Modelling and Management of Cyber Risk

Abstract

Cyber risks are an important point on the business agenda in every company, but they are difficult to assess due to the absence of reliable data and profound analyses. To improve this situation, we identify cyber losses from an operational risk database and analyze these with methods from the field of actuarial science. Specifically, we apply operational risk models in order to yield consistent risk estimates, depending on country, industry, size, and other variables. Our results show that human behavior is the main source of cyber risk and that cyber risks are very different compared to other operational risk. The results of the paper are useful for practitioners, policymakers and regulators in order to provide a better understanding of this new and important type of risk.

Keywords: Cyber risk, operational risk, risk management, insurance

1 Introduction

Although cyber risk is a crucial topic for the economy and society and is reported in the media every day[1], it has been subject of very limited academic research.[2] This is most likely due to the wide absence of reliable data. The aim of this paper is to go one step forward and provide a thorough empirical analysis of cyber risks. We extract 1,579 cyber risk incidents from an operational risk dataset and analyze these with methods from the field of actuarial science.

To get deeper insights into the nature and statistical properties of cyber risk, we apply the complete actuarial toolbox. Specifically, we test whether models which prove to be useful for analyzing operational risk can also be applied to cyber risk or whether other tools are needed. We are interested in the question whether cyber risks are structurally identical to other operational risks or exhibit distinct characteristics. Our results show that human behavior is the main source of cyber risk. Moreover, cyber risks are structurally different from other operational risk.

---

[1] Cyber attacks were denoted by the G20 group as a threat to the global economy (see Ackermann, 2013); the World Economic Forum (2014) estimates the probability of a critical information infrastructure breakdown with 10 percent and the financial consequences after a few days to about US$ 250 billion.

[2] The existing literature on cyber risk is mostly limited to papers from the field of technology. Within risk and insurance, our paper is closest to Biener, Eling, and Wirfs (2015) who analyze the insurability of cyber risk and illustrate their statistical properties using descriptive statistics. We build upon and extend their data and analyze it with a longer coverage period and a more thorough empirical analysis that goes beyond descriptive statistics. Other articles on cyber risk and cyber insurance emphasize its complexity (e.g., Hofmann and Ramaj, 2011; Öğüt, Raghunathan, and Menon, 2011) and adverse selection and moral hazard problems (e.g., Gordon, Loeb, and Sohail, 2003).

These results are important for (the CFO and CRO of) every company in order to get a better understanding of cyber risks and their consequences. A special importance is given in the financial services sector, since regulators require banks and insurance companies to hold risk capital for operational losses which might result from cyber risks.[3] Moreover, our results are useful for insurance companies which are developing cyber insurance policies and do not have enough data and experience with cyber risks. We illustrate the usefulness of our results for policymakers, regulators and practitioners in two applications (risk management, pricing). For the academic audience we present effective and contemporary modeling and solution approaches for the novel application area of cyber risk.

The remainder of this paper is structured as follows. In Section 2 we define the term "cyber risk" and introduce our data and methodology. Then, in Section 3 the empirical analysis is presented. We conclude in Section 4.

2 Material and Methods

Cyber risk is a dynamic loss category that has not been thoroughly discussed in academic literature yet (Biener, Eling, and Wirfs, 2015). An efficient data collection for cyber risks is just emerging. Typically, information on cyber risk is not publicly available since affected companies tend to not report it.[4] Another problem that hampers the collection of cyber risk data is the absence of a clear-cut definition.

The definition we employ here is based on how banking supervisors categorize operational risk and goes back to Cebula and Young (2010), who 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". Linking cyber risk to operational risk has several advantages: Firstly, it allows distinguishing cyber risk from other established risk categories.[5] Secondly, in structuring cyber risk we can use the established

[3] Also for firms outside the financial services sector, the results are important not only for internal risk management, but also in light of recent regulatory reforms. For instance, the regulatory approaches for new data protection (e.g., in December 2015 the 'Network and Information Security Directive' in the European Union was passed, that (1) improves cybersecurity capabilities in Member States and regulates their cooperation; and (2) requires operators of essential services in the energy, transport, banking and healthcare sectors, and providers of key digital services like search engines and cloud computing, to take appropriate security measures and report incidents to the national authorities; see European Commission, 2015).

[4] Both in the EU and the US there is a discussion on mandatory reporting requirements. If these become reality, then more data and information would be available.

[5] In banking supervision (e.g. BIS, 2006) market, credit, liquidity, legal and operational risks are separated. Insurance supervisors (e.g., CEIOPS, 2009) typically consider market, insurance, credit and operational risks.

subcategories from operational risk (see Table 1).[6] And thirdly, linking cyber risks to operational risks allows to clearly identifying relevant data.

Table 1 Categories of cyber risk (see Cebula and Young, 2010)

Category		Description	Elements
Actions of people			
1.1	Inadvertent	unintentional actions taken without malicious or harmful intent	mistakes, errors, omissions
1.2	Deliberate	actions taken intentionally and with intent to do harm	fraud, sabotage, theft, and vandalism
1.3	Inaction	lack of action or failure to act upon a given situation	lack of appropriate skills, knowledge, guidance, and availability of person to take action
Systems and technology failures			
2.1	Hardware	risks traceable to failures in physical equipment	failure due to capacity, performance, maintenance, and obsolescence
2.2	Software	risks stemming from software assets of all types, including programs, applications, and operating systems	compatibility, configuration management, change control, security settings, coding practices, and testing
2.3	Systems	failures of integrated systems to perform as expected	design, specifications, integration, and complexity
Failed internal processes			
3.1	Process design and/or execution	failures of processes to achieve their desired outcomes due to poor process design or execution	process flow, process documentation, roles and responsibilities, notifications and alerts, information flow, escalation of issues, service level agreements, and task hand-off
3.2	Process controls	inadequate controls on the operation of the process	status monitoring, metrics, periodic review, and process ownership
3.3	Supporting processes	failure of organizational supporting processes to deliver the appropriate resources	staffing, accounting, training and development, and procurement
External events			
4.1	Hazards	events, both natural and of human origin, over which the organization has no control and that can occur without notice	weather event, fire, flood, earthquake, unrest
4.2	Legal issues	risk arising from legal issues	regulatory compliance, legislation, and litigation
4.3	Business issues	risks arising from changes in the business environment of the organization	supplier failure, market conditions, and economic conditions
4.4	Service dependencies	risks arising from the organization's dependence on external parties	utilities, emergency services, fuel, and transportation

The latter argument is exactly the empirical strategy of this paper: Having defined cyber risk as a subgroup of operational risk, we use the world's largest collection of publicly reported operational losses – the SAS OpRisk Global data – and extract cyber risk events using the search and identification strategy described in Appendix A. The database consists of 26'541 observations between January 1995 and March 2014.[7] All losses are given in USD and adjusted for inflation to make them comparable.[8]

To analyze the statistical properties of cyber risk and to identify the model that describes the data best we use the standard toolbox from actuarial science. After presenting descriptive

[6] Following the operational risk frameworks in Basel II (BIS, 2006) and Solvency II (CEIOPS, 2009), we categorize cyber risk into four classes: (1) actions of people, (2) systems and technology failures, (3) failed internal processes, and (4) external events.

[7] The results in this paper are based on an extract of a much bigger version of this dataset. The original dataset contained 30,173 observations, from March 1971 until March 2014. Due to a very small number of cyber risk incidents before 1995, we cut off the years before 1995, ending up with the dataset mentioned.

[8] The dataset attempts to provide an estimate of the complete costs of operational risk events (both direct as well as indirect effects); however, reputational loss due to an operational risk event is not covered since this sort of loss is typically excluded from operational risk.

statistics, we fit the cyber loss data using extreme value theory. In particular, we implement the loss distribution approach (with applying the peak-over-threshold method for the loss severity), which is standard in modelling operational risk. We also present an extended version of this approach where the loss data depends on covariates (following Chavez-Demoulin, Embrechts, and Hofert, 2015) and fit the loss data to various other distributions which have proven to be useful for actuarial loss analysis (e.g., the g-and-h family of distributions, the Generalized Beta distribution of the second kind, and skewed distributions; see, e.g., Dutta and Perry, 2007, and Eling, 2012). To identify the model that works best, we apply standard goodness-of-fit tests and also more tailored tests for the advanced measurement approaches.

After having identified the best modelling approach, we present two applications: Firstly, a numerical study to estimate the risk measures value at risk and tail value at risk – two measures especially relevant in banking and insurance regulation (Basel II, Solvency II), which show how much capital a company needs to cover losses with a given confidence level (see Eling and Tibiletti, 2010, for definitions of the risk measures). Secondly, we use the numerical results for pricing of a typical cyber insurance policy. Here we use results from a recent market survey (Biener et al., 2015) and standard pricing methods from actuarial science (see, e.g., Bowers et al., 1997). Appendix B gives a detailed description of the methodology.

3 Results

3.1 Descriptive Statistics

Table 2 provides a summary of the cyber risk sample and compares its characteristics with non-cyber risk. All descriptive statistics for cyber risk are significantly smaller than those for non-cyber risk, i.e., the other operational risks.[9] The maximal loss in our sample is US$ 14.6 billion compared to US$ 97.7 billion for non-cyber risk.[10] The loss amounts for cyber risk are thus much smaller than for other operational risks.[11] Sorting into cyber risk subcategories (Panel B of Table 2) shows that most of the cyber risk incidents occur in the "actions of people"

[9] Mean and median are close to estimations of average losses found in the literature; Ponemon Institute (2015) finds that the mean annualized cost of cybercrime for an organization result in an average financial impact of US$ 7.7 million per year. Average losses from the theft of data are estimated at US$ 2.1 million by KPMG (2013).

[10] The largest cyber risk case occurred at the Bank of China in February 2005 when US$ 14,589.15 million were laundered through one of its branches, which was possible because the bank's internal money laundering controls were manipulated by employees. The largest non-cyber risk case involves the U.S. tobacco company Philip Morris, which, in November 2001, was ordered to pay US$ 97,687.34 million in punitive damages to sick smokers.

