European Journal of Operational Research 272 (2019) 1109–1119

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Innovative Applications of O.R.

What are the actual costs of cyber risk events?

Martin Eling*, Jan Wirfs

Institute of Insurance Economics, University of St. Gallen, Girtannerstrasse 6, 9010, St. Gallen, Switzerland

ARTICLE INFO

Article history:
Received 20 September 2017
Accepted 10 July 2018
Available online 17 July 2018

Keywords:
Risk analysis
Cyber risk
Operational risk
Risk management
Insurance

ABSTRACT

Cyber risks are high on the business agenda of every company, but they are difficult to assess due to the absence of reliable data and thorough analyses. This paper is the first to consider a broad range of cyber risk events and actual cost data. For this purpose, we identify cyber losses from an operational risk database and analyze these with methods from statistics and actuarial science. We use the peaks-over-threshold method from extreme value theory to identify "cyber risks of daily life" and "extreme cyber risks". Human behavior is the main source of cyber risk and cyber risks are very different compared with other risk categories. Our models can be used to yield consistent risk estimates, depending on country, industry, size, and other variables. The findings of the paper are also useful for practitioners, policymakers and regulators in improving the understanding of this new type of risk.

1. Introduction

Although cyber risk is a crucial topic for the economy and society and is reported in the media every day, it has been the subject of very limited academic research.[1] This is most likely due to the absence of reliable data. The aim of this paper is to take one step forward by conducting a thorough empirical analysis of cyber risks. While existing literature is limited to the consideration of the number of records lost in a data breach, this paper is the first to consider the whole range of cyber risks (not only data breaches) and to consider cost information (not only the number of records lost). For this purpose, we extract 1579 cyber risk incidents from an operational risk dataset and analyze them with methods from the field of statistics and actuarial science. We thus contribute to the growing literature on the use of stochastic modeling techniques in operations research (see Laengle et al., 2017).

To arrive at deeper insights into the nature and statistical properties of cyber risk, we apply the actuarial toolbox. Specifically, we analyze whether models, which prove useful for other loss categories, can also be applied to cyber risk. With this, we test whether cyber risks are structurally comparable to or distinctive from other risks. Answering this question is essential to improve the modeling and management of cyber risk, for example to calculate risk capital or to improve the provision of insurance. The consideration of cost information is much more relevant to business and leads to a different evaluation of cyber risk compared to the consideration of non-monetary information.

Our main results can be summarized as follows: Using the peaks-over-threshold method from extreme value theory, we identify what we call the "cyber risks of daily life" and "extreme cyber risks".[2] Especially given the magnitude of daily life cyber events, these risks are significantly less skewed and less heavy-tailed than other operational risks. At the same time, a few extreme cyber risks pose an enormous threat to the solvency of the afflicted

* Corresponding author.
E-mail addresses: martin.eling@unisg.ch (M. Eling), jan.wirfs@unisg.ch (J. Wirfs).

[1] The literature on cyber risk is limited to the field of IT and information security (e.g. Evans, Maglaras, He, & Janicke, 2016), but relatively little work has been done in business, finance, and economics. Our paper is closest to Biener, Eling, and Wirfs (2015) who analyze the insurability of cyber risk and present descriptive statistics to illustrate their statistical properties. Our paper is also related to the data breach analyses of Edwards, Hofmeyr, and Forrest (2015), Maillart and Sornette (2010), and Wheatly, Maillart, and Sornette (2016). Table A.1 in the Online-Appendix summarizes these papers and outlines the contribution of this paper. We build upon and extend these papers in that we analyze (1) actual costs instead of number of records affected by data breaches (2) a global dataset (instead of US data only), (3) a longer time period (1995–2014), and (4) the whole range of cyber risks (not solely data breaches). Furthermore, we include more explanatory variables (e.g., risk type, contagion, region, industry, company size, and time), derive a model under which the simultaneous analysis of all covariates is possible (i.e., interactions between covariates), and analyze the whole loss distribution (not only peaks over a threshold). Recently, Romanosky (2016) provides another attempt to quantify the costs of cyber events considering U.S. data from Advisen; he mainly presents descriptive statistics that we use to validate and verify the plausibility of our numbers; moreover, he presents a logistic regression model to analyze the costs of cyber events, but for data breaches only.

[2] This separation follows Rakes, Deane, and Rees (2012) who motivate their model with the fact that IT security incidents typically result in small operational disruptions or minimal recovery costs, but occasionally high-impact security breaches can have catastrophic effects on the firm. An example of a small cyber risk of daily life is a hacker attack on companies reported in the media almost every day. Extreme cyber risks, however, are major disruptions of the business operations that might threaten the survival of a company.

https://doi.org/10.1016/j.ejor.2018.07.021
0377-2217/

company. Our results show that cyber risk constitutes a distinct risk category where more research is warranted, especially given their high importance for the economy and society. The results also highlight that human behavior, be it criminal or not, is the main source of cyber risk. Another relevant result for risk managers, insurers and policymakers is that cyber risk are extreme, but not as extreme as the other operational risk in our dataset.

Our findings are important to reach a better understanding of cyber risks and their consequences. Special importance is given to the financial services sector, since regulators require banks and insurers to hold risk capital for operational losses, which might result from cyber risks. Moreover, our results are useful for insurance companies that are developing cyber insurance policies and do not have enough data and experience with cyber risks.[3] For firms outside the financial services sector, the results are important not only for internal risk management, but also in light of new reporting requirements for cyber incidents.[4] We illustrate the usefulness of our results for policymakers, regulators and practitioners in a risk measurement application. For the academic audience we present effective and contemporary modeling approaches for the novel application area of cyber risk. Our paper thus provides an innovative application of OR methods that is relevant not only to the finance and insurance domain (e.g., Adcock & Meade, 2017; Boonen, 2016), but to risk management in general (e.g., Aven, 2016). Moreover, it contributes to the growing literature on cybersecurity in the business domain (e.g. Nagurney & Shukla, 2017).

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

2. Data and methods

2.1. Data

Typically, information on cyber risk is not publicly available since companies with security breaches or that have been attacked do not report it. Another impediment to the collection of cyber risk data is the absence of a clear-cut definition of "cyber risk". Our definition 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".[5] Linking cyber risk to operational risk has several advantages. Firstly, it distinguishes cyber risk from other established risk categories (market, insurance, credit, liquidity, legal, operational risk; see BIS, 2006; CEIOPS, 2009). Secondly, in structuring cyber risk we can use the established subcategories from operational risk (actions of people, systems and technology failures, failed internal processes, and external events). Thirdly, linking cyber risks to operational risks allows a clear identification of data.

The latter advantage 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 Online-Appendix B. The database consists of 26,541 observations between January 1995 and March 2014.[6] An important aspect is the reliability and completeness of the dataset. Regarding reliability, each loss event has been confirmed at least by one major media source and is thus traceable and peer-reviewable. The dataset has already been used in numerous academic papers (e.g., De Fontnouvelle, Dejesus-Rueff, Jordan, & Rosengren, 2006; Hess, 2011) and is widely accepted in practice (e.g., regulators allow banking and insurance companies to enrich their own data with information from the dataset to calculate risk-based capital). In terms of completeness, one limitation is that the data provider only includes losses in excess of US$100,000; in general it holds that the larger the event, the more likely it is to be reported.[7] However, for the analysis of tail behavior, which is of principal interest in this paper, censorship is no problem. When applying extreme value theory the empirical distribution converges to the GPD, if the threshold is chosen appropriately; thus there is no bias in the tail fit. Another indication for the reliability and completeness is that the data reflects many characteristics known from industry studies (e.g., Ponemon Institute, 2015; PwC, 2015), but allows to dig much deeper into the topic.

