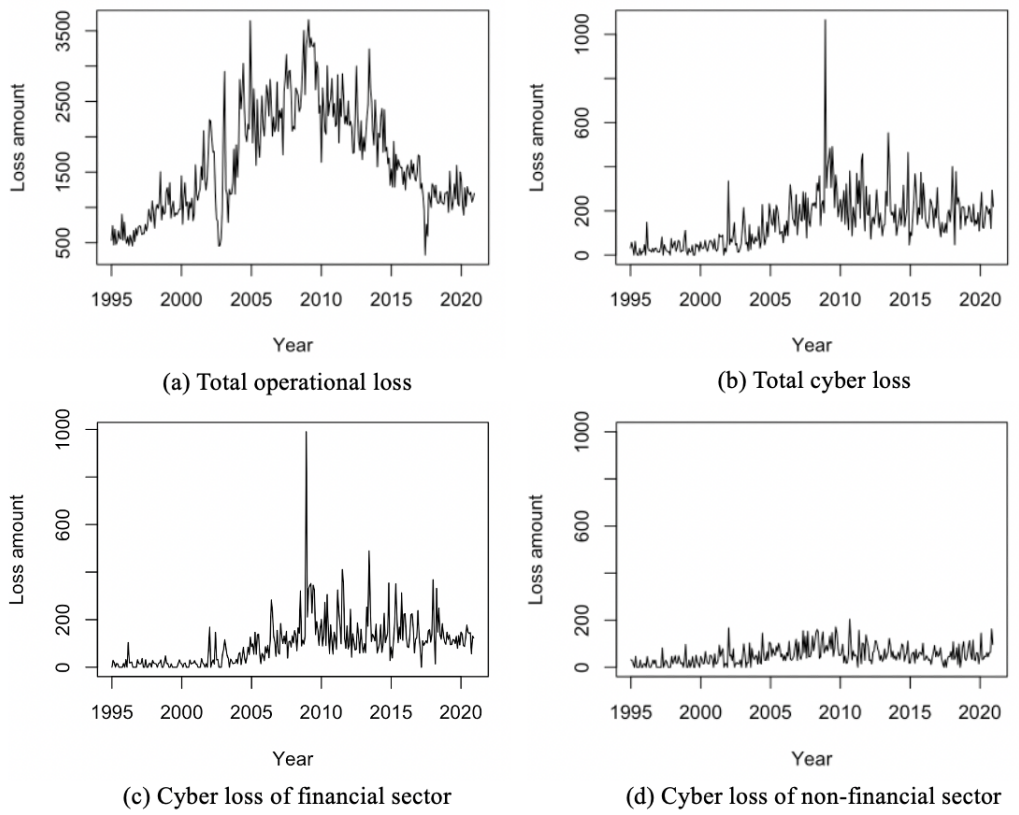

Figure 1: *Trends of monthly log loss amount.*

where $\alpha_1 + \beta_1 < 1$. For more details of GARCH, refer to Bollerslev (1986), Bera and Higgens (1993), and McNeil *et al*. (2015).

2.5. Copula

Copula is a method to create a joint distribution from given marginal distributions and dependence structure. According to Sklar (1959), there exists a unique copula function C defined on $[0,1]^n$ in (2.12) for a joint distribution function F and continuous univariate marginals $F_1, F_2, \ldots, F_n$.

$$F(x_1, x_2, \ldots, x_n) = C(F_1(x_1), F_2(x_2), \ldots, F_n(x_n)). \tag{2.12}$$

Conversely, C is expressed using distribution functions as

$$C(u_1, u_2, \ldots, u_n) = F\left(F_1^{-1}(u_1), F_2^{-1}(u_2), \ldots, F_n^{-1}(u_n)\right),$$

where $0 \le u_i \le 1$, $i = 1, \ldots, n$.

For example, t-copula is defined as

$$C_{v,R}^t(u_1, u_2, \ldots, u_n) = t_{v,\mathrm{R}}\left(t_v^{-1}(u_1), t_v^{-1}(u_2), \ldots, t_v^{-1}(u_n)\right), \tag{2.13}$$

where $t_{v,R}$ is the CDF of n-dimensional multivariate t-distribution with degrees of freedom v and correlation matrix R, and t_v^{-1} is the quantile function of t-distribution with degrees of freedom v. Similarly, Gaussian copula is defined as

$$C_R^{\text{Gauss}}(u_1, u_2, \ldots, u_n) = \Phi_R\left(\Phi^{-1}(u_1), \Phi^{-1}(u_2), \ldots, \Phi^{-1}(u_n)\right), \tag{2.14}$$

where Φ^{-1} is the inverse CDF of standard normal and Φ_R is the joint CDF of an n-dimensional normal distribution with zero mean vector and correlation matrix R. More details about copulas are included in, for example, Byun and Song (2021). Di Clemente and Romano (2004) stated that using t-copula captures dependence structure better than the Archimedean copulas because it has $n(n-1)/2+1$ parameters, whereas the Archimedean copulas have only one parameter for describing the entire dependence structure. Kole *et al.* (2007) suggests that t-copula overwhelms Gaussian and Gumbel copula in the goodness-of-fit test of copulas in risk management. Hence, we apply t and Gaussian copula to our analysis. Although Gaussian copula allocates less probability to the tail dependence, we also consider Gaussian copula to compare with t-copula.

3. Data analysis

3.1. Data description

Previous studies such as Carfora *et al.* (2010) and Edwards *et al.* (2016) examined cyber risk with PRC data. Although the PRC data is open to public, it only describes data-breach events and does not give details about loss amounts. On the other hand, SAS® OpRisk Global Data contains detailed information about loss amounts of operational risk. According to SAS® OpRisk Global Data document on the SAS homepage, Eling and Wirfs (2015), Eling and Wirfs (2016), and Eling and Wirfs (2019), the SAS dataset is complete and reliable because it only holds cases that official media have at least once reported. Note that it contains losses over $100,000. Here, we also use SAS® OpRisk Global Data to measure cyber risk.