[11] Cyber risk policies typically cover a maximum such as, e.g., US$ 50 million. Actual cover limits vary. If US$ 50 million is the limit, then 94% of the cases in our sample would be covered completely by the policy.

subcategory.[12] It thus seems that human behavior is the main source of cyber risk, while the other categories, such as external disasters, are very rare. The average losses across the different subcategories are relatively similar.[13]

Table 2 Losses per risk type (in million US$)

Category	N	Mean	Std. dev.	Min.	Quantiles			VaR	TVaR	Max.
					25%	50%	75%	(95%)	(95%)	
Panel A: Cyber versus non-cyber risk										
Cyber risk	1,579	43.49	426.36	0.10	0.43	1.53	7.43	100.55	730.52	14,589
Non-cyber risk	24,962	98.52	1,154.39	0.10	1.39	5.09	24.45	271.60	1,565.81	97,687
Panel B: Cyber risk subcategories										
Actions of people	1,203	42.66	475.53	0.10	0.42	1.35	5.39	77.75	743.20	14,589
Systems and technical failure	212	45.32	141.23	0.10	0.57	4.78	26.98	232.56	485.10	1,668
Failed internal processes	108	15.12	48.96	0.10	0.36	1.32	7.45	65.62	179.91	372
External events	56	109.12	431.92	0.10	1.04	4.25	19.53	331.06	1,585.58	2,949

Table 3 further separates the cyber and non-cyber risk loss data into several subcategories. The geographic separation (Panel A) shows that Northern American companies experience about twice as many (52.6%) cyber risk incidents than European firms (24.9%) and even more than twice as many as firms located on other continents. This might be due to reporting standards implemented in the US, while such mandatory regulations were not in place for the other continents at the time of data collection. For the loss severity, Asia shows the highest average cyber risk loss, whereas Europe and Northern America have much smaller cyber risk losses. This situation may be due to North American and European firms being more capable of and willing to invest in risk mitigating measures for extreme losses, which results from a longer tradition of recognizing and managing cyber risks as compared to Asia.[14]

Panel B of Table 3 provides a separation into financial and nonfinancial services industries. 75.9% of all cyber risk incidents occur in the financial services industry. This is not surprising since financial services firms, such as banks and insurance companies, store a significant amount of critical personal data.[15] However, the average loss resulting from cyber risk for firms in nonfinancial services industries is almost three times as high as for financial services firms. This finding might be explained by financial services firms having a higher awareness regarding critical data and better protection against cyber risk. For non-cyber risks, firms in the

[12] Hacking attacks, physical information thefts, human failures, and all incidents where employees manipulate data (un-/intentionally) are included here.

[13] The higher mean loss for category "External events" is due to the very small sample and a very high maximal value (average without that maximal value amounts to US$ 57.50 million).

[14] Similar patterns can be observed also for the non-cyber risk sample, although companies from Europe show much higher average losses than Northern American companies.

[15] The market survey of potential customers in the financial services industry (Biener et al., 2015) shows that banks are especially prone to cyber risk, i.e., the respondents from the banking sector had significantly more experience with cyber risk than the respondents from other financial service sectors.

nonfinancial services industries also face higher average losses than firms in the financial services sector; however, the difference is not as substantial as it is for cyber risk.

Table 3 Cyber and non-cyber risk losses (in million US$)

	Cyber risks				Non-cyber risks			
	N	**Share of cyber risk incidents**	**Mean**	**Median**	**N**	**Share of non-cyber risk incidents**	**Mean**	**Median**
Panel A: Region of domicile								
Africa	24	1.52%	30.90	1.86	278	1.11%	58.72	2.59
Asia	256	16.21%	104.31	1.52	3,375	13.52%	132.95	4.04
Europe	393	24.89%	31.09	1.78	5,596	22.42%	121.01	5.49
North America	830	52.56%	33.26	1.42	14,867	59.56%	85.31	5.27
Other	76	4.81%	18.44	1.55	846	3.39%	57.44	4.47
Panel B: Industry								
Nonfinancial	381	24.13%	84.11	4.47	13,665	54.74%	114.31	7.43
Financial	1,198	75.87%	30.57	1.16	11,297	45.26%	79.40	2.92
Panel C: Relation to losses in other firms								
One firm affected	1,283	81.25%	49.21	1.56	17,748	71.10%	87.59	5.02
Multiple firms affected	296	18.75%	18.65	1.45	7,214	28.90%	125.40	5.30
Panel D: Company size by number of employees[*]								
Small	67	4.24%	19.71	1.29	732	2.93%	33.27	1.97
Medium	73	4.62%	13.35	1.09	1,193	4.78%	27.81	2.17
Large	1,375	87.08%	46.17	1.49	20,005	80.14%	112.55	5.63

[*]: Small: Less than 50 employees; Medium: Less than 250, Large: More than 250. The total in each size group does not add up to the total sample, since for a few incidents, the number of employees is not available.

An important aspect of cyber risk is contagion, and thus our next separation of the data is between incidents affecting only one single firm and those affecting multiple firms (Panel C of Table 3). If just one firm is involved (81.3% of the cyber risk cases), the average loss per firm per case is more than twice as high as if more than one firm is involved. This result may appear counterintuitive; however, in the event that more than one firm is affected, cyber attacks are identified earlier and thus losses can be limited. Also, there may be economies of scale in solving the problems created by cyber incidents when multiple firms are involved (e.g., forensic investigation costs).

Panel D of Table 3 separates the sample based on firm size. With increasing size, the number of incidents increases, i.e., firms with more than 250 employees have more cyber losses. Interestingly, we observe a U-shaped pattern in the mean losses both for cyber and non-cyber risk.[16] It may be that smaller firms do not have the awareness and resources to protect against cyber risk, while large firms have diseconomies of scale due to complexity.[17]

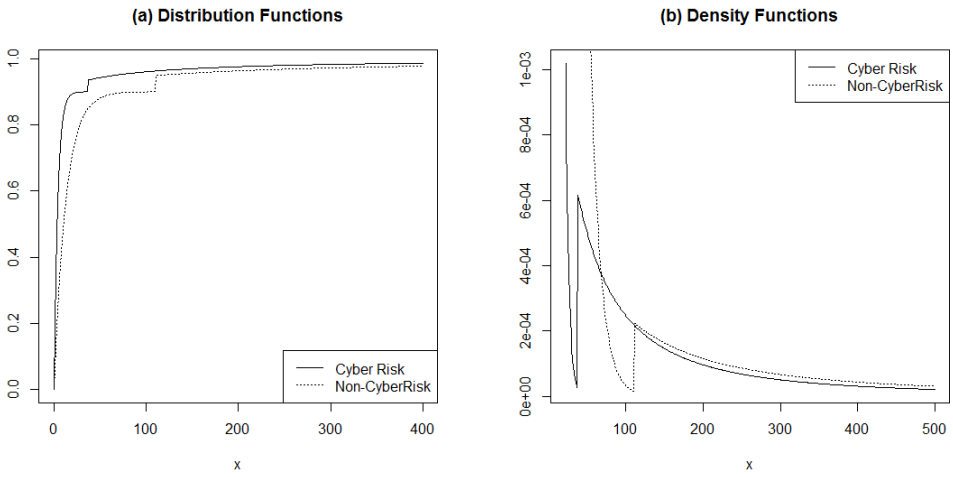
[16] The results are robust with regard to the size categorization. We estimated the values for a separation into Small: less than 100, Medium: less than 1,000, and Large: more than 1,000 employees and find no differences in this pattern.

[17] We also analyzed the development of cyber risks over time and found that the number of cyber risk incidents was relatively small before 2000. After that point, however, the number of incidents continuously increased and in the last years accounts for a substantial part of all operational risk incidents. These findings again emphasize the increasing economic importance of cyber risk in recent years. The average loss, however, has decreased over the last several years, which might indicate the increasing use of self-protection measures that reduce the loss amount in the event of a cyber attack. Detailed results are available in Appendix C.

3.2 Modelling Results and Goodness-of-Fit

To more closely analyze the distributional characteristics of cyber risk compared to other operational risk, we implement methods from extreme value theory when estimating the loss severity distribution (e.g., Peaks-over-Threshold method; POT). In this approach, losses above a predefined threshold are modeled by a generalized Pareto distribution (GPD) while losses below the threshold are modeled with a distribution common in loss modelling (e.g., the Exponential, Gamma, Log-normal or the Weibull distribution; see, e.g., Chapelle et al. 2008).[18/19] The estimated distributions for cyber and non-cyber risk are presented in Figure 1.

Figure 1 Estimated distribution and density function



In addition, we model losses with other parametric distributions common in actuarial science, such as the Exponential, Gamma, GPD, Log-logistic, Log-normal, or Weibull distribution (see, e.g., Dutta and Perry, 2007) and the skew-normal and skew-student t distributions (see, e.g., Eling, 2012).[20] Furthermore, we include a non-parametric transformation kernel estimation (see, Bolancé, Guillen, and Pitt, 2014). Results of the goodness-of-fit analysis are presented in Table 4. The measures we use are the value of the log-likelihood function, the Akaike Information Criterion (AIC), and the Kolmogorov-Smirnov-Tests (K-S tests). Since the K-S test puts more weight on the center of the distribution and by that does not show good evaluation

[18] For the distribution of the body we use the Log-normal distribution, because it will provide the best overall fit for the single parametric distributions (second best fit for cyber risk and best fit for non-cyber risk; see Table 4). Also, in the analysis with different body distributions it also provides the best fit for both risk categories; see Appendix D.

[19] We apply the bootstrap goodness-of-fit test by Villaseñor-Alva and González-Estrada (2009) and, based on this, choose a threshold at the 90% percentile. For purposes of comparison, we also computed results for a 92.5% threshold, with findings similar to those with a 90% threshold; thresholds below reveal a non-fit for non-cyber risks according to Villaseñor-Alva and González-Estrada (2009); raising thresholds much higher makes the sample used for the fit in cyber risk too small.

[20] We included the skew-student distribution into our analysis, however, we were not able to make the maximum-likelihood-estimation converge yet. The results will be contained in the next draft of this paper.

in the tails, we also apply the Anderson-Darling-Test (A-D test), where possible. This test puts more weight on the tail than on the center of the distribution (see, e.g., D'Agostino and Stephens, 1986). We also present a graphical goodness-of-fit analysis by Q-Q plots in the Appendix D.

The results from the K-S tests indicate that none of the five single parametric distributions models the cyber risk loss data adequately. Furthermore, these distributions also do not fit the non-cyber risk data, which motivates the use of more advanced modelling approaches. From the seven distributions, the GPD provides the best results for the cyber losses, but also for this model the null hypothesis in the K-S test is rejected at a 1% confidence level. Similar applies to the log-normal distribution for the non-cyber losses. Looking at the POT approach, we observe the best fit – for cyber risk and non-cyber risk – under all models in Table 4. This motivates the use of EVT and the further extension of this approach.

Table 4 Goodness-of-Fit Analysis

Model	Log-likelihood	AIC	Kolmogorov-Smirnov-Test	Anderson-Darling-Test
Panel A: Cyber risk (N = 1,579)				
Exponential	-7,535.78	15,073.55	0.60 ***	79.94 ***
Gamma	-5,368.23	10,740.46	0.24 ***	18.79 ***
GPD	**-4,553.42**	**9,110.84**	0.07 ***	7.18 ***
Log-logistic	-4,591.40	9,186.80	1.00 ***	13.22 ***
Log-normal	-4,588.09	9,180.19	0.08 ***	16.99 ***
Weibull	-4,886.78	9,777.57	0.16 ***	60.20 ***
Skew-normal	-10,718.32	21,442.63	0.82 ***	166.03 ***
Skew-student	No convergence	/	/	/
POT (threshold 90%)	**-4,529.26**	**9,066.53**	/	/
Transformation Kernel	**-4,402.48**	**/**	/	/
Panel B: Non-Cyber Risk (N = 24,962)				
Exponential	-139,542.80	279,087.60	0.54 ***	58.75 ***
Gamma	-109,184.80	218,373.60	0.21 ***	13.10 ***
GPD	-99,438.54	198,881.10	0.03 ***	50.27 ***
Log-logistic	-99,572.73	199,149.50	1.00 ***	61.63 ***
Log-normal	**-99,258.09**	**198,520.20**	0.03 ***	54.86 ***
Weibull	-102,587.30	205,178.60	0.10 ***	6.30 ***
Skew-normal	-194,267.50	388,541.00	0.81 ***	230.30 ***
Skew-student	No convergence	/	/	/
POT (threshold 90%)	**-99,074.40**	**198,156.80**	/	/
Transformation Kernel	**-98,120.48**	**/**	/	/

Note: AIC = Akaike information criterion; for the Kolmogorov-Smirnov- and the Anderson-Darling-Test we present the value of the test statistic and the significance level of rejecting the null hypothesis (H_0: the given distribution is equal to the sample distribution). *, **, ***, indicate confidence levels of 10%, 5%, and 1%, respectively.