In additional tests available in the Online-Appendix, we also compare our data and results with the data breach data considered by Edwards, Hofmeyr, and Forrest (2015) and Wheatley, Maillart, and Sornette (2016) that is we analyze the PRC data breaches with our methodology (Online-Appendix DIV) and test the methodology of Edwards et al. (2015) on our cyber risk data (Online-Appendix DV). We also filtered out data breaches (according to the PRC, 2017, definition) from our cyber risk data, resulting in about 25% of our events being classified as data breaches. The evaluations with this subsample show that (in contrast to Edwards et al., 2015, and Wheatley et al., 2016) also data breaches should be analyzed with extreme value theory. This is an important result, because it shows that the results documented for the number of lost data do not hold when actual costs are considered, emphasizing the relevance of the analysis presented here.

[3] Industry studies show that the modeling and pricing of cyber insurance policies is the main impediment to the insurability of cyber risks; see Biener et al. (2015) and Eling and Wirfs (2016).

[4] In the US, reporting requirements for data breaches have been introduced in many states since 2002 (NCSL, 2016). In the European Union, such reporting requirements will apply from 2018 (European Union, 2016). If these are enforced, more data and information will be available. This has happened already in the US with data samples as the Privacy Rights Clearinghouse "Chronology of Data Breaches".

[5] Our definition of cyber risk leads to a broader set of events than the existing data breach analyses based on the Privacy Rights Clearinghouse (PRC) dataset (e.g., Edwards et al., 2015; Wheatley, Maillart, & Sornette, 2016). Data breaches describe a "security violation in which sensitive, protected or confidential data is copied, transmitted, viewed, stolen or used by an individual unauthorized to do so" (PRC, 2017). While PRC only considers incidents where data became public (i.e., in which the confidentiality of the data is violated), we also consider the availability and integrity of data (i.e., destroyed, lost, or manipulated). Our definition further includes all effects on information systems and spillover effects to operational technology (e.g., business interruption) and not just the effect of the data breach itself. Data breaches are thus a subgroup of the overall cyber risk potentials that we analyze in this paper. Important is also that we analyze actual cost numbers while existing papers are restricted to non-monetary information (only the number of lost data is considered). Reputational effects are not incorporated in both definitions.

[6] The dataset provides an estimate of the complete costs of operational risk events (direct and indirect); however, reputational losses (and thus loss of potential business) are not covered since this sort of loss is typically excluded in regulation of operational risk. All losses are given in US$ and adjusted for inflation to make them comparable. The reported loss amount consists of six different loss categories (legal liability, regulatory action, loss or damage to assets, restitution, loss of recourse, write-down) and are before insurance. The data can be broken down into seven event categories (business disruption and system failures; clients, products and business practices; damage to physical assets; employment practices and workplace safety; execution, delivery and process management; external fraud; internal fraud). Cyber losses before the internet era are mainly data breaches or business interruptions due to computer system failures.

[7] Another limitation might be that early data might be backfilled and late data might be incomplete because the outcomes are not fully reported. To account for this we evaluate our findings with a subsample of losses from 2004 to 2013 and found no significant differences (results are available in Online-Appendix DV). This effect might also be small, because a team of seasoned operational risk research analysts maintains the database in accordance with strict data quality standards and reviews it periodically to provide updates to older loss events that change over time and to ensure accuracy and completeness.

2.2. Methodology

We adopt the loss distribution approach – the most common method for actuarial modeling – which considers two separate distributions for loss frequency and loss severity (see, e.g., Panjer, 2006). The losses L are described by:

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

where the frequency N is a discrete random variable and $X_1, \ldots, X_N$ are positive independent and identically distributed random variables. Therefore, a loss frequency distribution (random variable N) and a loss severity distribution (random variables X_i; equal to the first modeling case) are fitted separately and then combined to an aggregated loss distribution (see, e.g., McNeil, Frey, & Embrechts, 2015). In most models used in academia and practice, there are no closed-form formulas for the aggregates so that the aggregation is approximated by Monte Carlo simulation.

Loss frequency is typically modeled either by a Poisson or a negative binomial distribution (see, e.g., Panjer, 2006, Chapter 5; Boucher, Denuit, & Guillén, 2008; Liu & Pitt, 2017). Both distributions belong to the class of discrete probability distributions and describe random and independent events. The negative binomial distribution has the advantage over the Poisson distribution that it is more flexible in shape by allowing two parameters to be estimated.

For the loss severity we fit the data to distributions used in recent actuarial literature (e.g., exponential, Gamma, log-normal, log-logistic, generalized Pareto distribution (GPD), and Weibull; see, e.g., Eling, 2012). Furthermore, we include a non-parametric transformation kernel estimation (see Bolancé, Guillen, & Nielsen, 2003, 2008) and implement the peaks-over-threshold (POT) method from EVT (see, e.g., Chapelle, Crama, Huebner, & Peters, 2008). In the latter approach, losses above a threshold are modeled by a GPD, while losses below the threshold are modeled with another common loss distribution such as the exponential, log-normal or Weibull.

The POT approach was developed especially for random variables characterized by extreme values; we thus expect this approach to show a better fit for the cyber risk data compared to other standard models mentioned above. The POT 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 & de Haan, 1974). The *GPD* (ξ, β), with shape parameter $\xi \in [-\infty; \infty]$ and scale parameter $\beta > 0$, is defined by the distribution function:

$$GPD_{\xi,\beta}(x) = \begin{cases} 1 - \left(1 + \dfrac{\xi \cdot x}{\beta}\right)^{-1/\xi}, & if\ \xi \neq 0, \\ 1 - \exp\left(-\dfrac{x}{\beta}\right), & if\ \xi = 0, \end{cases} \tag{2}$$

for $x \geq 0$, if $\xi \geq 0$, and $x \in [0, -\beta/\xi]$, if $\xi < 0$. The body distribution is then fitted on one of the simple parametric distributions discussed above, such as exponential (Hess, 2011) or log-normal (Moscadelli, 2004). In our analysis we test different body models and apply them to different threshold values (see Online-Appendix DII). To identify the best models we apply various goodness-of-fit tests (log-likelihood value, the AIC, and the Chi-square goodness-of-fit tests; Anderson-Darling-test; modified Kolmogorov-Smirnov-test for discrete distributions; see Arnold & Emerson, 2011). We use the bootstrap goodness-of-fit test by Villaseñor-Alva and González-Estrada (2009) to identify the optimal threshold value for the POT approach.[8]

In Section 3.3 we go beyond the standard actuarial modeling and apply a new version of the POT method following Chavez-Demoulin, Embrechts, and Hofert (2016) where the loss data depends on covariates. Recent literature on operational risks has shown that the inclusion of covariates enhances model fit and thus improves estimation accuracy (see, e.g., Ganegoda & Evans, 2013). Instead of separating the data into subsamples and modeling the losses in each group individually (see, e.g., Edwards et al., 2015), this approach has the advantage of using the whole data (which is critical for the sparse cyber risk data) and that interactions among covariates can be analyzed. In the dynamic EVT model the distributional parameters depend on covariates. Let $\theta = (\lambda, \xi, \beta)$ be the POT model parameter vector. Due to simplicity and a better interpretation of the distributional parameter, we model the loss frequency by a Poisson distribution with intensity parameter λ. The severity is modeled by a GPD with the parameters ξ (shape) and β (scale). Each of these parameters is defined by a linear combination of three incident-specific variables (cyber subcategory, contagion A (multiple firms affected), contagion B (multiple losses caused)), three firm-specific variables (region of domicile, industry, company size) and time. This yields the following formulas for λ and ξ:

$$\begin{aligned} \ln(\lambda(X,t)) = {} & \beta_0 + \sum_{i=1}^{3} \beta_i \cdot X_i^{Dummy-Subcategories} \\ & + \beta_4 \cdot X^{Dummy-Contagion\ A} + \beta_5 \cdot X^{Dummy-Contagion\ B} \\ & + \sum_{i=1}^{4} \beta_{i+5} \cdot X_i^{Dummy-Domicile} + \beta_{10} \cdot X^{Dummy-Industry} \\ & + \beta_{11} \cdot X^{CompanySize} + \beta_{12} \cdot t, \end{aligned} \tag{3}$$

$$\begin{aligned} \xi(X,t) = {} & \beta_0 + \sum_{i=1}^{3} \beta_i \cdot X_i^{Dummy-Subcategories} + \beta_4 \cdot X^{Dummy-Contagion\ A} \\ & + \beta_5 \cdot X^{Dummy-Contagion\ B} + \sum_{i=1}^{4} \beta_{i+5} \cdot X_i^{Dummy-Domicile} \\ & + \beta_{10} \cdot X^{Dummy-Industry} + \beta_{11} \cdot X^{CompanySize} + \beta_{12} \cdot t, \end{aligned} \tag{4}$$

with $X = (X^{Dummy\text{-}Subcategories}, X^{Dummy\text{-}Contagion\ A}, X^{Dummy\text{-}Contagion\ B}, X^{Dummy\text{-}Domicile}, X^{Dummy\text{-}Industry}, X^{CompanySize})$ being a matrix of covariates, and t being the time at which the incident occurred. The selection of covariates is motivated by existing literature (see Online-Appendix C).