Cebula and Young (2010) defined cyber security risks as "operational risks to information and technology assets that have consequences affecting the confidentiality, availability, or integrity of information or information systems." As operational risk encompasses cyber risk, we applied the keyword searching method described in Eling and Wirfs (2015) to SAS® OpRisk Global Data to extract cyber risk data. Eling and Wirfs (2015) defined that to be categorized as cyber risk, the incident should satisfy three criteria: Critical assets, actor, and outcome. First, critical assets such as PC, personal information, or records should be affected. Secondly, business asset losses must have occurred by people's actions, systems, and technical failures, internal or external events. Lastly, there must be a clear outcome, e.g., lost, damaged or breached. We further included more keywords such as code, program, spyware, online, internet, and manipulation, which are not mentioned in keywords that Eling and Wirfs (2015) used, to extract more cyber loss events. After extracting cyber risk cases, we read every event description, evaluated whether it fits the category of cyber losses, and included only relevant events.

We obtained 3,048 cyber risk events through keyword filtering among 33,734 operational losses from January 1995 to December 2020. Our study is relatively more recent than Eling and Wirfs (2015), Eling and Wirfs (2016), and Eling and Wirfs (2019), where the data from January 1995 to March 2014 were used. Figure 1 shows the overall trend of operational and cyber losses during the time period that we used. We can see that monthly operational losses are larger than monthly cyber losses and overall, cyber losses tend to increase as time goes by. Also, there are some extreme observations in cyber loss data. Histograms of two losses show different shapes in Figure 2, indicating that

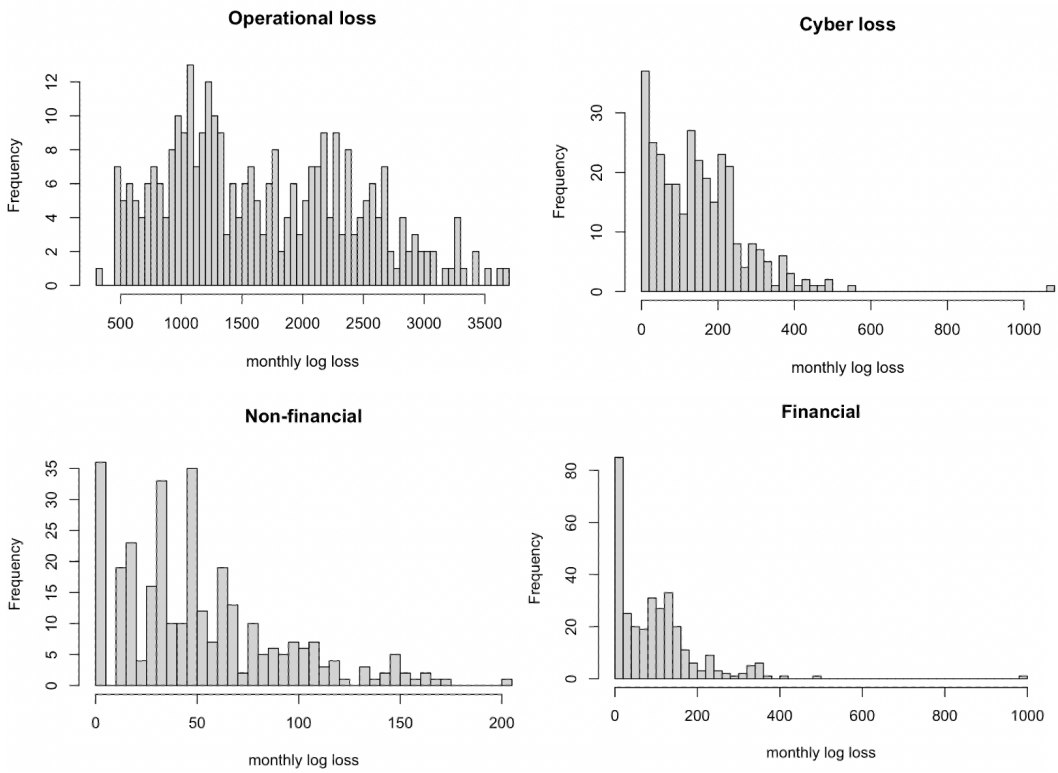

Figure 2: *Histograms of monthly log loss amount.*

Table 1: Descriptive statistics of monthly log operational and cyber losses

	N	Mean	SD	Median	Skewness	Kurtosis	Max
Operational loss	312	1663.2	768.4	1539.6	0.411	2.273	3655.8
Cyber loss	312	150.2	121.1	135.5	1.952	12.812	1065.7

the monthly cyber loss distribution has a heavier tail than the monthly operational loss distribution. Note that cyber losses are a part of operational losses, so it is a matter of fact that the total operational losses in a month are larger than the total cyber losses.

In Table 1, the mean and the median of monthly log operational losses are larger than those of monthly log cyber losses. But the skewness and the kurtosis of monthly log cyber losses are 1.952 and 12.812, respectively, which are about five times larger than those of monthly log operational losses. It indicates that the distribution of monthly log cyber losses is more asymmetric and heavy-tailed than monthly log operational losses. Also, from Figure 2, we notice that the monthly operational losses have lighter tail than the monthly cyber losses. This finding aligns with Eling and Wirfs (2015), but our study displays a more visible difference. Considering that cyber loss data shows a different pattern from operational loss data, it would be meaningful to separate cyber losses out from the operational losses and analyze cyber losses only.

Figure 1 (c) and (d) display the separate time series of monthly cyber losses in the financial sector

Table 2: Descriptive statistics of monthly log cyber losses in financial and non-financial sectors

	N	Mean	SD	Median	Skewness	Kurtosis	Max
Financial	312	98.9	102.5	86.5	2.924	21.205	989.7
Non-financial	312	51.3	39.8	45.4	0.998	3.783	203.7

Table 3: Train and test datasets in the rolling window method

	Train set	Test set
first window	January 1995 ~ December 2011	January 2012
second window	February 1995 ~ January 2012	February 2012
third window	March 1995 ~ February 2012	March 2012
⋮	⋮	⋮
last window	December 2003 ~ November 2020	December 2020

and in the non-financial sector. Losses in the financial sector show a similar pattern to the total monthly cyber losses. The size of monthly cyber losses from the non-financial sector is much smaller than that from the financial sector, especially after around 2005. On the other hand, as we looked at individual losses in both sectors, the sizes of losses in the non-financial sector are not smaller than those in the financial sector. According to Eling and Wirfs (2015), the average cyber loss amount from January 1995 to March 2014 in the non-financial sector is 1.7 times higher than that of the financial sector. The reason that we see larger sizes of monthly cyber losses from the financial sector would be mainly due to the high frequency of cyber losses in that sector.