In the following version of the paper we will also present an extended version of this POT approach where the loss distribution depends on covariates (following Chavez-Demoulin, Embrechts, and Hofert, 2015; preliminary results are attached in Appendix E) and fit the loss data to various other distributions which have proven to be useful for actuarial loss analysis

(e.g., the g-and-h family of distributions, the Generalized Beta distribution of the second kind, and skewed distributions; see, e.g., Dutta and Perry, 2007, and Eling, 2012).

3.3 Applications

We first conduct a numerical study to estimate the risk measures value at risk (VaR) and tail value at risk (TVaR). These measures are especially relevant for regulatory purposes in banking and insurance (Basel II, Solvency II). Table 5 presents the risk measurement results for cyber and non-cyber risk for the POT models and the five parametric distributions.

Table 5 Risk Measurement

Model	Cyber Risk (N = 1,579)		Non-Cyber Risk (N = 24,962)	
	VaR	TVaR	VaR	TVaR
Exponential	130.20	173.72	295.51	393.73
Gamma	213.03	352.20	474.48	772.17
GPD	**94.35**	222,554.60	**237.39**	24,730.33
Log-logistic	56.28	3,104.10	213.01	31,057.40
Log-normal	63.15	238.18	206.62	851.67
Weibull	88.61	196.32	232.64	496.80
Skew-normal	833.14	995.07	2,252.37	2,690.11
Skew-student	No Convergence		No convergence	
POT (threshold of 90%)	**104.63**	1,720.03	226.78	4,565.33
Empirical	100.55	730.52	271.60	1,565.81

Note: Value at risk (VaR) and tail value at risk (TVaR) at 95% confidence level.

The VaR estimator for cyber risk, applying the Exponential, Gamma, Log-normal, and Weibull distribution, is significantly different from the empirical VaR, which indicates that the distribution assumption does not fit the data well in the tails. The result for the GPD distribution is much closer to the empirical VaR than the other four parametric distributions. However, the estimate from the POT provides the best fit for the VaR. Similar results can be observed for the TVaR. The Exponential, Gamma, Log-normal, and Weibull distribution significantly underestimate the TVaR, while the GPD significantly overestimates the TVaR. This suggests that they are not modelling the tail-behavior appropriately. Although overestimating the tail-losses, the POT approach again provides the best fit. Moreover, a more conservative estimation might be appropriate for regulatory purposes.[21] In the comparison of non-cyber risks, GPD provides the best fit for VaR. The POT approach does not show a very good approximation of the VaR and the TVaR, which motivates a further analysis with an extended approach. Furthermore, the results show that the distribution of cyber risk differs significantly from the distribution of other operational risks. For example, the distribution of the non-cyber risk sample shows much higher VaR and TVaR than that of the cyber risk sample, explaining in

[21] An approximation of the loss distribution per category was not conducted, since the sample size would be too small for the computation of the tail distribution, in particular the category "External events". However, an analysis of this issue is possible by the approach described in Chavez-Demoulin, Embrechts, and Hofert (2015).

part the much higher maximal losses in these categories.[22] This finding implies that when modeling operational risk, cyber risk needs to be considered separately.

Secondly, we will use the numerical results to yield a price for a standard cyber insurance policy. This will help to get a sense for the economic relevance of these risks. The results will be contained in the next draft of this paper.

4 Conclusions

Insurance firms start to sign cyber risk policies and might not have a lot experience with cyber risks. For the pricing process of insurance contracts and the estimation of security capitals, a good understanding of the properties and behavior of the risk is vital. Furthermore, regulatory approaches for new data protection and regulation schemes are expected to come (see, e.g., a proposal for the EU; European Commission, 2012). Our findings can provide insight what parts it is most important to look at and what security levels they have to postulate in the risk management processes of their supervised companies. The results of the paper might thus offer important insights for the management of cyber risks, about their insurability and might also provide guidance for the pricing of cyber insurance policies. They are relevant for policymakers and regulators that need to develop sound policies for the treatment of this new, dynamic risk category. For the academic audience we present effective and contemporary modeling and solution approaches for the novel application area of cyber risk.

We need to highlight some limitations of the paper, which yield avenues for future research opportunities. For example, the identification strategy should not be interpreted as more as a first step towards a more thorough analysis of cyber risk. What we do is extracting cases with a predefined criteria catalogue; collecting an own database with cyber risk would be a useful avenue for future research. Also our risk estimates are only a first indication of the true cyber risk; e.g. since reputational risks are not incorporated. If the models from operational risk prove to be useful for an analysis of cyber risks, then recent papers from this field can also be used to estimate the potential reputational effect (see, e.g., Cannas, Masala, and Micocci, 2009, or Cummins, Lewis, and Wei, 2006). We will either directly integrate this issue in the paper or, alternatively, place it in the conclusion of the paper as an avenue for future research as it is right now.

[22] The modeled VaR for non-cyber risk is more than twice as high as for cyber risk.

References

Ackerman, G. (2013) 'G-20 urged to treat cyber-attacks as threat to global economy,' Bloomberg, from www.bloomberg.com/news/2013-06-13/g-20-urged-to-treat-cyber-attacks-as-threat-to-economy.html, accessed 18 January 2014.

Aue, F., and Kalkbrener, M. (2006) 'LDA at work: Deutsche Bank's approach to quantify operational risk', Operational Risk 1(4), 49-93.

Balkema, A. A., and de Haan, L. (1974) 'Residual life time at great age', Annual Probability 2, 792-804.

Bank for International Settlements (BIS) (2006) International Convergence of Capital Measurement and Capital Standards: A Revised Framework Comprehensive Version, from www.bis.org/publ/bcbs128.pdf, accessed 10 December 2013.

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.

Biener, C., Eling, M., Matt, A., and Wirfs, J. H. (2015) 'Cyber Risk: Risikomanagement und Versicherbarkeit', I.VW Schriftenreihe, Band 54, St. Gallen.

Bolancé, C., Guillen, M., and Pitt, D. (2014) 'Non-parametric Model for Univariate Claim Severity Distributions – an approach using R', UB Riskcenter Working Paper Series, 2014/01 – Research Group on Risk and Insurance and Finance, University of Barcelona.

Bowers, N., Gerber, H. U., Hickman, J., Jones, D., and Nesbitt, C. (1997) 'Actuarial Mathematics', 2nd ed., Society of Actuaries, Schaumburg, IL.

Cannas, G., Masala, G., and Micocci, M. (2009) 'Quantifying Reputational Effects for Publicly Traded Financial Institutions', Journal of Financial Transformation 27, 76-81.

Cebula, J. J. and Young, L. R. (2010) A Taxonomy of Operational Cyber Security Risks, Technical Note CMU/SEI-2010-TN-028, Software Engineering Institute, Carnegie Mellon University.

CEIOPS (2009) CEIOPS' Advice for Level 2 Implementing Measures on Solvency II: SCR Standard Formula – Article 111 (f): Operational Risk. CEIOPS-DOC-45/09, Frankfurt: Committee of European Insurance and Occupational Pensions Supervisors.

Chapelle, A., Crama, Y., Huebner, G., and Peters, J.-P. (2008) 'Practical methods for measuring and managing operational risk in the financial sector: a clinical study', Journal of Banking & Finance 32(6): 1049-1061.

Chavez-Demoulin, V., Embrechts, P., and Hofert, M. (2015) 'An Extreme Value Approach for Modeling Operational Risk Losses Depending on Covariates', Journal of Risk and Insurance, DOI: 10.1111/jori.12059.

Chavez-Demoulin, V., Embrechts, P., and Nešlehová, J. (2006) 'Quantitative models for operational risk: Extremes, dependence and aggregation', Journal of Banking & Finance 30(10): 2635-2658.

Cope, E., and Labbi, A. (2008) 'Operational loss scaling by exposure indicators: Evidence from the ORX database', The Journal of Operational Risk 3(4), 25-45.

Cope, E. W., Piche, M. T., and Walter, J. S. (2012) 'Macroenvironmental determinants of operational loss severity', Journal of Banking and Finance 36(5), 1362-1380.

Cummins, J. D., Lewis, C. M., and Wei, R. (2006) 'The Market Value Impact of Operational Loss Events for US Banks and Insurers', Journal of Banking and Finance 30(10), 2605-2634.

D'Agostino, R. B. and Stephens, M. A. (1986) 'Goodness-of-Fit Techniques', New York, NY: Marcel Dekker, Inc.

Dahen, H., and Dionne, G. (2010) 'Scaling models for the severity and frequency of external operational loss data', Journal of Banking and Finance 34(7), 1484-1496.

Davison, A. C. (1984) 'Modelling excesses over high thresholds, with an Application', in J. de Oliveira (ed.), Statistical Extremes and applications, D. Reidel, 461-482.

De Fontnouvelle, P., Dejesus-Rueff, V., Jordan, J. S., and Rosengren, E. S. (2006) 'Capital and risk: New evidence on implications of large operational losses', Journal of Money, Credit, and Banking 38(7), 1819-1846.

Degen, M., Embrechts, P., and Lambrigger, D.D. (2007) 'The quantitative modeling of operational risk: between g-and-h and EVT', ASTIN Bulletin 37(2), 265-291.

Dutta, K., and Perry, J. (2007) 'A tale of tails: an empirical analysis of loss distribution models for estimating operational risk capital', Working Paper No. 06-13, Federal Reserve Bank of Boston.

Edwards, B., Hofmeyr, S., and Forrest, S. (2016) 'Hype and Heavy Tails: A Closer Look at Data Breaches', Working Paper, http://www.econinfosec.org/archive/weis2015/papers/WEIS_2015_edwards.pdf, accessed 08 February 2016.

Eling, M. (2012) 'Fitting insurance claims to skewed distributions: Are the skew-normal and skew-student good models?', Insurance: Mathematics and Economics 51(2), 239-248.

Eling, M., and Tibiletti, L. (2010) 'Internal vs. external risk measures: How capital requirements differ in practice', Operations Research Letters 38(5), 482-488.

Embrechts, P., Klüppelberg, C., and Mikosch, T. (2003) 'Modelling Extremal Events for Insurance and Finance', Springer.