To ensure convergence, the estimation approach for the GPD's scale parameter β requires an orthogonally transformed parameter (see Chavez-Demoulin et al., 2016), such that the scale parameter cannot be directly connected to the covariates. The transformation that is estimated in the algorithm is defined by:

$$\nu(X,t) = \ln\left((1 + \xi(X,t)) \cdot \beta(X,t)\right), \tag{5}$$

with X being the covariates and t the time. Thus, the scale parameter is defined by:

$$\beta(X,t) = \frac{\exp(\nu(X,t))}{(1 + \xi(X,t))}, \tag{6}$$

and by that already depends on covariates if the shape parameter does (see Chavez-Demoulin et al., 2016). Effects of the covariates

[8] Several selection approaches have developed in literature; e.g., graphical methods like the mean excess and Hill plots (see Dutta and Perry, 2007, for a detailed listing of models). To identify the optimal threshold we use the bootstrap goodness-of-fit test by Villaseñor-Alva and González-Estrada (2009). We select the lowest possible threshold u^{low}, such that the data above u^{low} can still be modeled by a GPD. Furthermore, we choose exemplary thresholds (higher than u^{low}) and compared the model fits by their log-likelihood values (see Online-Appendix DII). The results show that the choice of u^{low} is appropriate.

on β are thus difficult to evaluate if also the transformed parameter ν depends on covariates. To account for this, the relationships can be plotted (plots are available upon request). For completeness we also allow the orthogonally transformed parameter to depend on covariates:

$$\nu(X,t) = \beta_0 + \sum_{i=1}^{3} \beta_i \cdot X_i^{Dummy-Subcategories} + \beta_4 \cdot X^{Dummy-Contagion\ A} + \beta_5 \cdot X^{Dummy-Contagion\ B} + \sum_{i=1}^{4} \beta_{i+5} \cdot X_i^{Dummy-Domicile} + \beta_{10} \cdot X^{Dummy-Industry} + \beta_{11} \cdot X^{CompanySize} + \beta_{12} \cdot t. \quad (7)$$

Chavez-Demoulin et al. (2016) implement the described approach in the R packages mgcv and QRM. In their paper, the authors restrict the analysis to the distribution of the loss excesses (i.e., the POT). As our analyses show that the GPD provides the best fit for the single parametric functions, we will also apply the dynamic EVT approach to the whole loss range (i.e., a threshold $u = 0$). This enables comparisons of the new findings with the results from standard models presented in Section 3.2. Afterwards, we apply the approach as it was originally done in Chavez-Demoulin et al. (2016) that is with a threshold of $u = u^{low}$. If the excess distribution under Chavez-Demoulin et al. (2016) approach provides a better fit than the excess distribution under the normal POT approach, we can improve the modeling of the excess and of the overall loss severity.

To ensure that the inclusion of specific covariates increases model fit, we apply the selection process used in Chavez-Demoulin et al. (2016). For the loss frequency model, we sequentially incorporate each variable into the model, and determine the extent to which the model fit improves considering likelihood ratio tests. Only if a covariate increases the overall model fit the variable is included in the final model setup. For the loss severity distribution, we include the covariates describing the shape parameter ξ of the GPD first (analogous to Chavez-Demoulin et al., 2016). We do so, because the shape parameter indicates the heaviness of the tail (the higher the parameter, the heavier the tail), which identifies distributional differences in tail behavior for different covariate combinations. The shape parameter ξ is also easier to interpret than the scale parameter β. After the identification of the optimal shape-model, we include the covariates for the orthogonally transformed scale parameter ν.

Finally, in Section 3.4 we estimate the value at risk, a risk measure that banking and insurance regulators use to determine how much capital a company needs to hold to cover losses with a given confidence level (e.g. in Basel II and Solvency II; see Eling & Tibiletti, 2010). Given that the LDA is applied with a loss frequency modeled by a Poisson distribution and the loss severity modeled by the POT with a GPD in the tail, the risk measure VaR at level α can be computed by the following closed-form formula:

$$VaR_\alpha = u + \frac{\hat{\beta}(X,t)}{\hat{\xi}(X,t)} \left(\left(\frac{1-\alpha}{\hat{\lambda}(X,t)/n'} \right)^{-\hat{\xi}(X,t)} - 1 \right), \quad (8)$$

with $\hat{\lambda}$, $\hat{\xi}$ and $\hat{\beta}$ being the parameter estimates from the dynamic EVT approach which depend on covariates X and time t, u being the threshold in the POT and $n' = n'(X, t)$ is the total number of losses for a fixed covariate X and time point t (see, e.g., McNeil, Frey, and Embrechts, 2005, p. 283). The advantage of our approach is that one obtains a closed form solution of the parameters. From this we can directly derive quantiles of the distribution and thus Value at Risk, also in closed form. For Monte Carlo complex simulations are necessary and it remains just an approximation.

3. Results

3.1. Descriptive statistics

Table 1 compares 1579 cyber losses extracted from the SAS OpRisk database with the 24,962 cases not classified as cyber. Mean, standard deviation, median, skewness, and kurtosis are significantly smaller for cyber than for the non-cyber losses (Table 1, Panel A).[9] The results of a non-parametric Mann-Whitney-Wilcoxon Test show that the two samples come from different populations, illustrating that cyber risks are different. [10] The tail risk measure (Tail Value at Risk/Value at Risk at a predefined confidence level; see McNeil et al., 2015, p. 283), however, is higher for cyber risk than for non-cyber losses. It thus seems that there is a large number of small losses (the "cyber risks of daily life") and a few large ones (the "extreme cyber risks") leading to high Tail Value at Risk (TVaR) values and a higher tail risk measure. This aspect will be considered in more detail in the later analyses. The mean loss (US$ 43.49 million) is also substantially higher than the median loss (US$ 1.53 million), illustrating the skewness of the loss distribution. This finding is important, because it indicates that mean loss numbers presented in a few industry studies (e.g. Ponemon Institute, 2015) should not be interpreted without further information such as median or skewness.[11]

The separation into subcategories (Panel B) shows that "actions of people" is the most frequent incident, while the other categories are rather rare. Hacking attacks, physical information thefts, human failures, and all incidents where employees manipulate data (un-/intentionally) are included here. This finding confirms results from the IT literature (see, e.g., Evans, Maglaras, He, & Janicke, 2016) that human behavior is the main driver of cyber risk. Regarding Panel C it might seem counterintuitive that the 296 cases in which more than one company is affected do not cause higher losses, but if several companies are affected the incident might be detected sooner and economies of scale in solving the problems might be realized. If one incident causes multiple losses in the same firm, however, the costs are higher (Panel D). Note that the rate at which the risks spread could make cyber risks structurally different from other risks. Since the focus is on costs, it is possible that the cyber costs may not be as high when compared to non-cyber risks, however, more assets could have got compromised (to a lesser extent). Five out of the ten largest cyber losses come from Asia, increasing the mean loss for this region (Panel E).