From Table 2, the median of the monthly cyber loss from the financial sector is about two times larger than that from the non-financial sector, and the maximum value from the financial sector is almost five times larger than that from the non-financial sector. The kurtosis of the financial sector is also much larger than that of the non-financial sector, which indicates that the financial sector has a more heavy-tailed loss distribution than the non-financial sector. The reason would be the chain of events, due to the correlation among companies. For example, the Madoff investment scandal (International Banker, 2021) that occurred in 2009 in the financial sector affected many financial firms and banks all around the world. The tendency of the cyber incidents that occur simultaneously in the financial sector could make the distribution of the monthly financial cyber losses heavy-tailed. Considering the different characteristics between two sectors, we had better investigate the loss distributions of financial and non-financial sectors separately. The heavier tail of the monthly cyber losses in the financial sector is depicted in Figure 2.

While previous studies, including Carfora *et al*. (2019), Eling and Wirfs (2015), and Edwards *et al*. (2016), did not focus on the correlation between the financial and non-financial business sectors, we would like to fit a joint distribution of losses of financial and non-financial sectors considering their correlation. In our dataset, Kendall's τ and Spearman's ρ between monthly cyber losses of the two industries are 0.39 and 0.54, respectively, both significant at the significance level of 0.05. For fitting a joint distribution to bivariate data of monthly cyber losses from financial and non-financial business sectors, we utilize a copula model. The joint distribution we obtain can be used to compute the risk measure for a group of companies from both sectors by reflecting the dependency.

3.2. Model selection

This section will describe the model selection process. We use the rolling window method with the window size of 17 years. Table 3 describes the train and test sets in the rolling window method

Table 4: frequency distribution of the first rolling window : Goodness-of-fit

		Log-likelihood	AIC	Chi-square test	K-S test
Financial	**Negative binomial**	**−555.51**	**1115.02**	227.63***	0.06
(n = 1,071)	Poisson	−1001.77	2005.54	1470.60***	0.37***
Non-financial	**Negative binomial**	**−454.53**	**913.06**	231.20***	0.04
(n = 627)	Poisson	−507.55	1017.10	508.26***	0.14***

Note : n is the number of individual loss observations in the first window.

Table 5: Severity distribution of the first window : Goodness-of-fit

		Log-likelihood	AIC	KS test	AD test
	Exponential	−3955.22	7912.45	0.54***	193.60***
	Normal	−2365.13	4734.27	0.19***	66.01***
	Skewed-normal	−2269.98	4545.96	0.27***	87.65***
	Skewed-t	**−2248.30**	**4504.60**	0.30***	107.71***
	Gamma	−2328.85	4661.70	0.20***	71.12***
Financial	Weibull	−2455.54	4915.09	0.19***	60.92***
(n = 1,071)	Logistic	−2367.57	4739.14	0.22***	90.23***
	Cauchy	−2529.41	5062.82	0.31***	120.32***
	GPD	−3575.89	7155.78	0.36***	120.88***
	NIG	−2276.13	4560.26	0.27***	88.80***
	Skewed-t with POT(95%)	**−2132.55**	**4277.10**	/	/
	Exponential	−2364.68	4731.356	0.51***	135.67***
	Normal	−1382.12	2768.250	0.09***	6.51***
	Skewed-normal	**−1375.68**	**2757.360**	0.10***	8.24***
	Skewed-t	−1375.68	2759.360	0.60***	8.24***
	Gamma	−1384.94	2773.870	0.10***	7.46***
Non-financial	Weibull	−1415.49	2834.970	0.07***	5.82***
(n = 627)	Logistic	−1391.19	2786.380	0.11***	9.14***
	Cauchy	−1507.60	3019.200	0.13***	14.89***
	GPD	−2462.19	4928.380	0.66***	271.48***
	NIG	−1375.87	2759.740	0.10***	7.90***
	Skewed-normal with POT(92%)	**−1261.06**	**2532.120**	/	/

Note : n is the number of individual loss observations in the first window.

that we use throughout Sections 3 and 4. In Section 3.2.1, 3.2.2, and 3.2.3, we explain the steps of selecting models with the train set of the first window of Table 3. We repeat the steps until the last window and use the results in Section 4 to calculate and backtest VaR. Train sets are used for modeling the distribution of cyber losses in Section 3.2.1 and test sets are used for measuring value-at-risk in Section 4.

3.2.1. LDA

LDA assumes the loss amount from each cyber loss incidence independent and models the frequency and severity distributions separately. Considering that Eling (2012) depicted skewed distributions as "good" models for severity compared to other benchmark distributions such as t, normal inverse Gaussian (NIG), and hyperbolic distributions, we include the skewed-normal and the skewed-t distributions as candidates of our cyber loss severity and choose the most appropriate distribution using AIC. We go through the following steps in order to find the distributions of loss severity and loss frequency of financial and non-financial business sectors.

- **Step 1** : We count the number of cyber loss occurrences per month and set this number as the frequency of cyber loss in LDA. The log loss amount of each occurrence is the severity of cyber

loss in LDA.

- **Step 2** : We fit negative binomial and Poisson distributions to frequency data and select one with the smaller AIC. For the severity, exponential, normal, skewed-normal, skewed-*t*, logistic, gamma, Weibull, Cauchy, GPD, NIG with POT are considered. As before, we select the distribution that has the smallest AIC value. The threshold of POT is chosen based on the mean excess plot. Repeat this step for every window in Table 3.

- **Step 3** : We generate 5,000 observations from the selected frequency distribution. According to the generated frequency observations, we generate the appropriate number of severity observations to obtain monthly log loss amounts. We will call this process Monte Carlo simulation as compared to the modified historical simulation in Section 4.

The Goodness-of-fit results for the train set of the first window in Table 3 are shown in Tables 4 and 5.