European Commission (2015) 'Network and Information Security Directive: co-legislators agree on the first EU-wide legislation on cybersecurity', from http://ec.europa.eu/digital-agenda/en/news/network-and-information-security-directive-co-legislators-agree-first-eu-wide-legislation, accessed 15 December 2015.

Ganegoda, A., and Evans, J. (2013) 'A scaling model for severity of operational losses using generalized additive models for location scale and shape (GAMLSS)', Annals of Actuarial Science 7(1), 61-100.

Giacometti, R., Rachev, S., Chernobai, A., Bertocchi, M., and Consigli, G. (2007) 'Heavy-Tailed Distributional Model for Operational Losses', Working Paper.

Gilli, M., and Këllezi, E. (2006) 'An application of extreme value theory for measuring financial risk', Computational Economics 27(2-3), 1-23.

Gourier, E., Farkas, W., and Abbate, D. (2009) 'Operational risk quantifying using extreme value theory and copulas: from theory to practice', The Journal of Operational Risk 4(3), 1-24.

Gordon, L.A., Loeb, M.P. and Sohail, T. (2003) 'A framework for using insurance for cyber-risk management', Communications of the ACM 44(9): 70–75.

Gustafsson, J, Nielsen, J. P., Pritchard, P., and Roberts, D. (2006)' Operational risk guided by kernel smoothing and continuous credibility: A practitioner's View', The Journal of Operational Risk 1(1), 43-56.

Hess, C. (2011) 'The impact of the financial crisis on operational risk in the financial services industry: Empirical evidence', Journal of Operational Risk 6(1): 23-35.

Hofmann, A. and Ramaj, H. (2011) 'Interdependent risk networks: The threat of cyber attack', International Journal of Management and Decision Making 11(5/6): 312–323.

KPMG (2013) 'e-Crime – Computerkriminalität in der deutschen Wirtschaft mit Kennzahlen für Österreich und Schweiz', KPMG Forensic Services, from www.kpmg.com/CH/de/Library/Articles-Publications/Seiten/e-crimesurvey-2013.aspx, accessed 18 January 2014.

McNeil,A. J., Frey, R., and Embrechts, P. (2005) 'Quantitative Risk Management: Concepts, Techniques, Tools', Princeton University Press.

Moscadelli, M. (2004) 'The modelling of operational risk: Experience with the analysis of the data collected by the Basel Committee', Technical Report 517, Banca d'Italia.

Öğüt, H., Raghunathan, S. and Menon, N. (2011) 'Cyber security risk management: Public policy implications of correlated risk, imperfect ability to prove loss, and observability of self-protection', Risk Analysis 31(3): 497-512.

Pickands, J. (1975) 'Statistical inference using extreme order statistics', Annals of Statistics 3, 119-131.

Ponemon Institute (2015) '2015 Cost of Cyber Crime Study: Global', http://informationsecurity.report/Resources/Whitepapers/5fe1dd7e-b46d-49f6-833a-d192cecb29e3_2015-cost-cyber-crime-study-global-pdf-10-w-2093.pdf, accessed 03 February 2016.

Privacy Rights Clearinghouse (PRC) (2016) 'About the Privacy Rights Clearinghouse', https://www.privacyrights.org/node/1398, accessed 08 February 2016.

Reiss, R.-D., and Thomas, M. (2007) 'Statistical Analysis of Extreme Values', Birkhäuser Verlag, Basel – Boston – Berlin.

Shevchenko, P. V. (2010) 'Implementing loss distribution approach for operational risk', Applied Stochastic Models in Business and Industry 26(3), 277-307.

Shih, J., Khan, A., and Medepa, P. (2000) 'Is the size of an operational loss related to firm size?', Operational Risk Magazine 2(1), 1-2.

Soprano, A., Crielaard, B., Piacenza, F., and Ruspantini, D. (2009) 'Measuring operational and reputational risk – A Practitioner's Approach', Wiley, Finance.

Villaseñor-Alva, J.A. and González-Estrada, E. (2009) 'A bootstrap goodness of fit test for the generalized pareto distribution', Computational Statistics and Data Analysis 53(11): 3835-3841.

World Economic Forum (2014) 'Global Risks 2014 – Ninth Edition', from www3.weforum.org/docs/WEF_GlobalRisks_Report_2014.pdf, accessed 17 April 2014.

Wilson, S. (2007) 'A Review of Correction Techniques for Inherent Biases in External Operational Risk Loss Data', Australian Prudential Regulation Authority.

Supplemental Material for Online Publication Only

Appendix A: Search and Identification Strategy

To be categorized as a cyber risk incident, a loss event must meet three criteria: (1) a *critical asset* such as a company server or database needs to be affected, (2) a relevant *actor* needs to be involved in the cause of the cyber risk incident (e.g., hackers, employees, system, nature), and (3) a relevant *outcome* such as the loss of data or misuse of confidential data needs to be present (see Table A1 for more information). For each category we defined a comprehensive set of keywords, which we then systematically scanned for in the incident descriptions of our SAS OpRisk Global Data database (see Table A2). The resulting dataset includes a total of 1,579 cyber risk incidents, or about 5.9% of the total sample of operational risks.

Table A1 Data Search Strategy

Step	Description
1.	For all three criteria – critical asset, actor, and outcome – we identify keywords that describe terms in the appropriate group
2.	We searched the descriptions of each observation in our sample data for a combination of keywords, where each combination consisted of one word from each group (three-word combinations)
3.	We checked all identified observations individually (reading each description) for their affiliation to cyber risk or non-cyber risk and if necessary we excluded the incidents from the cyber risk term; while checking the observations we also decided in which of the cyber risk categories they fit best
4.	For all observations that were not identified by one of our keyword combinations we checked randomly chosen incidents and included them if necessary; furthermore, if we could identify keyword combinations that we missed in the first round, we started all over at Step 2 with these new words

Table A2 Keywords per Criteria

Critical Asset	Actor	Actor (cont.)	Outcome
account	*(1) Actions by people*	*(2) Systems and technical failure*	availability
accounting system	administrator	defect	available
address	deadline	hardware	breach
code	denial of service, DoS	loading	breakdown
communication	destruction	malicious code	confidential
computer	devastation	software	congestion
computer system	employee	stress	constrain
confidential	extortion	system crash	control
confidential document	forgot, forget, forgotten		delete
consumer information	hacker, hacked	*(3) Failed internal processes*	deletion
data	hacking	unauthorized access	disclosure
disk	human error		disorder
document	infect	*(4) External events*	disruption
file	infection	Blizzard	disturbance
hard-disk	infiltrate	Earthquake	encryption
hard-drive	infiltrated	Eruption	espionage
homepage	key logger	Explosion	failure
info(rmation)	lapse	Fire	false
information system	logic bomb	Flood	falsification
internet site	maintenance	Hail	falsified
names	malware	heat wave	falsifying
network	manager	Hurricane	incompatibility
numbers	manipulate	Lightning	incompatible
online banking	miscommunication	natural catastrophe	incomplete
payment system	mistake	Outage	integrity
PC	misuse	pipe burst	interruption
personal information	omission	Riot	limit
phone	online attack	Smoke	lose
purchase information	oversight	Storm	loss
record	phish	Thunder	lost
reports	phishing	Tornado	malfunction
server	spam	Tsunami	missing
site	Trojan	Typhoon	modification
social security number	vandalism	Unrest	modified
stored information	virus	Utilities	modify
tablet	worm	War	overload
trade secret		Weather	publication
webpage		Wind	restrict
website			sabotage
			steal
			stole
			theft

Note: We used regular expressions to ensure that different spellings were captured (e.g., “homepage” and “home page”).

Appendix B: Methodology

Fitting of Single Parametric Distributions

As a first insight on the distributional properties and because of its easy implementation, the data can be fitted by single parametric distributions. For operational risk, literature suggests the use of more advanced methodologies (mostly from extreme value theory) to model these kind of losses, since simple distributions were found to provide no perfect fit for the extreme events in operational risk (e.g. Moscadelli, 2004). In Section 3.1, we observed that cyber risk losses are much smaller than operational losses. Thus, there might be reasons to believe that the simple parametric distributions can provide a better fit for cyber risk than for operational risk. Thus, we fit parametric distributions first, for instance, Log-normal, Exponential, Gamma, Weibull, Log-Weibull, GPD, Burr, symmetrized alpha-stable, Log-alpha-stable (see Giacometti et al., 2007).

Approaches from Extreme Value Theory (EVT)

If the distribution fitting is used to compute risk capitals (e.g. required in Basel II and Solvency II for operational risks) the single estimation of severity distributions is not adequate, since the distribution of events over time is neglected. Furthermore, the extreme events that can occur in operational risk might not be modelled adequately. For these purposes, methods from EVT have proven to be the right choice in operational risk modelling (see, e.g., McNeil, Frey, and Embrechts, 2005; Embrechts, Klüppelberg, and Mikosch, 2003; Reiss and Thomas, 2007, for an introduction). In this area the loss distribution approach (LDA) has become the most common model, where a loss-frequency distribution and a loss-severity distribution are fitted separately on historical data. The first describes the occurrence of losses over time, while the latter provides information on the potential size of the losses. Afterwards these two distributions are combined to an aggregated loss distribution (see, e.g., McNeil, Frey, and Embrechts, 2005). Since in most of these models there are no closed-form formulas for the aggregates available, the aggregation is typically done by Monte Carlo Simulation.[23] For the modeling of frequency and severity distributions a variety of different approaches exist that we briefly introduce in the following parts.

[23] In banking this approach is applied on a yearly basis and for each business line an aggregated loss distribution is estimated. From those an overall annual loss distribution is estimated by a copula approach that enables to account for diversification effects between business lines (see, e.g., Gourier, Farkas, and Abbate, 2009). This final distribution is then used for the calculation of capital requirements. For an example in the industry we refer to Soprano et al. (2009) for UniCredit Group, or Aue and Kalkbrener (2006) for Deutsche Bank.

Loss-frequency Distribution

The loss-frequency distribution is commonly modeled as a homogeneous Poisson process. It is assumed that the mean number of events occurring in a fixed time interval is constant over time. In practice this could not be confirmed for operational risk (Giacometti et al., 2007). Thus, if it is assumed that the mean number of events in a given time period changes over time, non-homogeneous Poisson processes are used (Giacometti et al., 2007). Those processes assume that the intensity parameter (which defines the average number of events) can be expressed by a mathematical function depending on time. Giacometti et al. (2007) assumes the intensity functions to be Log-normal or Log-Weibull.