Moreover, a high portion of incidents occurred in the financial industry (76%; Panel F), while it is more balanced for other operational risks. This illustrates that the financial industry might be an especially attractive target, though obviously better protected

[9] This holds also for the maximum losses, which are US$ 14.6 billion for cyber risk and US$ 97.7 billion for non-cyber risk. The maximum loss incident for cyber risk is a money-laundering incident at the Bank of China in February 2005. The largest non-cyber risk incident was in November 2001 when Philip Morris had to pay smokers who became sick from their products. The Bank of China case is the 23rd largest loss in the dataset, illustrating that there are many other big operational risk events, which are not in the cyber risk category.

[10] We also conducted the Levene- and Fligner-Killeen-Test to test for the homogeneity in variances and again observe a significant difference. Furthermore, a Kolmogorov-Smirnov-Test for the empirical distributions for cyber and non-cyber risks indicates that the two samples are different. The results for these additional tests are available upon request.

[11] 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. Romanosky (2016) comes to an average of US$ 7.8 million with a median of US$ 0.25 million and a standard deviation of US$ 47.28 million. Overall, it seems that our values are higher than those known from other studies. One part of the difference might be explained by the fact that the data provider only includes losses in excess of US$100,000. Another explanation is that a few very large non-US losses are part of our dataset (e.g. the above-described Bank of China case).

Table 1
Comparison of cyber risk and non-cyber risk losses (in million US$).

	Cyber risks							Non-cyber risks						
	N	Mean	Std. dev.	Median	Skewness	Kurtosis	Tail risk	N	Mean	Std. dev.	Median	Skewness	Kurtosis	Tail risk
Panel A: Total sample														
Total	1579	43.49	426.36	1.53	27.12	873.33	7.265	24,962	98.52	1154.39	5.09	49.95	3388.68	5.765
Panel B: Subcategories														
Actions of people	1203	42.66	475.53	1.35	25.42	739.22	9.559							
Systems and technical failure	212	45.32	141.23	4.78	7.87	81.21	2.086							
Failed internal processes	108	15.12	48.96	1.32	5.30	30.90	2.745							
External events	56	109.12	431.92	4.25	5.49	31.56	4.789							
Panel C: Relation of one incident to losses in other firms (Contagion Type A)														
One firm affected	1283	49.22	470.84	1.56	24.73	720.95	6.835	17,748	87.59	983.16	5.02	48.44	3212.91	5.382
Multiple firms affected	296	18.65	90.74	1.45	11.35	154.13	4.083	7214	125.40	1494.10	5.30	45.74	2693.57	6.235
Panel D: Relation of one incident to losses in the same firm (Contagion Type B)														
One single loss caused	1426	39.33	437.28	1.43	27.58	873.25	8.109	22,533	81.55	1140.32	4.45	55.43	3907.68	6.264
Multiple losses caused	153	82.21	304.82	7.51	7.47	64.70	3.264	2429	255.90	1267.00	19.65	13.35	226.28	3.966
Panel E: Region of domicile														
Africa	24	30.90	131.99	1.86	4.29	17.17	14.264	278	58.72	286.60	2.59	9.34	96.98	3.828
Asia	256	104.31	942.39	1.52	14.32	215.39	14.255	3375	132.95	1828.87	4.04	34.00	1330.99	8.019
Europe	393	31.09	126.56	1.78	8.12	84.02	3.144	5596	121.01	892.16	5.49	19.40	447.42	4.781
North America	830	33.26	250.90	1.42	15.55	264.37	5.575	14,867	85.31	1083.29	5.27	60.22	4800.16	5.883
Other	76	18.44	71.46	1.55	7.50	59.05	1.246	846	57.44	231.27	4.47	9.14	103.85	3.763
Panel F: Industry														
Non-financial	381	84.11	408.49	4.47	8.62	82.87	4.944	13,665	114.31	1449.82	7.43	44.27	2459.44	6.576
Financial	1198	30.57	431.26	1.16	32.23	1081.47	7.384	11,297	79.40	633.58	2.92	25.53	857.65	4.980
*Panel G: Company size by number of employees**														
Small	506	36.46	251.79	1.45	15.66	288.87	6.796	7310	49.64	473.25	3.60	41.79	2120.66	5.048
Medium	504	25.86	154.55	1.52	16.16	308.05	3.805	7310	90.62	1015.99	5.70	38.04	1614.73	4.730
Large	505	67.91	691.61	1.44	18.83	384.22	11.191	7130	175.63	1787.05	6.62	37.84	1825.05	5.717
NA	64	45.19	141.89	4.67	5.40	32.58	3.264	3032	49.47	470.72	5.36	33.99	1444.98	5.869

Note: The non-cyber losses are the losses not classified as cyber in the SAS OpRisk Global dataset. The tail risk measure (Tail Value at Risk/Value at Risk at 95% confidence level) describes the extent to which the TVaR captures the tail risk that is not gauged by the VaR. *: Size-classification is based on the lower, middle and upper 33% quantiles of number of employees; Cyber risk – Small ($\leq$5882 employees), medium (between 5900 and 56,137 employees), and large ($\geq$56,200 employees); Non-cyber risk – Small ($\leq$5956 employees), medium (between 5960 and 47,979 employees), and large ($\geq$48,000 employees). NA denotes firms where the number of employees is unavailable, representing mainly relatively old cases.

(mean and median loss are much lower). Finally, in Panel G we observe a U-shaped relation between the loss amount and the number of employees as well as a U-shape in the tail risk measure, indicating heavier tails for small- and large-sized companies. Additional tests available upon request show however that these observations are driven by a few very large losses and are not significant in a multivariate regression model.

Fig. 1(a) illustrates that the number of cyber risk incidents was small before 2005. After that point, however, the number of incidents continuously increased until May 2009 and then decreased slightly. A similar pattern can be observed in Maillart and Sornette (2010), who analyzed data breaches in the US between 2000 and 2008, indicating a pattern change as early as July 2006. The mean and median loss per event (see Fig. 1(b) and (c)) and their variance have since decreased, perhaps because of the increased use of self-protection measures that reduce the loss amount in a cyber incident. The developments of cyber and non-cyber losses are comparable, but mean and median losses for non-cyber risk exhibit a higher variance. The low median for cyber again illustrates that the majority of the cyber losses are small events of daily life and that the risk profile is different from that of other risk categories.

We also tested for serial dependence between the individual 1579 loss events and found the autocorrelation at lag 1, 2, and 3 to be close to 0 (0.005, 0004 and −0.005). We also tested for serial dependence in the monthly losses and added the Augmented Dickey-Fuller test, again indicating that there is no serial dependence. Moreover, we checked the correlation between the monthly losses from the cyber and non-cyber data, which is 0.02. It thus seems that the monthly losses from the cyber and non-cyber data are not correlated.

3.2. Standard models and goodness-of-fit

The log-likelihood function and AIC values indicate that for the loss frequency – both measured by year and month – the negative binomial distribution provides better fit than the Poisson distribution (see Table 2). This is underlined by the results of the K-S test: the null hypothesis for the Poisson distribution is rejected at a 1% level, while it cannot be rejected for the negative binomial distribution for the yearly data. We also test the zero-inflated versions of the Poisson and negative binomial distribution (see, e.g., Edwards et al., 2015), which do not provide a better fit than the negative binomial distribution for frequency data measured on a yearly basis. This indicates that modeling the yearly loss frequency of cyber risks by a negative binomial distribution is most appropriate. This result is in line with the findings of Edwards et al. (2015), who identified the negative binomial distribution to be optimal for the daily frequencies of data breaches and also provide an explanation (the negative binomial can be generated by a mixture of different types of breaches, with each type occurring at a different, but constant rate).