In Table 4, we present Log-likelihood, AIC, and the values of test statistics of Chi-square test and Kolmogorov-Smirnov (K-S) test. *** after a number indicates the significance at the level of 1%. Negative binomial distribution has smaller AIC values than Poisson distribution for both of financial and non-financial business sectors. Also, with a K-S test result not rejecting the null hypothesis, we say that the negative binomial distribution is suitable for modeling the number of monthly cyber loss occurrences of financial and non-financial industries. Edwards *et al*. (2015) analyzed total daily cyber loss frequency through the PRC data and concluded that the negative binomial distribution fits the daily frequency better than Poisson, binomial, and zero-inflated Poisson distributions. Eling and Wirfs (2019) also suggested that the negative binomial distribution is better than Poisson distribution in modeling monthly and yearly cyber loss frequency distributions. Our result coincides with Eling and Wirfs (2019).

Among single distributions for severity from Table 5, the skewed-*t* distribution and the skewed-normal distribution have the smallest AIC values in the financial and the non-financial sectors, respectively. When we apply POT to these distributions; that is, fit the GPD to the data larger than the threshold, the log-likelihood and AIC are all improved although we do not include all the results in Table 5. It is worth noting that the single parametric distribution with the lowest AIC also gives the smallest AIC value with POT method. This finding confirms the result of Eling and Wirfs (2015) but is different from the result of Edwards *et al*. (2015) that suggests that cyber loss amounts be modeled with the lognormal distribution. Considering that some of the monthly cyber losses are quite extreme in Figure 2, it would be reasonable that POT works well in modeling the severity distribution. For choosing the threshold values in POT, we employ the mean excess plot based on the Equations in (2.5) and (2.6). Figure 3 shows the mean excess plots of financial and non-financial cyber losses. The point where the slope of the graph changes to be positive could be an optimal-threshold as seen in (2.6). Such values for the financial and the non-financial sectors are the 95^{th} and 92^{th} percentiles, respectively. We use these percentiles for all rolling windows.

3.2.2. GARCH

Previous studies with SAS ® oprisk global data such as Eling and Wirfs (2015) viewed each cyber loss as independent. However, as Eling and Wirfs (2016) suggested, one of cyber risks' central properties is dependency, either temporal or cross-sectional. Here, we present a time series model that can incorporate the time dependency and in Section 3.2.3, we suggest a copula model to describe the cross-sectional dependency. McNeil and Frey (2000) applied a two-step time series model to the

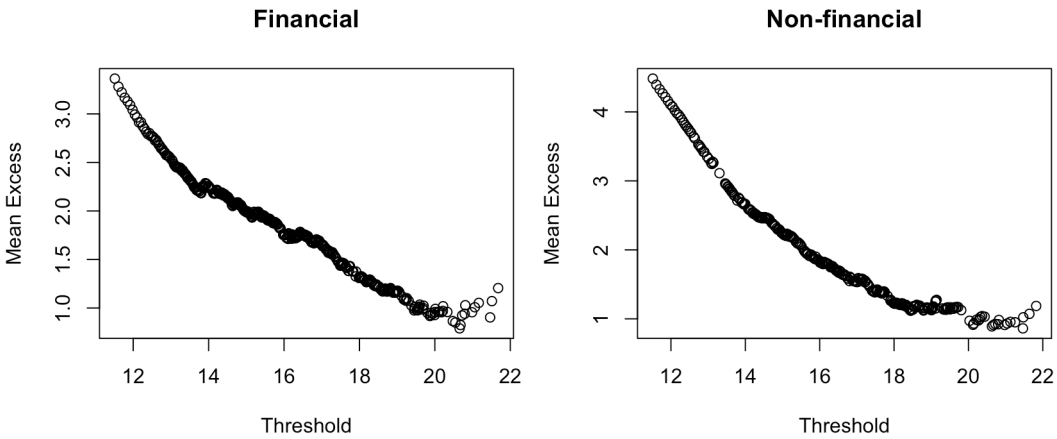

Figure 3: *Mean excess plot.*

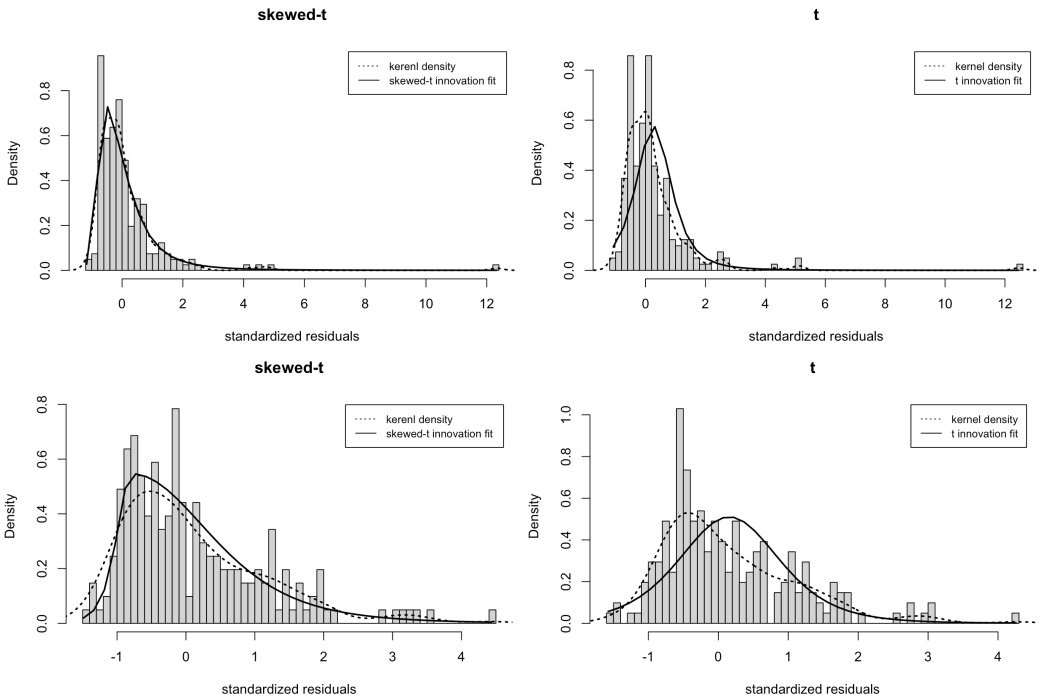

Figure 4: *skewed-t and t innovation distributions for the financial (top) and non-financial (bottom) sectors.*

stock return data. As a first step, they fit an appropriate GARCH-Type model to the return data using a pseudo-maximum-likelihood approach, predict μ_{t+1} and σ_{t+1} from the fitted model, and obtain the residuals. Secondly, they viewed the residuals as i.i.d. white noise and modeled the tail of residuals using the extreme value theory (EVT). They used the EVT to estimate the tail estimator as described in (2.7).