Loss-severity Distribution

For the loss-severity distributions, a variety of different approaches are discussed in the literature. For instance, the severity can be modeled by a simple parametric distribution (e.g., Pareto, Log-normal, etc.; Giacometti et al., 2007). However, those approaches do not cover the extreme events of operational risk adequately (Moscadelli, 2004). Thus, an approach that is often applied is the Peak-over-threshold (POT) approach (Embrechts, Klüppelberg, and Mikosch, 2003), in which the extreme values (losses above a predefined threshold u) of the severity distribution are modeled separately from the main body of the losses. The approach is based on the Balkema-de Haan-Pickands Theorem, which states that if the threshold u is chosen reasonably high, the distribution above the threshold can be modeled by a GPD (Pickands, 1975, and Balkema and de Haan, 1974). The body is then fitted on one of the simple parametric distributions discussed before, e.g., exponential (Hess, 2011) or Log-normal distribution (Moscadelli, 2004).[24]

More Advanced Methods from EVT

In literature, several limitations in the estimation of operational losses by the standard EVT approaches are discussed. Those in particular occur, if external data is used. Wilson (2007), for instance, provides a review on biases inherent in external operational risk loss data and discusses potential correction techniques:

- Reporting bias: occurs when different thresholds are used to report losses (e.g. the SAS OpRisk Global data covers losses above US$ 100’000 only, overestimating the losses since it has been fitted to a too large number of higher losses). An approach to correct for reporting

[24] In Biener, Eling, and Wirfs (2015) we provide a first analysis of the loss-severity distribution using POT following Hess (2011) and an exponential distribution for the body.

bias is proposed in De Fontnouvelle et al. (2006). In our case, we know the reporting threshold and can apply the method used in Hess (2011).

- Control bias: occurs because data is generated by institutions with different control mechanisms. Some losses might be irrelevant for some firms, thus not collected, while they might for others, which then cannot be used. We assume that this poses not a problem for our dataset, since we look at publicly reported incidents, reported in media.
- Scale bias: occurs because data is generated by institutions with different sizes (i.e., the loss severity of firms in external databases depends on the size of the firm). This problem can be incorporated, e.g. by adjusting the loss height depending on firm size and further covariates (e.g. business line, and event type; see Ganegoda and Evans, 2013).

Since, Ganegoda and Evans (2013) only adjust the loss-severity distribution by covariates, Chavez-Demoulin, Embrechts, and Hofert (2015) discuss an approach for loss-frequency and loss-severity. In addition, they add a time-dependence to their model, such that changes in loss-frequency / -severity can be modelled appropriately. One of the advantages of this approach is, that data can be pooled (data does not need to be separated into different groups, e.g., business lines, for which the fitting of the distribution must be done separately to figure out differences in the distributions across business lines), and by that sample size does not reduce. Furthermore, interactions between different covariates can be measured (e.g. an interaction between type of loss and change in frequency can be analyzed). The following covariates can and should be modelled by our approach:

- Time: For operational losses Chavez-Demoulin, Embrechts, and Hofert (2015) observed changes in loss-severity and loss-frequency over time. Chavez-Demoulin, Embrechts, and Nešlehová (2006) find a significant relationship between loss-frequency and time.

  Cyber risk's economic importance increased heavily in recent years and thus suggests that cyber risk losses developed over time also. In Biener, Eling, and Wirfs (2015) we observe a relatively small amount of cyber risk incidents before 2000; however, a continuous increase in the last years was shown. For loss severity, average losses decreased over the last years and gave reason to believe that increased use of self-insurance measures reduced the losses occurred.
- Size: The relationship between firm size and the loss severity is extensively discussed in the literature. For instance, Shih et al. (2000), Cope and Labbi (2008), and Ganegoda and Evans (2013) all discuss this relationship and find a positive correlation between size and loss height. This phenomenon is also called the scaling problem or scale bias, which occurs when data is collected from institutions with different sizes.

For our analysis of cyber risk, firm size might also have an influence, in particular on severity AND frequency. The larger the firm, the more sensitive data might be available, the higher the potential losses. The larger the company, the more complex the operations and the more often mistakes and incidents happen. In the descriptive analyses of Section 3.1 we observed an increasing number of incidents with increasing size (measures by number of employees). For the mean losses we observe a u-shape, which makes the inclusion of non-linear size variables in the approach appropriate.

- Business Line: The relation between business lines and loss severity has been analyzed in Dahen and Dionne (2010), Ganegoda and Evans (2013), and Chavez-Demoulin, Embrechts, and Hofert (2015). In the latter paper, the relationship has been analyzed also for the loss-frequency. The results show significant differences for business lines and suggest the analysis also for our approach.

  Unfortunately, in our approach we consider all industries and are not focused on the banking industry as the existing studies. Thus, a separation into the business lines, as, e.g. in Chavez-Demoulin, Embrechts, and Hofert (2015), is not applicable in our case. However, we can differentiate into firms from the financial and nonfinancial industry. In Section 3.1 we observed essential differences for this covariate in cyber losses (most incidents occur in the financial industry group, however, the average losses are just about half of those from the nonfinancial industry).
- Event type: The approaches in Dahen and Dionne (2010) and Ganegoda and Evans (2013) incorporate the event category for operational losses coming from the Basel regulations. As before for business lines, this was modelled for the banking industry specifically. We can easily adjust this to the cyber risk event types discussed in Cebula and Young (2010). In Section 3.1 we identified most of the incidents to fall in the category “Actions of people”, and by that showed that the human behavior is the main source of cyber risk. The average losses per category however, are very similar.
- Geographical region: To the best of our knowledge, we would be the first to differentiate by a geographical covariate. We believe that for cyber risk it is essential, since regulatory / legal responsibilities are completely different for different areas in the world and self-protection standards might be different or regulated differently. In Section3.1 we show that Northern American companies experience more than twice as many cyber risk incidents than European firms, however, for loss severity they show one of the smallest average losses. It could be worthwhile to incorporate this covariate into our analysis.

- For potential further macroenvironmental determinants, see Cope, Piche, and Walter (2012) (e.g., executive power, prevalence of insider trading, shareholder protection laws, restrictions on banking activity, supervisory power, per capita activity, and a government index).

In the further work, we might identify additional covariates that could be interesting to analyze and which could be covered by our dataset.

Further Approaches Beyond Operational Risks

In the paper we will also go beyond standard operational risk models and look at recent developments in the field of actuarial science in order to identify the model which best describes the cyber risk data.

For instance, Dutta and Perry (2007) fitted the g-and-h family of distributions and the Generalized Beta distribution of the second kind (GB2) to operational losses. These approaches were found to provide reasonable fits for non-EVT approaches. Both distributions can be used as approximations for many of the previously mentioned one- and two-parameter distributions and accommodate a wide variety of tail-thicknesses and permit skewness as well (see Dutta and Perry, 2007). Thus, on the one hand, the approach from Dutta and Perry (2007) could provide new insights on the distributional behavior of cyber risk, but on the other hand could serve as a robustness test of the results found earlier. Degen, Embrechts, and Lambrigger (2007) extend Dutta and Perry (2007)'s approach by discussing some fundamental properties of the g-and-h distribution and their link to EVT. They show that under some instances the quantile estimation by EVT approaches might be inaccurate if data is well modelled by a g-and-h distribution.

Another approach that has been applied to operational losses in literature is the one explained in Gustafsson et al. (2006). The approach is based on a non-parametric smoothing technique, utilizing the Generalized Champerowne distribution (GCD). The advantage of his approach is that the tail behavior can be modelled adequately, but unlike EVT, the modelling is done by data from the full distribution (the POT approach models body and tail separately, not considering information from the other part of the distribution). Finally, we fit skewed distributions to the loss data (e.g. skew-normal and skew-student), that have proved to be adequate in describing property-liability insurance claims (Eling, 2012).[25]

[25] All the approaches discussed before, are based in their estimation procedure on maximum-likelihood estimation. Shevchenko (2010) describes an alternative to maximum likelihood based on Bayesian Inference. This approach could be worth to implement since previous knowledge could be incorporated in this approach (e.g. properties of operational risk could be interesting for cyber risks).

Methodology for Comparison of Models

To compare the different fitting approaches with each other, and identify the one that works best, we first apply goodness-of-fit tests that compare the fitted distributions with the empirical data (Kolmogorov-Smirnov and Anderson-Darling test). These tests are standard in the fitting of parametric distributions (Moscadelli, 2004). More tailored tests, in particular for the POT approach, are given in Davison (1984), and Reiss and Thomas (2007). Further tests that provide insights about the appropriateness of estimations are severity VaR performance analyses (Moscadelli, 2004), which we will conduct in a second step. Since most of the fitting algorithms are based on maximum likelihood estimation, we can also compare Log-likelihood values and ground our analysis on the Akaike information criterion (AIC) or the Bayesian information criterion (BIC). A variety of further tests might come up during our implementation period. For the approach in Chavez-Demoulin, Embrechts, and Hofert (2015), further comparison techniques are needed. To identify the best combination of covariates (that model the cyber losses best) and to find the best model specification, Chavez-Demoulin, Embrechts, and Hofert (2015) compare models via likelihood-ratio tests. Furthermore, Ganegoda and Evans (2013) describe an information criterion that could be used to compare two competing models.[26]

Risk Measurement and Pricing of Cyber Insurance Policies

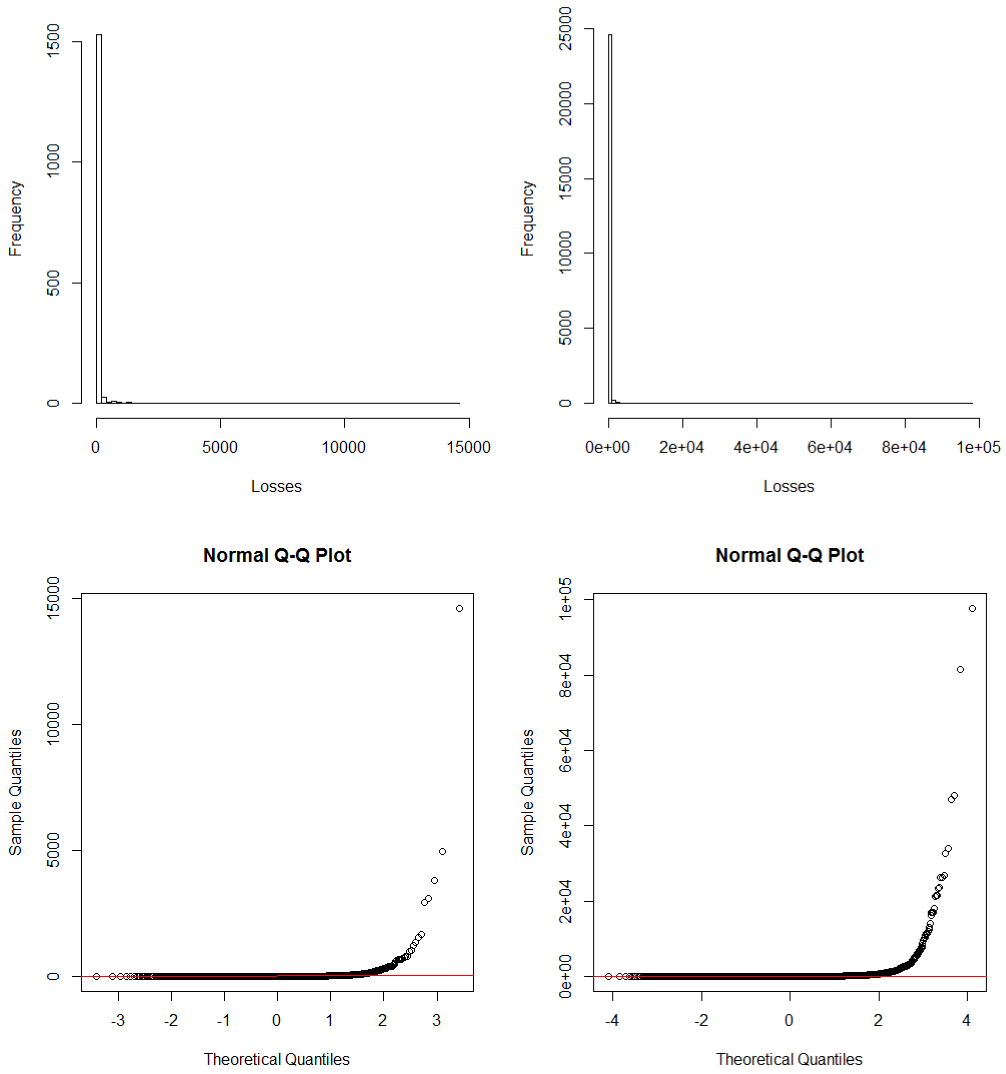
After we have identified the best modelling approach, we will present two applications: Firstly, we will conduct a numerical study to estimate the risk measures value at risk and tail value at risk. These measures are especially relevant for regulatory purposes in banking and insurance (Basel II, Solvency II). Secondly, we will use the numerical results to yield a price for a standard cyber insurance policy. This will help to get a sense for the economic relevance of these risks.