For the loss severity (Table 3) the results from the Kolmogorov-Smirnov-tests (K-S-tests) indicate that none of the single parametric distributions adequately models the loss data (the null hypothesis for each model in the K-S-test is rejected). Furthermore, these distributions do not fit the non-cyber risk data, necessitating the use of modeling approaches from EVT. From the seven distributions, the GPD provides the best results for cyber losses measured by log-likelihood values and AIC. This finding is in contrast to that of Edwards et al. (2015), who estimate the log-normal distribution as best model for data breach sizes. These observed differences are an important result, because they indicate that relying on the

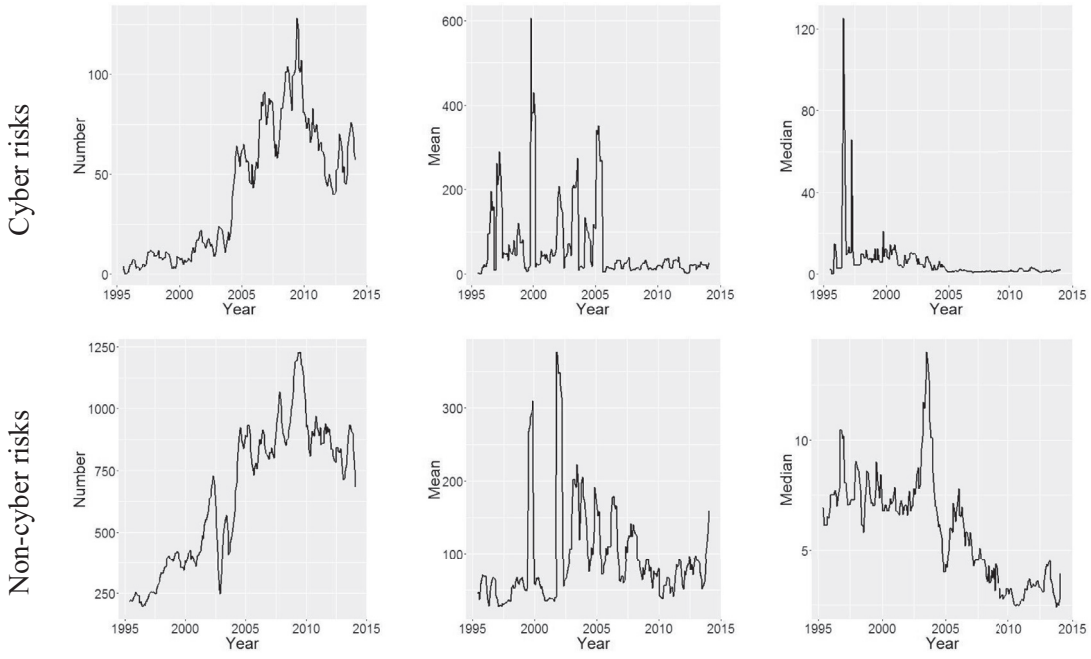

Fig. 1. Number of loss events, mean and median losses over time (in million US$).

Table 2
Goodness-of-fit analysis – frequency.

Model	Frequency	Log-likelihood	AIC	Chi-square-test	K-S-test
Panel A: Cyber risk (N = 1579)					
Poisson	Yearly	−622.18	1246.36	1130.08***	0.500***
Poisson	Monthly	−958.49	1918.97	1261.22***	0.611***
Negative binomial	Yearly	**−107.39**	**218.78**	100.49***	0.218
Negative binomial	Monthly	**−345.57**	**695.15**	35.40**	0.699***
Panel B: Non-cyber risk (N = 24,962)					
Poisson	Yearly	−3267.56	6537.12	> 10000.00***	0.500***
Poisson	Monthly	−10327.28	20656.55	19719.11***	0.661***
Negative binomial	Yearly	**−157.08**	**318.16**	194.33***	0.235
Negative binomial	Monthly	**−651.46**	**1306.93**	367.49***	0.627***

Note: AIC = Akaike information criterion; for the Chi-square- and the Kolmogorov-Smirnov-test (K-S-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). *, **, and ***, indicate significance levels of 10%, 5%, and 1%.

log-normal distribution often used in actuarial practice might misestimate the costs of cyber risks. This will also be illustrated in the numerical application (Section 3.4) when we compare risk measurement results for the log-normal and the POT. We also note that the difference between the log-normal and POT are bigger for the cyber risks compared to the non-cyber risks.

The GPD shape parameter for the cyber risk data has a value of 1.5998, which is smaller than that for non-cyber risk (1.6385). The distribution for the non-cyber risk losses is thus heavier than that of cyber risks. Looking at the POT approach, we observe a better fit for both cyber risk and non-cyber risk compared with the single parametric distributions,[12] also the non-parametric transformation kernel performs very well in fitting the data, providing the overall best fit (which again confirms findings from other papers; see, e.g. Eling, 2012).

The distributions estimated with the POT approach are presented in Fig. 2. These again illustrate that cyber risk is different from other operational risks; the distribution function is quite a bit away from the non-cyber risks, and approaches one much faster than the non-cyber losses, indicating less heavy tails. The visual inspection also again illustrates the large number of daily life events and confirms the good fit of the POT approach, motivating the use of EVT and its dynamic extension. The POT approach from EVT is thus useful to highlight the difference between the cyber risks of daily life (those in the body of the distribution) and the extreme cyber risks (those in the tail). The statistical results (Table 3) also illustrate that the POT provides a better goodness-of-fit compared with other candidates, especially also compared with the log-normal distribution widely used in insurance

[12] The optimal threshold value is the 56% percentile for the cyber and the 80% percentile for the non-cyber sample; for robustness, we also computed results for different threshold values, yielding similar findings. See Online-Appendix DII. Note also that we do not show K-S- and A-D-results for the POT approach in Table 3, since the models are not continuous at the threshold, which is a prerequisite for the application of those tests (see, e.g., Lehmann & Romano, 2005, p. 584). It is possible to estimate models with continuity constraints in the POT approach, but the overall fit of the constrained models is not as good as for the unconstrained models (see, e.g., Scarrott & MacDonald, 2012), which is why we do not show them here.

Table 3
Goodness-of-fit analysis – severity.

Model	Log-likelihood	AIC	Kolmogorov- Smirnov-test	Anderson- Darling-test
Panel A: Cyber risk (N = 1579)				
Exponential	−7535.78	15073.55	0.60***	79.94***
Gamma	−5368.23	10740.46	0.24***	18.79***
GPD	**−4553.42**	**9110.84**	0.07***	7.18***
Log-logistic	−4591.40	9186.80	1.00***	13.22***
Log-normal	−4588.09	9180.19	0.08***	16.99***
Weibull	−4886.78	9777.57	0.16***	60.20***
Skew-normal	−10718.32	21442.63	0.82***	166.03***
POT (threshold 56%)	**−4485.76**	**8979.50**	/	/
POT (threshold 80%)	−4510.16	9028.32	/	/
Transformation kernel	**−4402.59**	/	/	/
Panel B: Non-cyber risk (N = 24962)				
Exponential	−139542.80	279087.60	0.54***	58.75***
Gamma	−109184.80	218373.60	0.21***	13.10***
GPD	−99438.54	198881.10	0.03***	50.27***
Log-logistic	−99572.73	199149.50	1.00***	61.63***
Log-normal	**−99258.09**	**198520.20**	0.03***	54.86***
Weibull	−102587.30	205178.60	0.10***	6.30***
Skew-normal	−194267.50	388541.00	0.81***	230.30***
POT (threshold 56%)	**−98964.83**	**197937.70**	/	/
POT (threshold 80%)	−98998.47	198004.90	/	/
Transformation kernel	**−98120.48**	/	/	/

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). *, **, and *** indicate significance levels of 10%, 5%, and 1%.