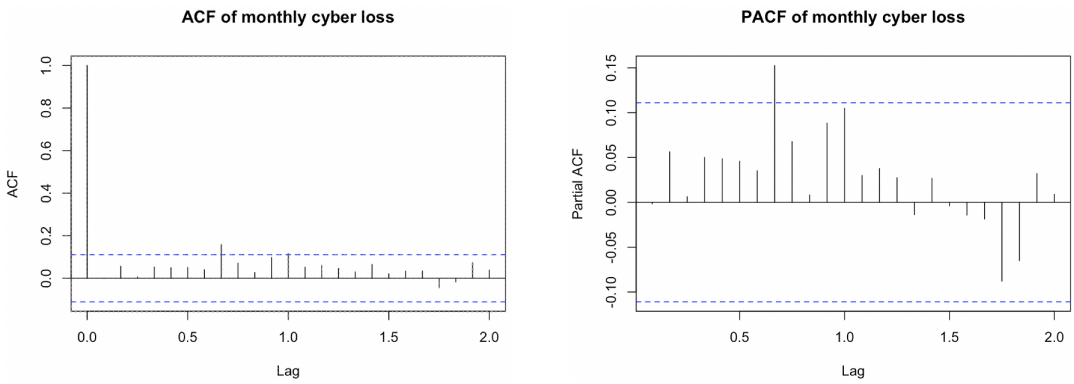

Figure 5: *Auto correlation function (ACF) and partial auto correlation function (PACF) of squared residuals of ARMA*(1, 1).

We fit ARMA(1, 1) to explain the conditional mean and GARCH(1, 1) to describe the conditional variance part as follows. If X_t is the monthly loss at time t, our model is written as

$$\begin{aligned} X_t &= \mu_t + \sigma_t \epsilon_t, \\ \mu_t &= \phi_0 + \phi_1 X_{t-1} - \theta_1 a_{t-1}, \\ a_t &= \sigma_t \epsilon_t, \\ \sigma_t^2 &= \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 \sigma_{t-1}^2. \end{aligned} \tag{3.1}$$

Here, ϵ_t is a white noise with zero mean and unit variance, μ_t is the mean term of the time series and σ_t^2 is the conditional variance of the time series.

Our study applies McNeil and Frey (2000)'s two-step method to monthly cyber loss amounts. We practice GARCH models to incorporate the time-dependent attribute of monthly cyber loss. The Lagrange multiplier test, known as the ARCH test, rejects the null hypothesis that the residuals from the model are a series of white noise and Figure 5 shows that there are some autocorrelation between the squared residuals of ARMA(1,1) model, indicating the necessity to apply a GARCH model.

Monthly cyber losses of financial and non-financial business sectors are analyzed separately following the next steps.

- **Step 1** : For the train set of each window in Table 3, we fit ARMA(1, 1)-GARCH(1, 1) model. We selected the order of ARMA model based on the criterion of small AIC values, and the order of GARCH model from literature, as mentioned in Section 2.4.

- **Step 2** : We use the AIC criterion to choose the innovation distribution, among Gaussian, t, and skewed-t distributions.

- **Step 3** : We extract standardized residuals from the fitted model and apply POT to get $\hat{\xi}$ and $\hat{\beta}$ in (2.7). Then construct the tail estimator given in (2.7) using those values until the last window in Table 3.

Since the tail of the cyber loss is generally heavier than the stock return data, we consider skewed-t distribution as well as t-distribution to better explain the skewness and tail heaviness. In Table 6, we

Table 6: time series models for the first window : Goodness-of-fit

	Model	Innovation	Log-likelihood	AIC
Financial	GARCH(1, 1)	Gaussian	−1138.019	11.196
		t	−1105.699	10.889
		skewed-t	−1095.651	10.801
	ARMA(1, 1)-GARCH(1, 1)	Gaussian	−1123.756	11.076
		t	−1039.490	10.260
		skewed-t	**−1014.358**	**10.023**
Non-financial	GARCH(1, 1)	Gaussian	−1051.310	10.346
		t	−1048.028	10.324
		skewed-t	−1002.051	9.8828
	ARMA(1, 1)-GARCH(1, 1)	Gaussian	−1014.501	10.005
		t	−1007.646	9.9475
		skewed-t	**−984.5953**	**9.7313**

present the result of the selected model for monthly cyber loss of financial and non-financial sectors, for the first window in Table 3.

Table 6 shows that the AIC values of ARMA(1, 1)-GARCH(1, 1) with skewed-t innovations are the smallest as 10.023 and 9.7313 for the financial and non-financial sector, respectively. We also compare the goodness-of-fits of t and skewed-t distributions for the innovation distribution in Figure 4 and find that skewed-t distribution fits the standardized residuals of financial and non-financial sectors better than t-distribution. Table 1, Table 2, and Figure 4 show that the standardized residuals of monthly cyber log loss is not just fat-tailed and leptokurtic but also right-skewed. Therefore, we choose ARMA(1, 1)-GARCH(1, 1) with skewed-t innovations for modeling monthly cyber log loss data for both of financial and non-financial sectors.

3.2.3. Copula

So far, we considered LDA and GARCH as models that explain the tail heaviness and the time-dependency of cyber loss data. In this section, we build a copula model to incorporate the correlation structure between financial and non-financial business sectors. LDA is applied for modeling marginal distributions. Here are the steps that we follow.

