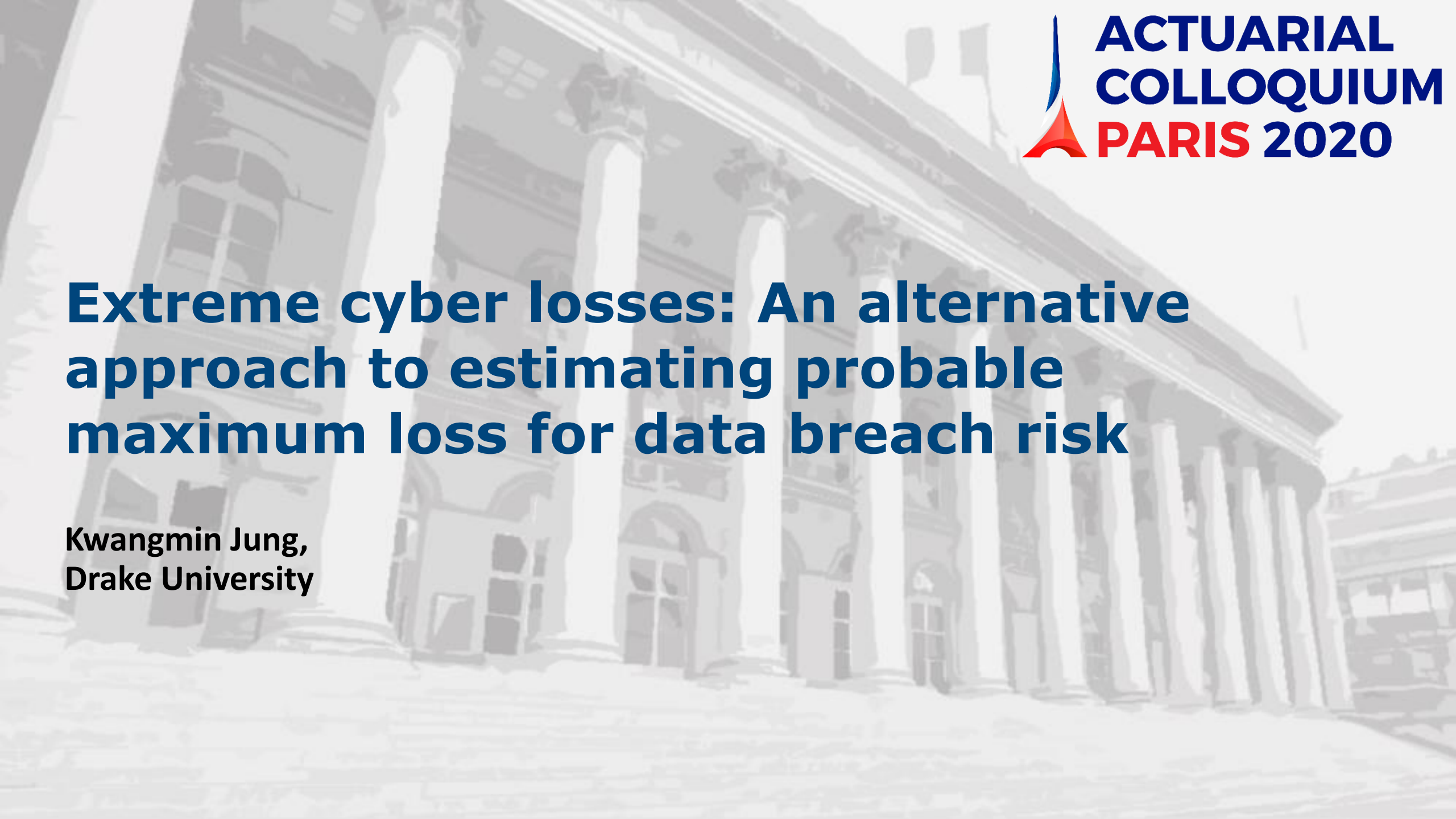


ACTUARIAL COLLOQUIUM PARIS 2020

Extreme cyber losses: An alternative approach to estimating probable maximum loss for data breach risk