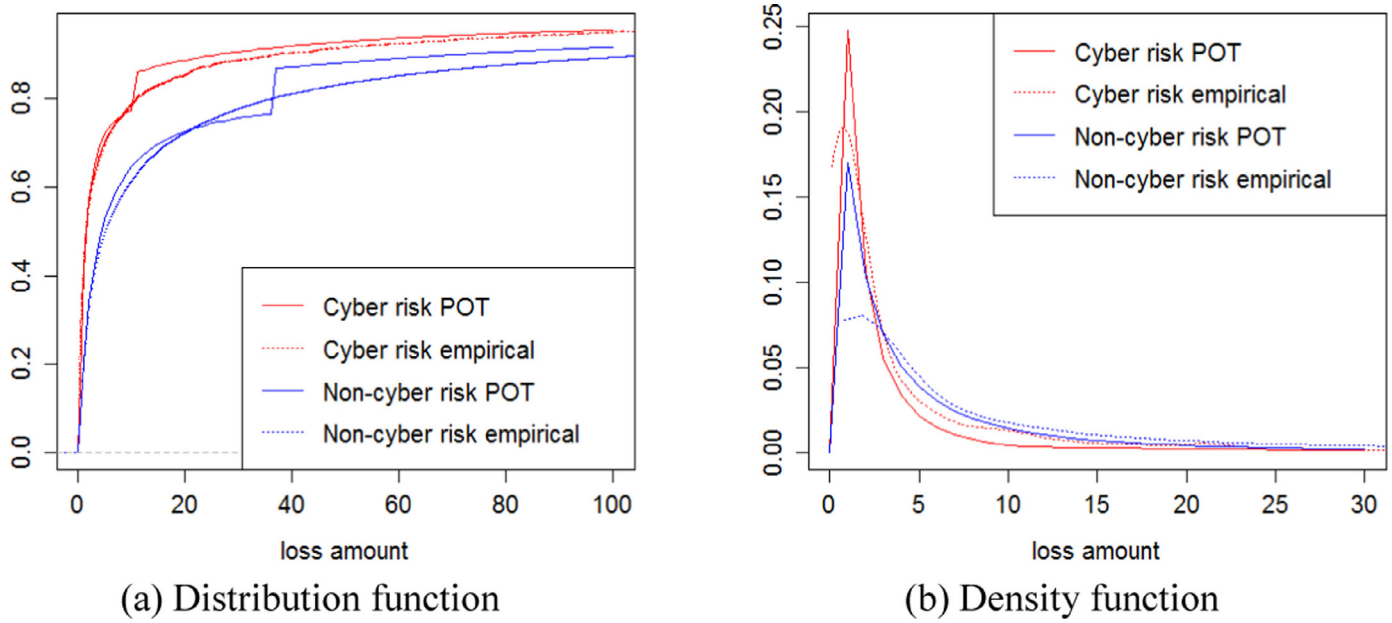

Fig. 2. Estimated distribution and density function.

regulation and practice; it thus seems that although cyber risks have less heavy tails than other operational risks, EVT is still the best approach to describe the data.

3.3. Extended models – The dynamic EVT approach

Column (1) in Table 4 presents the results for the loss frequency considering a Poisson Generalized Linear Model (GLM) with log-link function.[13] Columns (3) and (4) show the loss severity results considering the POT for two different threshold values (0% as a test without POT and 56% as optimal threshold with POT). Since

[13] We present the Poisson distribution, because its parameter is easier to interpret when depending on covariates. We also tested the negative binomial distribution and received similar results (see Online-Appendix DV). Note also that we use the log-link function since it is the only way to guarantee the convergence of our estimation approach. The convergence properties of EVT have been widely discussed in literature (see e.g. De Haan & Ferreira, 2007). In the case of Chavez-Demoulin, Embrechts, and Hofert (2016) in order to ensure convergence, a reparameterization of β in terms of the parameter ν is done via $\beta = \exp(\nu)/(1+\xi)$. The convergence is checked within the algorithm of Chavez-Demoulin et al. (2016) and is reached after 25 iterations. Based on the log-link in the Poisson GLM (see formula (3)) the interpretation for the actual loss frequency parameter is that ceteris paribus a one-unit increase in X_1, leads to a proportional change in λ of $e^{\beta 1}$. Furthermore, the variable "No. of Employees" exhibits few missing values, which is the reason why Table 4 is based on 1,515 data points. The code to obtain Table 4 runs through in less than two minutes and thus should not be consider as computational expensive. In the preceeding selection of the optimal threshold value, we use the bootstrap goodness-of-fit test by Villaseñor-Alva and González-Estrada (2009), but this is also not computational extensive and only one of several possible selection approaches.

Table 4
Results for the dynamic EVT approach.

Variable		(1) Frequency Poisson GLM	(2) Logistic regression (threshold 56%)	(3) Severity (threshold 0%)		(4) Severity (threshold 56%)	
		λ		Shape ξ	Transf. Scale	Shape ξ	Transf. Scale
Panel A: Cyber incident specific covariates							
Subcategories	Systems and technical failures	−1.736***	0.649***	0.963***	0.266	–	0.735
		(0.0761)	(0.1667)	(0.1155)	(0.7014)		(0.5770)
	Failed internal processes	−2.395***	0.205	0.153	0.049	–	0.009
		(0.1019)	(0.2183)	(0.1368)	(0.7687)		(0.8388)
	External events	−3.042***	0.517	0.556***	0.350	–	0.009
		(0.1380)	(0.3175)	(0.2014)	(1.2133)		(1.0350)
Contagion Type A	Multiple firms affected	−1.420***	0.191	−0.453***	–	–	–
		(0.0649)	(0.1405)	(0.0741)			
Contagion Type B	Multiple losses caused	−2.216***	0.812***	0.448***	1.237*	–	1.274**
		(0.0863)	(0.1932)	(0.1259)	(0.7516)		(0.6094)
Panel B: Company specific covariates							
Region of Domicile	Africa	−3.542***	0.763*	−0.113	–	–	–
		(0.2115)	(0.4473)	(0.2378)			
	Asia	−1.196***	−0.026	0.278***	–	–	–
		(0.0737)	(0.1611)	(0.0968)			
	Europe	−0.729***	0.347***	0.033	–	–	–
		(0.0622)	(0.1333)	(0.0780)			
	Other	−2.360***	0.365	0.099	–	–	–
		(0.1208)	(0.2552)	(0.1666)			
Industry	Financial	1.206***	−1.146***	−0.120	−1.247***	–	−0.515
		(0.0610)	(0.1378)	(0.0835)	(0.4868)		(0.4784)
Company size	No. of Employees	0.001***	0.000	−0.001*	–	0.001**	–
		(0.0001)	(0.0005)	(0.0003)		(0.0007)	
Panel C: Time + Intercept							
Time	Years (continuous)	0.111***	−0.109***	−0.011***	−0.110*	−0.060***	–
		(0.0050)	(0.0170)	(0.0092)	(0.0580)	(0.0184)	
Intercept		−6.685	1.699***	1.683***	3.242***	1.921***	2.732***
		(0.0903)	(0.2580)	(0.1451)	(0.8621)	(0.2498)	(0.4412)
Panel D: Model fit							
LLV		−10442.670	–	−4034.929		−2773.335	
LLV_null		−13376.640	–	−4151.807		−2809.587	
AIC		20915.340	–	8109.858		5564.670	
AIC_null		26755.290	–	8307.625		5623.174	

Note: Numbers in parentheses represent standard errors of the coefficients. LLV = log-likelihood value, AIC = Akaike Information Criterion. LLV_null and AIC_null show the results for the null model, i.e., the fit for the model without covariates. *, **, and *** indicate significance levels of 10%, 5%, and 1%. In column (4) the goodness-of-fit estimates in Panel D only compare the tail fit (as done in Chavez-Demoulin et al., 2016) and not the overall fit of the severity distribution as in Table 3.

the modeling algorithm of Chavez-Demoulin et al. (2016) requires a transformation of the original scale parameter, the interpretation of dependences is relatively complex; we thus focus on the shape parameter here. The results in column (1) show the loss frequency over the whole data sample; it is interesting to look at covariates that could determine tail events. For this purpose, the results of the logistic regression for the occurrence of tail events are given in column (2). The selection of variables in models (1), (3) and (4) is oriented to the approach used by Chavez-Demoulin et al. (2016) and provides the covariate combinations that best fit the data. The results presented for the logistic regression are based on the covariate combination that was optimal in the loss frequency model (1).