- **Step 1** : Carry out **Step 1** and **Step 2** in Section 3.2.1 and obtain the aggregated monthly log loss distributions for both of financial and non-financial sectors. Since we use the same dataset as in Section 3.2.1, same frequency and severity distributions are selected, which are negative binomial distribution and skewed-t with POT. These distributions are the marginal distributions of the copula model. Then we form the correlation matrix from the original dataset.
- **Step 2** : Using the correlation matrix obtained from **Step 1**, we build copula models using Gaussian copula and t-copula for the joint distribution of financial and non-financial business sectors.
- **Step 3** : Generate 5,000 observations from the model obtained in **Step 2**. These are a sample of size 5,000 from the joint distribution of monthly cyber log loss data from two business sectors. We compute the VaR predictions in Section 4 with random samples generated in this step.

For the train set of the first window in Table 3, Kendall's τ between monthly cyber log losses of financial and non-financial sectors was 0.386, which was statistically significant. Likewise, every rolling window provided a statistically significant Kendall's τ-value. By a copula model, we can derive a joint distribution of the cyber losses of financial and non-financial sectors that incorporates the cross-sectional dependency.

Table 7: VaR estimates for the first window

		90%	95%	99%
Financial	LDA	207.054	269.386	457.510
	LDA-POT	196.112	265.937	445.872
	GARCH	261.021	329.789	557.654
	GARCH-POT	309.781	387.473	738.258
Non-financial	LDA	107.935	132.042	197.028
	LDA-POT	106.454	131.798	183.979
	GARCH	137.077	168.499	249.358
	GARCH-POT	150.928	185.210	250.920

4. VaR and backtesting results

This section provides the VaR prediction of monthly cyber log loss data and backtesting results from models we obtained in Section 3. For LDA, 90%, 95%, and 99% VaRs are computed using Monte Carlo simulation and a modified historical simulation. The modified historical simulation we used selects a parametric distribution for the frequency but generates observations for the severity from historical observations. 'LDA' in Tables 7–9 are the results of VaR calculation through modified historical simulation. For GARCH, we forecast the one-step ahead conditional mean and conditional variance to calculate the VaR as

$$\hat{\mu}_{t+1} = \hat{\phi}_0 + \hat{\phi}_1 X_t - \hat{\theta}_1 \hat{a}_t \,,$$
$$\hat{\sigma}^2_{t+1} = \hat{\alpha}_0 + \hat{\alpha}_1 \hat{a}^2_t + \hat{\beta}_1 \hat{\sigma}^2_t.$$

Then, we compute 90%, 95%, and 99% VaRs using $\hat{x}_q$ in (2.7) as

$$\mathrm{VaR}_q = \hat{\mu}_{t+1} + \hat{\sigma}_{t+1} \hat{x}_q.$$

We calculate VaRs for each methodology and every window in Table 3. Table 7 shows VaR estimates using the train dataset of the first window. For the financial sector, 90%, 95%, and 99% VaRs for GARCH-POT are 309.781, 387.473, 738.258, respectively, which are larger than the corresponding VaR estimates from GARCH. In the case of non-financial sector, using POT generally gives larger VaR values than those without POT.

Note that not all the windows give the same result as Table 7. In Table 7, the VaR estimates of LDA are larger than the VaR estimates of LDA-POT in both financial and non-financial business sectors. However, for other windows, the VaR estimates are generally higher when the POT is applied since POT models the heavy tail. We observe the same phenomenon in the VaR estimates of GARCH and GARCH-POT in Table 7. Both the financial and non-financial sector's GARCH-POT VaR estimates are higher than those of GARCH. This tendency is confirmed in Figure 6 which displays the VaR prediction when GARCH and GARCH-POT are applied to every rolling window. The solid and dashed lines, which are the predictions of GARCH-POT, are above the dotted and dash-dotted lines, respectively.

Table 8 contains the backtesting results of the LDA and GARCH for financial and non-financial business sectors. For historical simulation, the negative binomial distribution was used for the frequency. When LDA-POT is applied, the negative binomial distribution and skewed-t body with POT were used for the frequency and the severity, respectively. ARMA$(1,1)$-GARCH$(1,1)$ is fitted for the GARCH model, and the POT is additionally applied for the standardized residuals to obtain the VaR for GARCH-POT. Violation rate is defined as the portion of exceedance obtained by counting the number of actual losses larger than the predicted VaR value. We desire this rate to be as close as