[26] Note, that Ganegoda and Evans (2013) model only the loss-severity, and not loss-frequency AND loss-severity as in the approach of Chavez-Demoulin, Embrechts, and Hofert (2015). Thus, some adjustments might be necessary.

Appendix C: Further Information on the Data Analysis

The following results provide further analyses on the two data sets used. In particular, the results underline the assumption of heavy-tailedness for both data sets and motivate the modelling approaches from extreme value theory (EVT), applied in Section 3.2.

Figure C1 Visualization of original cyber risk (left) and non-cyber risk losses (right)



The histograms in Figure C1 prove that both datasets exhibit a very high number of small losses, and only very few but extreme losses. In addition, the Q-Q plots show that the distributions for cyber risk and non-cyber risk are by far not normally distributed. The skewness and kurtosis parameters presented in combination with the descriptive statistics of Table C1 underline these results.

Table C1 Descriptive Statistics for Original Data

	Cyber risk losses	Non-cyber risk losses
No. of Obs.	1,579	24,962
Mean	43.49	98.52
Std. Dev.	426.36	1,154.39
Skewness	27.12	49.95
Kurtosis	873.33	3,388.68
Minimum	0.10	0.10
Maximum	14,589.15	97,687.34

For a more detailed analysis we also look at the logarithmic losses in both categories. The histograms (see Figure C2) indicate that cyber losses seem to have more very small losses compared to medium-high losses, compared to non-cyber risks. Furthermore, the graphical inspection illustrates the heavy-tails of both of the distributions. The Q-Q plots however indicate that logarithmic non-cyber risk losses seem to be closer to a normal distribution that cyber losses.

Figure C2 Visualization of logarithmic cyber risk (left) and non-cyber risk losses (right)

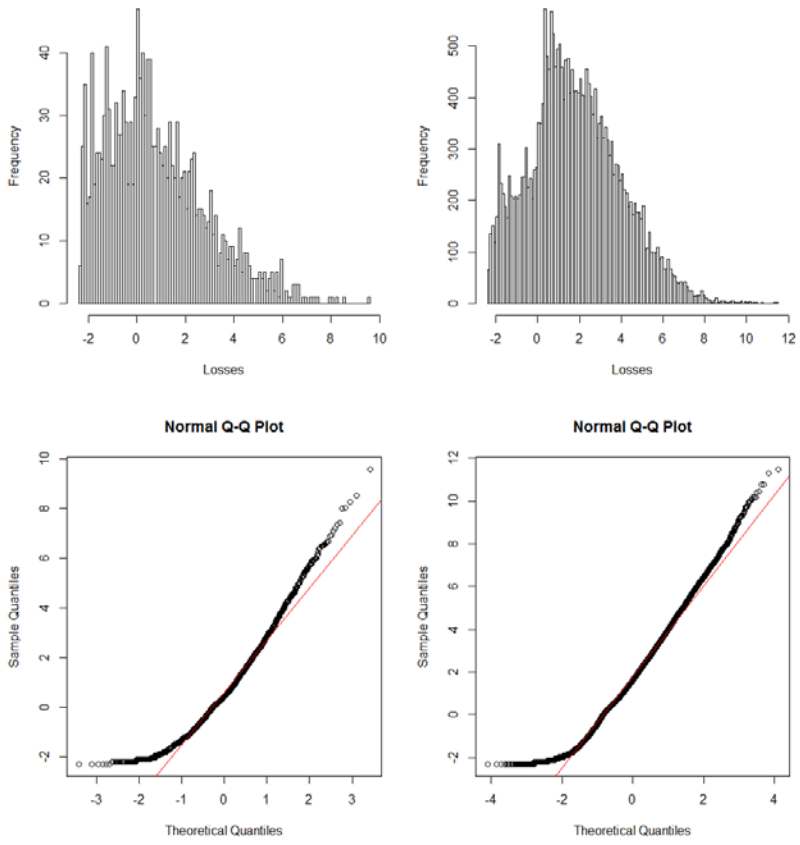


Table C2 Descriptive Statistics for Logarithmic Data

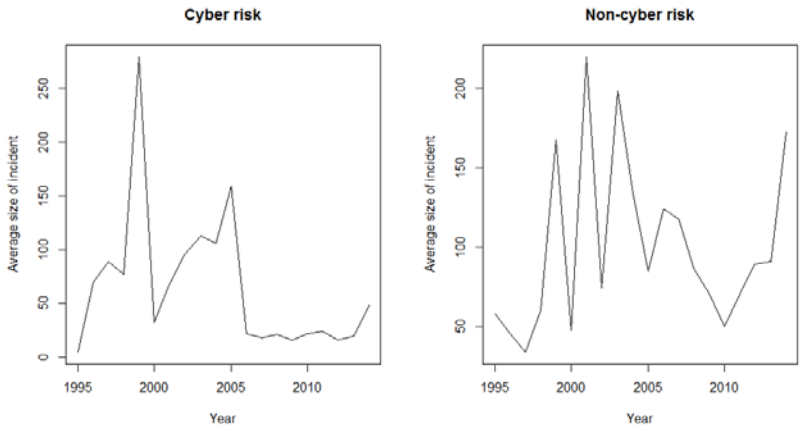
	Cyber risk losses	Non-cyber risk losses
No. of Obs.	1,579	24,962
Mean	0.76	1.79
Std. Dev.	2.06	2.16
Skewness	0.79	0.38
Kurtosis	0.39	-0.09
Minimum	-2.30	-2.30
Maximum	9.59	11.49

Also the descriptive statistics for the logarithmic data (Table C2) underline the results from the visualization in Figure C2 that the logarithmic non-cyber risk data is not as heavy tailed as the cyber risk data. With a skewness parameter of 0.38 and a kurtosis parameter of -0.09 the non-cyber losses show significant lower skewness and kurtosis. In addition, the log-non-cyber losses are closer to normal (platycurtic) kurtosis than the leptocuritc behavior that seems to be present in the cyber losses. These results, thus, provide the first evidence that cyber and non-cyber losses are structurally different. However, both distributions exhibit heavy tails and therefore an analysis by modelling approach from EVT is appropriate.

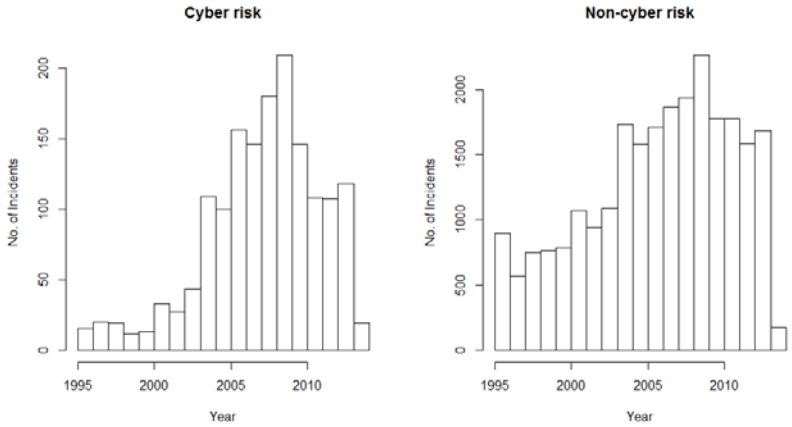
Finally, also analyze both data categories with respect to their development over time. While the average cyber losses over the years (see Figure C3(a)) where very volatile from 1995 to 2005, their development stabilized over the last few years of our observational period.[27] The non-cyber losses show a very volatile behavior over the whole observation period. The analysis of the number of incidents per year (see Figure C3(b)) shows no real difference between the two subsamples. For both datasets the number of incidents increases from 1995 to 2008, and afterwards decreases again. However, the decrease for the cyber loss data seems to be more rapidly.

[27] For year 2014, the data for cyber losses shows a small upward trend, which might be due to the very short observation period in 2014 (only until March 2014).

Figure C3 Visualization of the Development over Time



(a) Average loss per year



(b) Number of incidents per year

Appendix D: Detailed Analyses for the POT-Approach

Analysis of Different Body Distributions:

In the following, we analyze the impact of different body distributions on the overall fit of the POT approach. We assume the single-parametric distributions from Section 3.2 for the different bodies and present the result in Table D1.