All covariates are important predictors for the loss frequency and confirm the direction seen in the descriptive analysis (see column (1)). The loss frequency shows a growth of about 11.7% $(e^{0.111\cdot(\text{year}+1)} - e^{0.111\cdot\text{year}})/e^{0.111\cdot\text{year}} = e^{0.111} - 1)$ per year over the whole sample period, although we observed a stagnation and decrease after May 2009 in Fig. 1. Frequency in general increased, but the probability of an extreme loss occurring decreased over time (Panel C, column (2)). This is another result, which is in contrast to existing literature; Wheatley et al. (2016) show that the rate of large data breach events has been stable in the US, but increased for non-US.

In addition, we observe that the loss severity's shape parameter decreased also for both threshold models (Panel C, columns (3) and (4)). Our results might be explained by the increase in the awareness of cyber risks, and thus companies learned how to protect themselves better (i.e., by reducing the size and probability of extreme losses). Similar observations can be made for different subcategories (i.e., incidents connected to human behavior are more frequent but less often lead to extreme events and exhibit less severe losses; Panel A) and the financial industry (i.e., members of the financial industry are more frequently victims of cyber risk, but are less vulnerable to extreme losses; Panel B). A significant effect on the severity for the financial industry, however, cannot be detected.

The frequency of losses increases with size, but size has no effect on the frequency of extreme losses (Panel B, columns (1) and (2)). We also note that bigger firms face less severe losses in general (column (3)), but measuring the tails only (column (4)), this relationship is reversed, meaning that bigger companies have heavier losses. However, this effect is not robust with respect to the model selection (e.g., if model (3) is computed with the variable combination of model (4) only, size is not significant).

In Panel D we see that the inclusion of the covariates significantly improves the model fit, confirming the results from other analyses (see, e.g., Ganegoda & Evans, 2013). We also note the size of the shape parameter for loss severity in columns (3) and (4). A shape parameter greater than 1 indicates infinite moments and thus an infinite expected value (see, e.g. McNeil et al., 2015). This is true for almost all covariate combinations and underlines the heavy-tailedness of cyber losses. In Online-Appendix DV we also consider the complete SAS OpRisk Global dataset with an indicator

Table 5
Risk measurement (in US$ million).

	No. of observations	VaR (90%)	VaR (95%)	VaR (99%)	VaR (99.5%)
Company 1					
Log-normal	390	290.232	400.130	804.449	1035.434
Transformation kernel	390	192.671	224.540	290.480	316.379
Dynamic POT 56%	1579	20.696	54.981	479.597	502.783
Empirical	390	19.513	43.969	373.899	479.691
Company 2					
Log-normal	106	784.284	1414.175	4617.081	7403.161
Transformation kernel	106	142.700	180.170	242.740	252.971
Dynamic POT 56%	1579	93.403	257.709	2572.252	2965.903
Empirical	106	73.300	264.790	2868.344	3359.552

Note: The log-normal and transformation kernel are simulated with one million random numbers. The number of observations in the Dynamic POT 56% is higher than in the other models, because this model can be fitted to all data simultaneously which is the particular advantage of our model. In contrast, the data for the other models in Table 5 cannot be pooled and thus must be fitted only to the subsample with the company characteristics.

variable for a cyber incident. The significance of this cyber variable in this case again illustrates that cyber and non-cyber losses exhibit different characteristics.

3.4. Application

In Table 5 we simulate loss frequency and loss severity for two sample companies and determine the risk measure value at risk (VaR) for three different modeling approaches: The log-normal widely used in actuarial practice, the transformation kernel which provides the second best fit in Section 3.2 and the dynamic POT, which provides the best fit in Section 3.2. Also, the empirical estimates based on the actual loss data are given. Company 1 is a small bank with 5000 employees. Company 2 is a retail company with 100,000 employees. Note that in Table 5 we reduced our analysis to the two covariates size and industry. This reduction is necessary since otherwise the empirical samples for the determination of the empirical VaR are too small. The model estimates are relatively far from the empirical estimates in the last row, but the dynamic EVT approach provides in most cases the most accurate results. The log-normal distribution leads to a substantial overestimation of the actual cyber losses. This is an important result since it is widely used and since the existing literature (Edwards et al., 2015; Wheatley et al., 2016) indicates that the log-normal is the most appropriate model to analyze the number of data breaches. This result, however, cannot be transferred to actual cost data.

Note that we do not present the tail value at risk (TVaR) since the models fitted here are almost all infinite mean models (i.e., the shape parameter is greater than 1), which leads to very heavy-tailed models with extreme uncertainties for high quantile estimates (see, e.g., Chavez-Demoulin et al., 2016). Approaches to correct these uncertainties in infinite mean models are tapering or truncation from above (see, e.g., Kagan & Schoenberg, 2001). However, these models face the disadvantage that the outcomes highly depend on the mechanism used to adjust for the infinite mean model (see, e.g., Chavez-Demoulin et al., 2016). Another application that cannot be presented without such adjustments is the pricing of cyber insurance policies.

4. Discussion and business implications

Two major limitations of existing papers in the cyber risk domain are their limited scope in terms of event types (only data breaches) and the fact that they do not consider cost information (only the number of data lost in a data breach). The results presented in this paper thus expand the existing knowledge on cyber risks in two important aspects. First, we analyze a broad set of cyber events and show that the data breaches account for only 25% of the actual loss events. Second, we analyze actual costs and show that models from extreme value theory are needed to evaluate the costs of cyber events. This is a very important finding, because the consideration of data breaches and non-monetary information (number of lost data) would result in a completely different evaluation and might thus lead to misinterpretations. For example, in our numerical illustration the cyber losses would be overestimated by factor 2 to 12 if the log-normal distribution widely used in actuarial practice is applied.

Cyber risk is thus different. Different not only compared to other risk categories, but also different from what we know on them (from data breaches) so far. The cost of cyber risk should not be measured based on standard actuarial tools, but the more complex extreme value theory, which is widely used in operational risk analysis, also needs to be applied to cyber losses. This insight also brings a number of new problems: Applying extreme value theory leads to infinite mean models where standard measures like TVaR or insurance prices cannot be estimated. We believe that the insurance industry is anticipating this problem and is thus reluctant to offer cyber insurance on a broader scale.[14] Our results also give some additional insights relevant to risk managers, insurers and policymakers: Although cyber risk are extreme, they are not as extreme as the other operational risk in our dataset. Moreover, we show that cyber risks are not correlated with other operational risk events in our dataset.

For the provision of insurance and the estimation of risk capitals, understanding the properties and behavior of cyber risk is vital. The results of this paper might thus offer insights for cyber risk management and the insurability of cyber risks. The findings 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 modeling approaches that go beyond the first modeling papers for this novel application area of risk management, but also note that today's limitations present an abundance of opportunities for future research.

Cyber risk is a dynamic risk category that has substantially evolved over time; also, the protective processes and systems are fundamentally evolving. Our dataset takes a historical perspective that might not necessarily reflect the future so we need to be clear about the limitation of our study. Still we believe that the analysis is useful, because it fills some important gaps in the literature and

[14] Although industry studies estimate the annual losses from cyber risks to be at least in a three-digit billion area, the annual premium volume is very low. For example, McAfee (2015) estimates the annual losses caused only by cybercrime at US$ 445 billion. The premium volume for the year 2015 is estimated at US$ 3.5 billion globally (see Marsh, 2016).

is among the first to thoroughly analyze the costs of cyber events. However, given the dynamic nature of cyber risk, the time dynamics of cyber losses need to be carefully studies in future research.