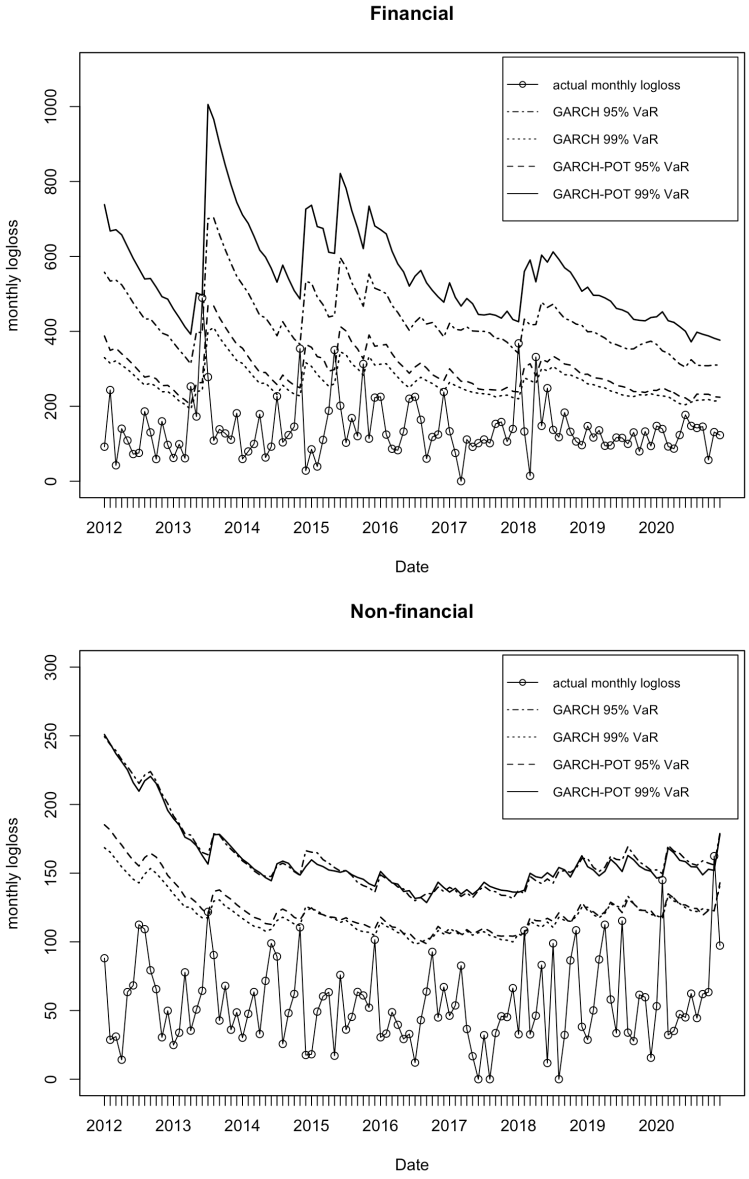

Figure 6: *VaR prediction using GARCH and GARCH-POT.*

possible to the suggested significance level, 0.1, 0.05, and 0.01. Kupiec test (Kupiec, 1995), known as the proportion of failure test, is performed to determine whether or not the prediction of VaR is proper. We expect not to reject the null hypothesis that the observed failure rate is the same as the proposed confidence level. We confirm that the p-values of Kupiec tests are large enough so that the predictions are proper.

From Table 8, we can see that applying POT to the LDA in the financial sector performs bet-

Table 8: Violation rates of VaR prediction

	Financial				Non-financial			
	LDA	LDA-POT	GARCH	GARCH-POT	LDA	LDA-POT	GARCH	GARCH-POT
90% VaR	0.083	0.102	0.120	0.092	0.056	0.093	0.111	0.074
95% VaR	0.056	0.056	0.065	0.056	0.019	0.019	0.037	0.037
99% VaR	0.009	0.009	0.019	0	0	0.009	0.009	0.009

LDA means the modified historical simulation, LDA-POT means the LDA method with POT, GARCH means the GARCH model, and GARCH-POT means the GARCH model with POT.

Table 9: Violation rate with joint distributions: Bootstrap assumes independence between two business sectors

	Violation rate		
	Bootstrap	Gaussian copula	t-copula
90% VaR	0.144	0.111	0.111
95% VaR	0.418	0.056	0.046
99% VaR	0.006	0.009	0.009

ter than the historical simulation. When tried on the non-financial sector, the LDA-POT seemed to overestimate 95% VaR but properly evaluate 90% and 99% VaR. Although the 95% VaR calculated through the LDA method is overestimated in the non-financial sector, modeling the severity of cyber loss with the LDA-POT in the financial and non-financial sectors is generally better than modeling it with the LDA in terms of preventing extreme losses.

As for the time series model, GARCH-POT overestimated 99% VaR of the financial sector but showed a better performance in 90% and 95% VaRs than GARCH model. On the other hand, GARCH model provides a more accurate result in terms of violation rates than GARCH-POT in case of non-financial sector. This could be explained by Table 2 where non-financial sector's monthly cyber log loss has a much smaller kurtosis than that of financial sector's loss. Since non-financial sector's monthly cyber log losses are not as extreme as those of the financial sector, applying POT to GARCH model does not improve the VaR prediction much. Hence, implementing GARCH-POT seems more advantageous only for the financial sector.

Table 9 shows the VaR backtesting results for copula models. To see the effect of correlation between two business sectors, we include the backtesting results from bootstrap samples. The bootstrap results are obtained through resampling the monthly log losses of financial and non-financial sectors separately 5,000 times without considering time order and correlations between the two business sectors. Then, two sectors' VaR were added by giving each sector's VaR a 50% weight. This method destroys the correlation structure of original data, so that it can be used as a benchmark for bivariate analysis. Both Gaussian and t-copulas performed better than bootstrap, especially for 95% VaR. The violation rate of t-copula is slightly closer to the target rate than Gaussian when the confidence level is 95%, but there is not much difference. The results we obtained here implies that modeling through copulas derives satisfactory outcomes with a statistically significant correlation. Besides, copula models are so flexible that we can reflect the dependency among business sectors while we fit each marginal distribution separately.

5. Conclusion

This study focused on modeling cyber loss distributions and the computation of their risk measures. We divided cyber loss data into financial and non-financial business sectors and modeled the loss distribution separately. Generally, financial sector had more extreme monthly cyber losses than non-financial sector. And the monthly cyber loss data showed characteristics as asymmetry and tail

heaviness, time dependency, and correlation among business sectors. The long-tail of cyber losses led us to combine POT with the LDA, a widely used operational risk management model. In the actuarial context, the LDA is widely used for risk management due to the ease of obtaining parametric loss distribution and calculating minimum capital amounts. Thus, we applied the same technique to cyber losses based on previous studies such as Edwards *et al.* (2016) and Eling (2012). Aligned with these studies, we fitted the loss frequency with negative binomial distribution rather than Poisson distribution. For severity, modeling with GPD in the tail and skewed-t or skewed-normal body distributions provided a good fit. As for the risk measure, value-at-risk values calculated using LDA-POT were better than those from the historical simulation in both of financial and non-financial sectors.

Moreover, as we observed volatility clustering from the monthly cyber loss data in Section 3.1 and Section 3.2.2, we tried GARCH-Type models. We fitted ARMA$(1,1)$-GARCH$(1,1)$ and found that the standardized residuals displayed a heavy and right-skewed tail. When we compared GARCH and GARCH-POT, GARCH-POT performed better in the financial sector. It is expected from the extreme observations in the monthly cyber log loss of financial sector in Figure 2. In case of non-financial business sector, GARCH-POT overestimated 90% VaR, which was expected because monthly cyber log losses of non-financial sector is not as extreme as those of financial sector. In practice, since calculating the correct minimum capital amount for cyber loss will benefit financial institutions in forestalling the worst scenario, we suggest that financial industries model their monthly cyber loss with time series model and POT. On the other hand, we conclude that GARCH would be just fine for non-financial sector.