Table D1 POT Approach for Different Body Distributions

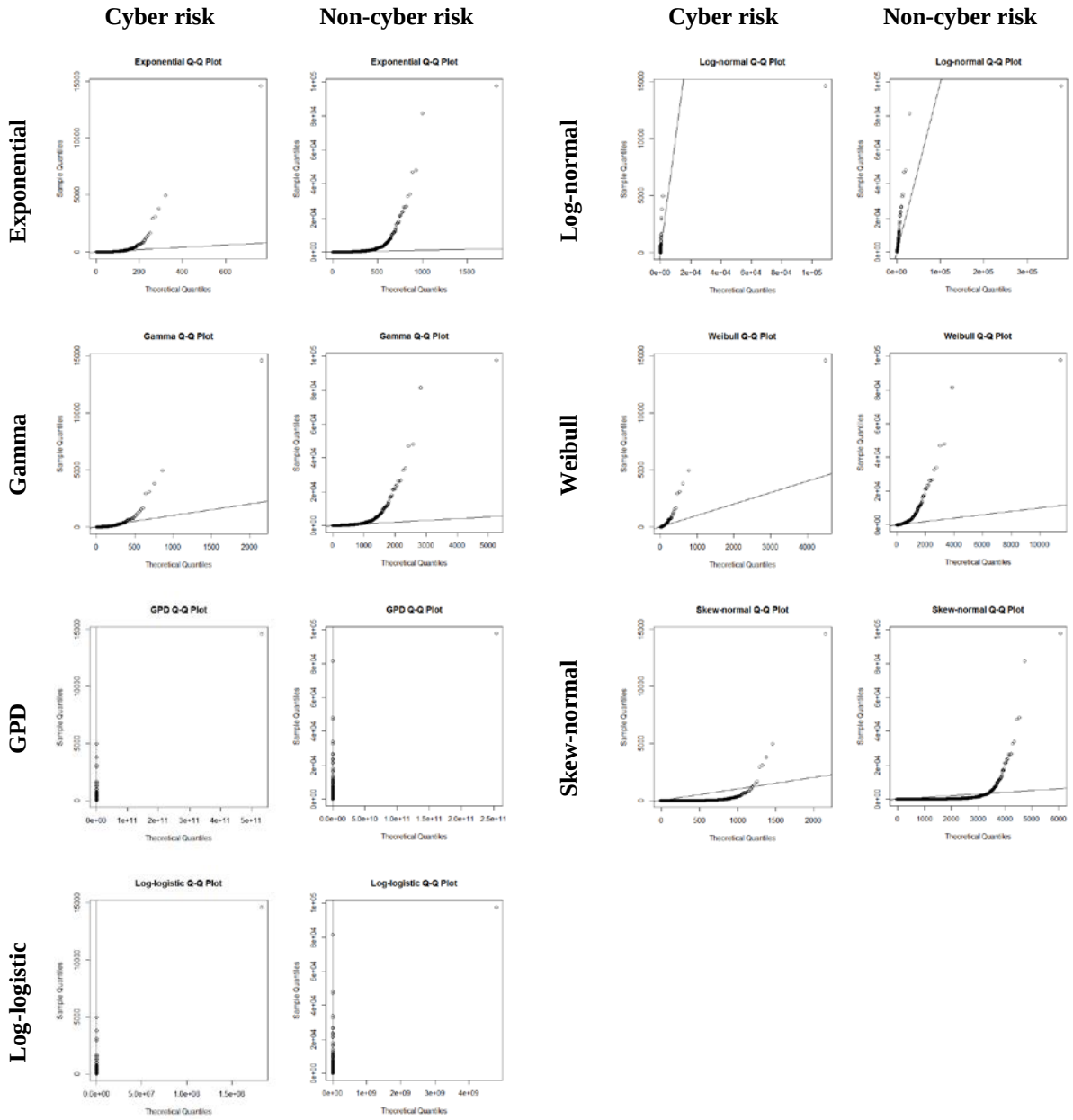
Body	Log-likelihood	AIC
Panel A: Cyber risk (N = 1,579)		
Exponential	-4,931.74	9,869.47
Gamma	-4,722.73	9,453.46
GPD	-4,547.08	9,102.17
Log-normal	**-4,529.26**	**9,066.53**
Weibull	-4,657.31	9,322.31
Skew-normal	-6,679.16	13,368.32
Skew-student	No convergence	/
Panel B: Non-Cyber Risk (N = 24,962)		
Exponential	-105,665.70	211,337.40
Gamma	-101,201.00	202,410.10
GPD	-99,182.28	198,372.60
Log-normal	**-99,074.40**	**198,156.80**
Weibull	-100,673.60	200,673.60
Skew-normal	-131,315.96	262,641.90
Skew-student	No convergence	/

As we can see from the log-likelihood value and the AIC, the POT approach with the log-normal distribution as body provides the best fit in both data subsamples. This was not surprising since this distribution already had one of the best fits for the single parametric distribution analysis (see Table 4). However, in combination with the POT approach this can be even more improved (see also results in Table 4).

Visualization of the Goodness-of-fit (Q-Q plots):

Finally, we also provide the graphical goodness-of-fit analysis of the single parametric distributions with respect to Q-Q plots (see Figure D1). The analysis shows the sample quantiles in comparison to the theoretical quantiles. If the data points produce an almost straight line with the bisectrix (also included in each graph), the data can be adequately described by proposed distribution.

Figure D1 Goodness-of-fit by Q-Q Plots



The results presented in Figure D1 underline the findings from Section 3.2. The GPD and the log-normal (single parametric) distributions provide the best fit for the cyber risk and non-cyber risk losses. Except for the highest quantile observation the results are not on one line, which indicates that the overall fit is very good for both distributions, but still the tails are not fitted perfectly. For all the other distributions the fit is not good already for lower quantiles, and by that indicates that those distributions are inappropriate in describing the cyber and non-cyber losses used in this paper.

Appendix E: Preliminary Results[28]

In Section 3.2 we identified the POT approach to provide the best fit for the loss data (cyber risk and non-cyber risk), compared to the single parametric distributions and motivated the use of more advanced EVT models to fit (non-)cyber losses. Such an advanced model is the extension of the POT approach by modeling the loss data depending on covariates. Chavez-Demoulin, Embrechts, and Hofert (2015) describe the approach in detail, and provide an example for operational loss data. However, the analysis in Chavez-Demoulin, Embrechts, and Hofert (2015) and the modelling of losses depending on covariates are restricted only to the distribution of the loss excesses (i.e., the distribution of the losses above a threshold). To be able to make comparisons with our models in Section 3.2 we implement the approach by Chavez-Demoulin, Embrechts, and Hofert only for the loss severity distribution, and compare it with the results for the excess distribution from the POT model (see, Table 4). If we can prove that the excess distribution under the Chavez-Demoulin, Embrechts, and Hofert (2015) approach provides a better fit than the excess distribution under the normal POT approach, we can replace the modelling of the excess and end up with a better fit of the overall loss severity than in Table 4 presented.

As in the POT approach, the losses above a threshold can be modelled by a GPD when the threshold is chosen appropriately. The extension in Chavez-Demoulin, Embrechts, and Hofert (2015) for the loss severity distribution further assumes that the distribution parameters of the GPD (ξ = shape parameter, β = scale parameter) can each be described by a function that depends on covariates. The covariate which we will use in the first analysis are the region of domicile, the industry, the size of the firms (measured by the number of employees), the time, and the cyber risk subcategory (only in the analysis of cyber risk losses). The definition of the functions that describe the parameters best (i.e., which combination of covariates estimates the respective parameter best) will be based on likelihood-ratio tests. These tests prove that including all five covariates yields significant improvements in the Log-likelihood values for the estimation of the shape and scale parameter.[29] The goodness-of-fit for the Chavez-Demoulin, Embrechts, and Hofert (2015) model, and its comparison with the original POT approach, is presented in Table C1. Although, the changes in the Log-likelihood values and the

[28] All estimations in this part of the appendix are based on an older dataset that included observations for the years March 1971 to September 2009, as it was done in Biener, Eling, and Wirfs (2015). However, we hope we can provide an idea how the approach works and we will proceed.

[29] Only exception from this rule is covariate size that is not included in the estimation of the shape parameter ξ. Including size for this parameter would lead to a poorer fit overall which is why we only look at size for the scale parameter β. In addition, in the analysis of non-cyber risk losses we exclude covariate 'cyber risk subcategory'.

AIC are rather small, the results in Table E1 indicate the appropriateness of the approach described in Chavez-Demoulin, Embrechts, and Hofert (2015).

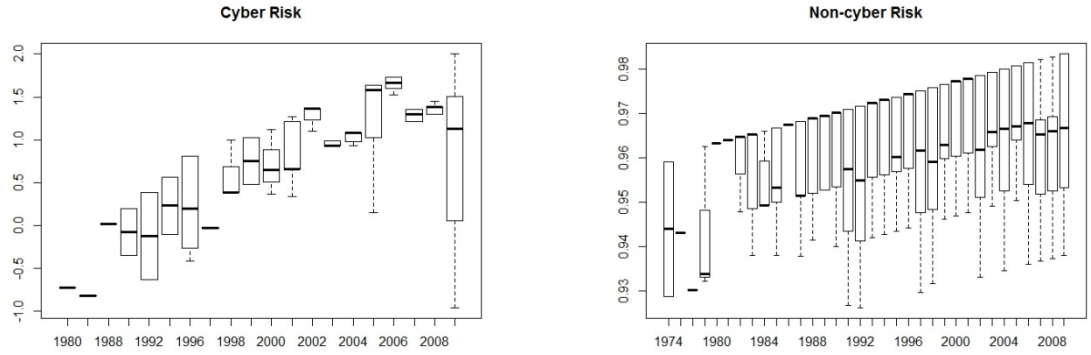
Table E1 Comparison POT and extended POT

Model	Cyber Risk (N = 1,579)		Non-Cyber Risk (N = 24,962)	
	Log-likelihood	AIC	Log-likelihood	AIC
POT (threshold 90%)	-990.76	1,985.52	-17,431.12	34,866.24
POT Extension (threshold 90%)	**-682.92**	**1,370.84**	**-14,597.50**	**29,199.00**

In the following analyses, we will particularly focus on the analysis of the shape parameter, since its value provides an indicator for the heaviness of the tail; the higher the parameter, the heavier the tail (see, e.g., Gilli and Këllezi, 2006). Thus, we are able to identify distributional differences – in terms of tail behavior – for different covariate combinations, which we then compare for cyber and non-cyber losses. In the following part we will show some preliminary results from this analysis.

The analysis of the shape parameter by time (see Figure E1) shows that in both risk categories – cyber risk and non-cyber risk – the tails become heavier. This shows the increasing importance of risk management not only for cyber risk but also for non-cyber risk. Moreover, Figure E1 also indicates that the distribution of the excesses in cyber risk seem to be heavier than in non-cyber risk (for the last 10 years the shape parameters in cyber risk were mostly above one, while those for non-cyber risk were below). This underlines the importance of cyber risk and the importance to think about the appropriateness of loss modelling approaches. Furthermore, it indicates that cyber losses are structurally different to other operational risks over time.