Moreover, the identification strategy based on OpRisk data should not be interpreted as more than a first step towards a more thorough analysis of cyber risk. International organizations and reinsurance firms are now starting to set up cyber loss databases and it will be interesting to apply our techniques when these data pools reach sufficient size. This seems especially relevant, because the results presented here indicate a certain lack of insurability, for example because insurance prices cannot be calculated. Furthermore, the inclusion of further covariates could fine-grain our analysis. For instance, the operational risk literature suggests further covariates (e.g., macro-environmental determinants, Cope, Piche, & Walter, 2012). Our risk estimates are also only a first indication of the true cyber risk, since reputational risks and thus loss of potential business are not incorporated. Other papers from the risk and insurance field could be used to estimate the reputational loss on top of the directly observable loss (see, e.g., Cannas, Masala, & Micocci, 2009, or Cummins, Lewis, & Wei, 2006) and linked with the results of this paper.[15] Finally, the fact that almost all estimated models in our analyses are infinite mean models limits the range of applications today (e.g., TVaR and insurance prices cannot be meaningfully calculated). But it also illustrates where we stand today in the application of sparse cyber risk data and as such the generation of more data and profound analyses of cyber risk constitutes an important area of future work.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ejor.2018.07.021.

References

Adcock, C. J., & Meade, N. (2017). 'Using parametric classification trees for model selection with applications to financial risk management'. *European Journal of Operational Research, 259*(2), 746–765.

Arnold, T. B., & Emerson, J. W. (2011). 'Nonparametric goodness-of-fit tests for discrete null distributions'. *The R Journal, 3*(2), 34–39.

Aven, T. (2016). 'Risk assessment and risk management: Review of recent advances on their foundation. *European Journal of Operational Research', 253*(1), 1–13.

Balkema, A. A., & 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', www.bis.org/publ/bcbs128.pdf, accessed 10 December 2013.

Biener, C., Eling, M., & Wirfs, J. H. (2015). 'Insurability of Cyber Risk – An Empirical Analysis'. *The Geneva Papers on Risk and Insurance – Issues and Practice, 40*(1), 131–158.

Bolancé, C., Guillen, M., & Nielsen, J. P. (2003). 'Kernel density estimation of actuarial loss functions'. *Insurance: Mathematics and Economics, 32*, 19–36.

Bolancé, C., Guillen, M., Pelican, E., & Vernic, R. (2008). 'Skewed bivariate models and nonparameteric estimation for the CTE risk measure'. *Insurance: Mathematics and Economics, 43*, 386–393.

Boonen, T. J. (2016). 'Nash equilibria of Over-The-Counter bargaining for insurance risk redistributions: The role of a regulator'. *European Journal of Operational Research, 250*(3), 955–965.

Boucher, J. P., Denuit, M., & Guillén, M. (2008). 'Models of insurance claim counts with time dependence based on generalization of Poisson and negative binomial distributions'. *Variance, 2*(1), 135–162.

Campbell, K., Gordon, L. A., Loeb, M. P., & Zhou, L. (2003). 'The economic cost of publicly announced information security breaches: Empirical evidence from the stock market'. *Journal of Computer Security, 11*(3), 431–448.

Cannas, G., Masala, G., & Micocci, M. (2009). 'Quantifying reputational effects for publicly traded financial institutions'. *Journal of Financial Transformation, 27*, 76–81.

Cavusoglu, H., Mishra, B., & Raghunathan, S. (2004). 'The effect of internet security breach announcements on market value: Capital market reactions for breached firms and internet security developers'. *International Journal of Electronic Commerce, 9*(1), 70–104.

Cebula, J. J., & 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., & Peters, J.-P. (2008). 'Practical methods for measuring and managing operational risk in the financial sector: A clinical study'. *Journal of Banking and Finance, 32*(6), 1049–1061.

Chavez-Demoulin, V., Embrechts, P., & Hofert, M. (2016). 'An extreme value approach for modeling operational risk losses depending on covariates'. *Journal of Risk and Insurance, 83*(3), 735–776.

Cope, E. W., Piche, M. T., & 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., & 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.

De Fontnouvelle, P., Dejesus-Rueff, V., Jordan, J. S., & 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.

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

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., & Tibiletti, L. (2010). 'Internal vs. external risk measures: How capital requirements differ in practice'. *Operations Research Letters, 38*(5), 482–488.

Eling, M., & Wirfs, J. H. (2016). Cyber Risk: Too Big to Insure? – Risk Transfer Options for a Mercurial Risk Class'. *I.VW Schriftenreihe, Band 59, St. Gallen*.

European Union (2016). 'Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)', http://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R0679&from=EN, accessed 24 June 2016.

Evans, M., Maglaras, L. A., He, Y., & Janicke, H. (2016). 'Human behaviour as an aspect of cybersecurity assurance'. *Security and Communication Networks, 9*(17), 4667–4679.

Ganegoda, A., & 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.

De Haan, L., & Ferreira, A. (2007). *'Extreme value theory: An introduction*'. Springer Science & Business Media.

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.

Hovav, A., & D'Arcy, J. (2003). 'The impact of denial-of-service attack announcements on the market value of firms'. *Risk Management and Insurance Review, 6*(2), 97–121.

Kagan, Y. Y., & Schoenberg, P. (2001). 'Estimation of the upper cutoff parameter for the tapered pareto distribution'. *Journal of Applied Probability, 38*, 158–175.

Laengle, S., Merigó, J. M., Miranda, J., Słowiński, R., Bomze, I., Borgonovo, E., Dyson, R. G., Oliveira, J. F., & Teunter, R. (2017). 'Forty years of the European Journal of Operational Research: A bibliometric overview'. *European Journal of Operational Research, 262*(3), 803–816.

Lehmann, E. L., & Romano, J. P. (2005). *Testing statistical hypotheses – Third Edition', Springer Texts in Statistics*. Springer.

Liu, F., & Pitt, D. (2017). 'Application of bivariate negative binomial regression model in analysing insurance count data'. *Annals of Actuarial Science, 11*(2), 390–411.

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

Marsh (2016). Continental European Cyber Risk Survey: 2016 Report. https://www.marsh.com/de/de/insights/research-briefings/continental-european-cyber-risk-survey-2016-report.html.

McNeil, A. J., Frey, R., & Embrechts, P. (2015). *'Quantitative risk management: Concepts, techniques, tools – Revised edition'*. Princeton University Press.

Nagurney, A., & Shukla, S. (2017). 'Multifirm models of cybersecurity investment competition vs. cooperation and network vulnerability'. *European Journal of Operational Research, 260*(2), 588–600.

National Conference of State Legislatures (NCSL) (2016). 'Security Breach Notifications Laws', http://www.ncsl.org/research/telecommunications-and-information-technology/security-breach-notification-laws.aspx, accessed 06 June, 2016.

Panjer, H. (2006). *Operational risk: Modeling analytics*. John Wiley & Sons.

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.

PwC (2015). 'Insurance 2020 & beyond – Repeating the dividends of cyber resilience', https://www.pwc.com/gx/en/insurance/publications/assets/reaping-dividends-cyber-resilience.pdf, accessed 02 March, 2016

[15] In fact it is not fully clear how substantial the reputational effects are, because the few existing event studies based on data breaches (e.g., Campbell, Gordon, Loeb, & Zhou, 2003; Cavusoglu, Mishra, & Raghunathan, 2004; Hovav and Arcy, 2003) only show relatively small effects.

Rakes, T. R., Deane, J. K., & Rees, L. P. (2012). 'IT security planning under uncertainty for high-impact events'. *Omega, 40*(1), 79–88.
Romanosky, S. (2016). 'Examining the costs and causes of cyber incidents'. *Journal of Cybersecurity, 2*(2), 121–135.
Scarrott, C., & MacDonald, A. (2012). 'A review of extreme value threshold estimation and uncertainty quantification'. *REVSTAT – Statistical Journal, 10*(1), 33–60.
Villaseñor-Alva, J. A., & González-Estrada, E. (2009). 'A bootstrap goodness of fit test for the generalized pareto distribution'. *Computational Statistics and Data Analysis, 53*(11), 3835–3841.
Wheatley, S., Maillart, T., & Sornette, D. (2016). 'The extreme risk of personal data breaches and the erosion of privacy'. *The European Physical Journal B, 89*(7), 1–12.