Lastly, we observed the statistically significant correlation between monthly cyber losses of financial and non-financial sectors. Thus we utilized copula models to construct a joint distribution of financial and non-financial sectors' cyber losses and compared them with the bootstrap method. Once we obtain a joint distribution reflecting the correlation structure, calculating the risk measure for different business sectors would become more accurate. Gaussian and t-copulas both performed quite well in computing VaR as seen from the backtesting results.

Due to the lack of information in the dataset, we could not analyze the patterns of individual cyber loss observations. It would be good if we could see the distributional properties of individual observations in the future. Also, we would like to supplement our cyber loss model with the enlarged data in the future by combining various cyber loss datasets. Furthermore, we would like to analyze monthly cyber losses in more detailed classification by applying high-dimensional analysis.

References

Aue F and Kalkbrener M (2007). LDA at work: Deutsche Bank's approach to quantifying operational risk, *Journal of Operational Risk*, **1**, 49-93.

SAS, Retrieved June 20, 2021, Available from: https://www.sas.com/content/dam/SAS/en_us/doc/produ\protect\@normalcr\relaxctbrief/sas-oprisk-global-data-101187.pdf

Balkema AA and De Haan L. (1974). Residual life time at great age, *Annals of Probability*, **2**, 792–804.

Basel Committee on Banking Supervision (2006). International Convergence of Capital Measurement and Capital Standards: A Revised Framework - Comprehensive Version, Bank for International Settlements, Retrieved December 13, 2021, Available from: www.bis.org

Basel Committee on Banking Supervision (2021). Newsletter on cyber security, Retrieved November 25, 2021, Available from: https://www.bis.org/publ/bcbs_nl25.htm

Beirlant J, Goegebeur Y, Segers J, Teugels JL, Waal DD, and Ferro C. (2004). *Statistics of Extremes:*

Theory and Applications, John Wiley & Sons, New Jersey.

Bera AK and Higgins ML (1993). ARCH models: Properties, estimation and testing, *Journal of Economic Surveys*, **7**, 305–366.

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

Bollerslev T (1986). generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics*, **31**, 307–327.

Bollerslev T, Chou RY, and Kroner KF (1992). ARCH modeling in finance: A review of the theory and empirical evidence, *Journal of Econometrics*, **52**, 5–59.

Byun K and Song S (2021). Value at risk of portfolios using copulas, *Communications for Statistical Applications and Methods*, **28**, 59–79.

Carfora MF and Orlando A (2019). Quantile based risk measures in cyber security, In *proceedings of 2019 International Conference on Cyber Situational Awareness*, Data Analytics And Assessment (Cyber SA), Oxford, 1–4.

Cebula JL and Young LR (2010). A taxonomy of operational cyber security risks, *Carnegie-Mellon Univ Pittsburgh Pa Software Engineering Inst.*

Chernobai AS, Rachev ST, and Fabozzi FJ (2008). *Operational Risk: A Guide to Basel II Capital Requirements, Models, and Analysis*, John Wiley & Sons, New Jersey.

Coles S, Bawa J, Trenner L, and Dorazio P. (2001). *An Introduction to Statistical Modeling of Extreme Values*, Springer, London.

Cont R (2007). *Long Memory in Economics*, Springer, Heidelberg.

Di Clemente A and Romano C (2004). A copula-extreme value theory approach for modelling operational risk, *Operational Risk Modelling and Analysis*, (pp. 189–208), Risk Books, London.

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

Eling M (2012). Fitting insurance claims to skewed distributions: Are the skew-normal and skew-student good models?, *Insurance: Mathematics and Economics*, **51**, 239–248.

Eling M and Wirfs JH (2015). Modelling and management of cyber risk, University of St.Gallen, Oslo, Available from: https://www.actuaries.org/oslo2015/presentations/IAALS-Wirfs&Eling-P.pdf

Eling M and Wirfs JH (2016). *Cyber risk: Too big to insure? Risk transfer options for a mercurial risk class*, I. VW HSG Schriftenreihe, St. Gallen.

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

Embrechts P, Klüppelberg C, and Mikosch T (2013). *Modelling Extremal Events: For Insurance and Finance*, Springer Science & Business Media, Berlin.

Engle RF (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, *Econometrica: Journal of the Econometric Society*, **50**, 987–1007.

Engle R (2001). GARCH 101: The use of ARCH/GARCH models in applied econometrics, *Journal of Economic Perspectives*, **15**, 157–168.

Gilli M (2006). An application of extreme value theory for measuring financial risk, *Computational Economics*, **27**, 207–228.

Ghosh S and Resnick S (2010). A discussion on mean excess plots, *Stochastic Processes and their Applications*, **120**, 1492–1517.

International Banker (2021), Retrieved May 05, 2022, Available from: `https://internationalbanker.com/history\-of-financial-crises/bernie-madoffs-ponzi-scheme-2008/`

Jacobs J (2014) Analyzing ponemon cost of data breach, Retrieved December 10, 2021, Available

from: http://datadrivensecurity.info/blog/posts/2014/Dec/ponemon/

Kole E, Koedijk K, and Verbeek M (2007). Selecting copulas for risk management, *Journal of Banking & Finance*, **31**, 2405–2423.

Kupiec P (1995). Techniques for verifying the accuracy of risk measurement models, *The Journal of Derivatives*, **3**, 73–84.

McNeil AJ and Frey R (2000). Estimation of tail-related risk measures for heteroscedastic financial time series: An extreme value approach, *Journal of Empirical Finance*, **7**, 271–300.

McNeil AJ, Frey R, and Embrechts P (2015). *Quantitative Risk Management: Concepts, Techniques and Tools-revised Edition*, Princeton university press, New Jersey.

Philippe J (2001). *Value at Risk: The New Benchmark for Managing Financial Risk*, McGraw-Hill Professional, New York.

Pickands III J (1975). Statistical inference using extreme order statistics, *Annals of Statistics*, **3**, 119–131.

Rydman M (2018). Application of the peaks-over-threshold method on insurance data, Available from: https://www.diva-portal.org/smash/get/diva2:1231783/FULLTEXT01.pdf

Sklar A (1959). Fonctions de reprtition an dimensions et leursmarges, *Publications de l'Institut Statistique de l'Universit é de Paris*, **8**, 229–231.

Received May 06, 2022; Revised June 01, 2022; Accepted June 01, 2022