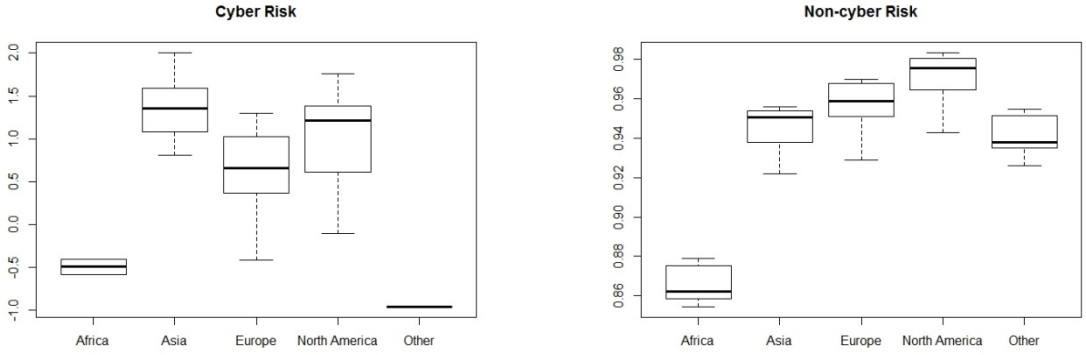
Figure E1 Comparison of the shape parameter over time



As for the descriptive analyses of Section 3.1, we look at the differences for the region of domicile (see Figure E2). We can observe that, in particular, Asia, Europa, and North America show heavier tails for the excess distribution in cyber and non-cyber risk than Africa and the other countries. Moreover, especially for Asia and North America be observe again higher

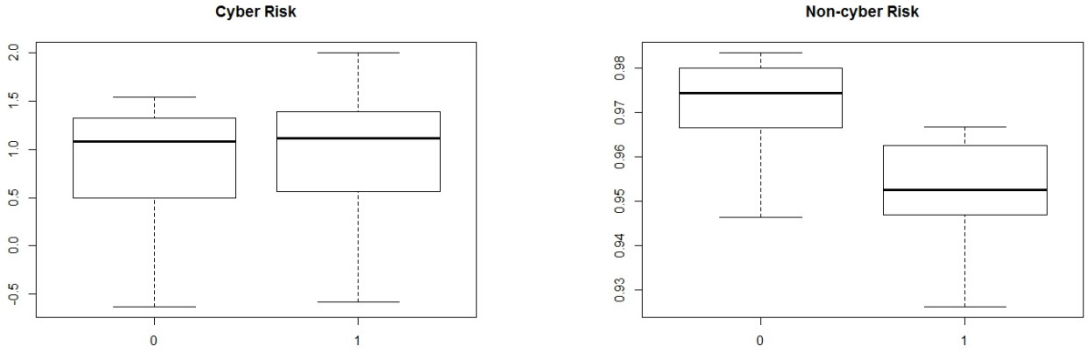
shape parameters in cyber than in non-cyber risk, underlining the findings from the analysis over time.

Figure E2 Comparison of the shape parameter by region of domicile



For the analysis of the shape parameter by industry (see Figure E3), we observe no significant distributional difference between firms from the financial services industry and those from the non-financial services industry in cyber risk. This is supervising because we identified differences for this covariate in the descriptive analysis (Table 3). For the non-cyber risks we prove that firms from the nonfinancial services industry exhibit heavier tails than those from the financial sector. This is in the line with the descriptive findings of Section 3.1.

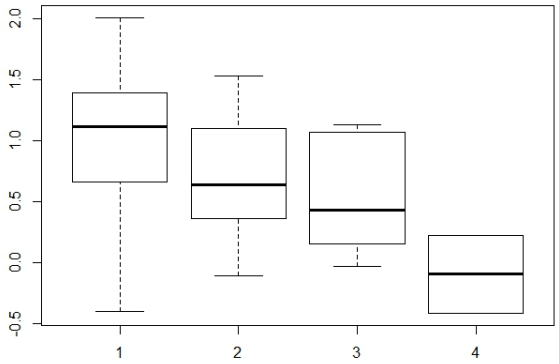
Figure E3 Comparison of the shape parameter by industry



0 = Nonfinancial industry, 1 = Financial industry

Finally, we discuss differences in the excess distribution by cyber risk subcategories. In Table 2 (descriptive analysis) we observe that human behavior is the main source of cyber risk, while the other three categories are less significant. A similar result can be found in the excess distribution analysis (see, Figure E4). The category ‘Actions of people’ is the one with the heaviest tails.

Figure E4 Shape parameter for cyber risk by subcategory



1 = Actions of people, 2 = Systems and technical failures,
3 = Failed internal processes, 4 = External events

The analyses in this appendix part are preliminary results and far from being complete (analysis of the scale parameter by covariates). However, they already provide interesting findings which will be elaborated and investigated further.

Appendix F: Analyses with Different Datasets

There exist other datasets connected with cyber risk, in particular on data breaches. For instance, the "Chronology of Data Breaches" from the Privacy Rights Clearinghouse (PRC). The PRC is an organization with the goal to engage, educate, and empower individuals to protect their privacy (PRC, 2016). The dataset they provide can be downloaded freely from their website and is regularly updated. The data sample consists of data breaches that occurred between January 10th, 2005 and December 15th, 2015 in the United States. After erasing all the observations under which no record was breached, we have a sample of 2,266 observations. Note further, that the data sample only contains number of records affected from data breach and not directly losses. However, this can be easily accounted for, because there exist average costs per breached record (see, e.g., Ponemon Institute, 2015; and the reports before). Since this is only a scaling of losses this will not directly affect the results from our distribution fitting.

The dataset here has been used in other academic works before; e.g., in Edwards, Hofmeyr, and Forrest (2015). Note, that the analysis in this part of the appendix is only for comparison of our finding for cyber losses generated from operational risks with another data source. The results might be different since the PRC data only focusses on data breaches and only incudes US data. However, we believe that the analysis can enrich our overall analysis.

As for the analyses of the SAS OpRisk Global data, we first provide a graphical illustration of the losses (see Figure F1) and present descriptive statistics (see Table F1). Similar to the results for the OpRisk data, we can observe skewness and kurtosis from the histograms. In addition, the results from the Q-Q plots for the original data show that the data breach losses in the PRC dataset are not normally distributed. However, the Q-Q plot for the logarithmic data shows a relatively good fit for the normal distribution.

Figure F1 Visualization of Original PRC Data (left) and Logarithmic PRC Data (right)

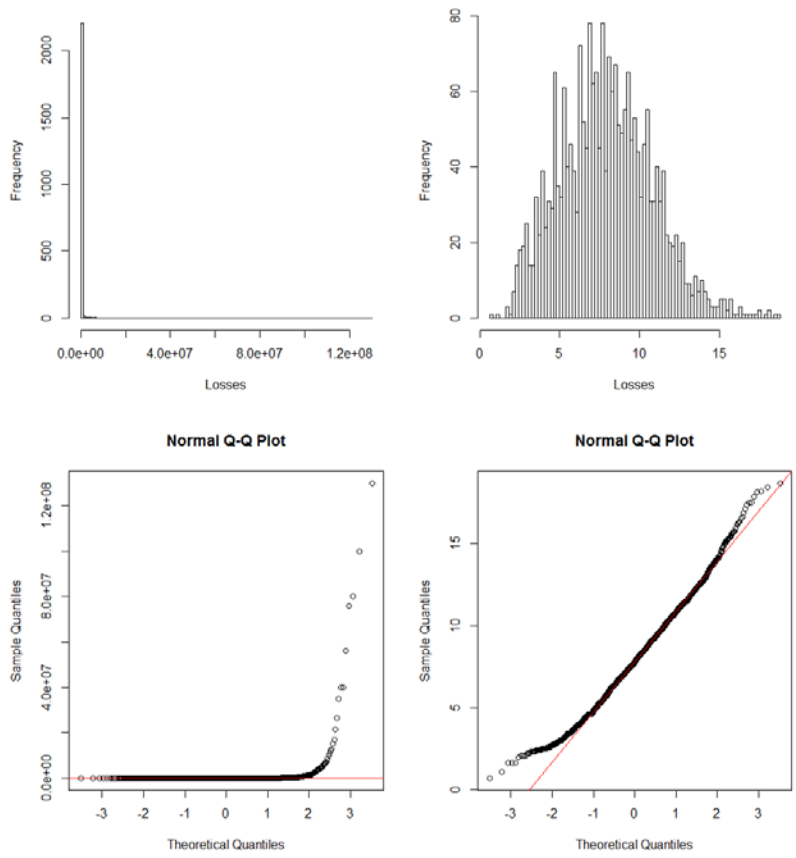


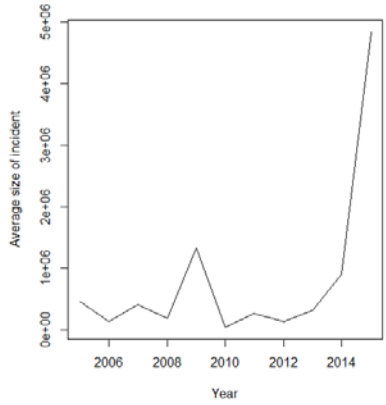
Table F1 Descriptive Statistics for PRC Data

	Original RPC	Logarithmic RPC
No. of Obs.	2,266	2,266
Mean	395,200.00	7.91
Std. Dev.	4,655,598.00	2.91
Skewness	19.60	0.33
Kurtosis	438.68	-0.01
Minimum	2.00	0.69
Maximum	130,000,000.00	18.68
VaR	330,000.00	12.71
TVaR	7,408,812.00	14.38

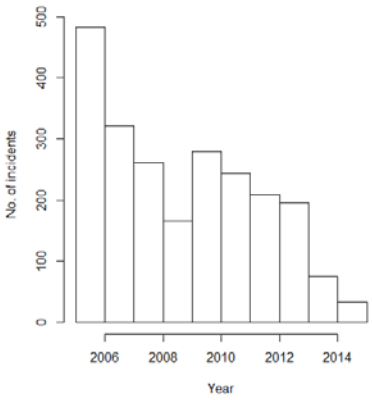
The results from the Q-Q plots and histograms can be underlined by the descriptive analysis. While the original dataset exhibits relatively high skewness and kurtosis parameters, these numbers are similar to the normal distribution for the log-RPC data. This indicates that the (original) PRC data is log-normally distributed. This is then also shown for the goodness-of-fit test in Table F2.

Firstly however, we will also look at the development over time for the PRC data (see Figure F2). The results for the average loss over time shows (with only one exception) relatively low, and constant average data breach losses per year. Only in the last year (2015) the average size per incident increased significantly.[30] The average number of incidents over the observation period however, seems to decrease. We found a similar result for the SAS OpRisk Global Data, which might be explained by better protection standards available than a few years ago.

Figure F2 Visualization of the Development over Time



(a) Average loss per year



(b) Average number of incidents per year

We also analyze the goodness-of-fit as in Table 4 for the OpRisk data. We already hypothesized by the descriptive statistics and visualizations that the RPC data might log-normally distributed. This results can be confirmed by the following goodness-of-fit analysis (see Table F2). The log-normal distribution is the only one for which the null hypothesis of the K-S test is not rejected (and for the A-D test the null hypothesis is only rejected at a 5% confidence level). For all other distributions, the fit is not as good as for the log-normal distribution. As for the OpRisk data, the POT approach with a log-normal body provides the best fit.

[30] This effect might be due to a relative small number of incidents in year 2015 (reporting reasons), and on very huge incident of the health insurer Anthem Inc., where almost 80 million records were breached.

Table F2 Goodness-of-fit Analysis

Model	Log-likelihood	AIC	Kolmogorov-Smirnov-Test		Anderson-Darling-Test	
Exponential	-31,468.30	62,938.60	0.73	***	73.73	***
Gamma	No convergence					
GPD	-23,647.95	47,299.90	0.07	***	18.76	***
Log-logistic	-23,599.71	47,203.42	1.00	***	4.34	***
Log-normal	**-23,572.74**	**47,149.48**	0.03		2.51	**
Weibull	-23,870.53	47,745.05	0.10	***	40.51	***
Skew-normal	-36,447.85	72,901.70	0.88	***	202.72	***
Skew-student	No convergence					
POT (threshold 90%)	**-23,133.81**	**47,133.81**	/		/	
Transformation Kernel	**-23,524.59**	/	/		/	