Kwangmin Jung,
Drake University

About the speaker

ACTUARIAL COLLOQUIUM PARIS 2020

Kwangmin Jung

- Robb B. Kelley Visiting Distinguished Assistant Professor
- Joined Drake University in January, 2020 after Ph.D. in Finance at the University of St. Gallen, Switzerland
- Research interests: Cyber risk, InsurTech, Extreme risk modeling

---

Drake University

- Department of Actuarial Science and Risk Management
- One of the original 15 CAE Designated schools in the U.S.
- Located in Des Moines, Iowa, one of the country's leading center for insurance and financial services

Status-quo of the cyber-insurance market

Status-quo

- Market growth: **37%** per annum between 2016 and 2017
- Global premium volume (2017): $ 3.9bn ($ 2,234 bn of total non-life premium globally)
- 80% of the premium volume from the U.S. and the rest from Europe and Asia.
- 528 cyber-insurers in the U.S. in 2018 (6,000 insurers in total in the U.S.)

Challenges

Demand side

1. Lack of understanding of risk
2. Purchasing behavior relative to effect of risk control measures

Supply side

1. Lack of data
2. Challenge in modeling and pricing
3. Limited coverage (Cover limit)

Source: “Cyber Overview”, Munich Re
“Ten key questions on cyber risk and cyber risk insurance”, Eling and Schnell (2017) with Geneva Association
“Content analysis of cyber insurance policies: How do carriers write policies and price cyber risk?”, Romanosky et al. (2019)

Extreme cyber events

"Wannacry" ransomware 2017

Estimated economic loss = $4 billion

Source: ZDNet & Krebs on Security

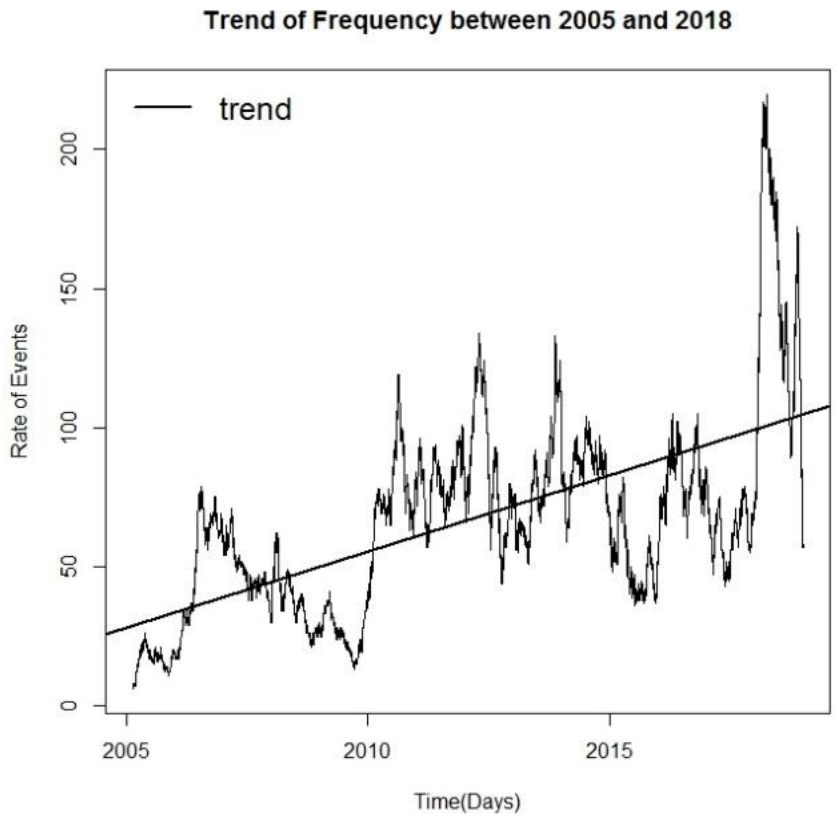
"NotPetya" virus 2017

Total economic damage = $10 billion

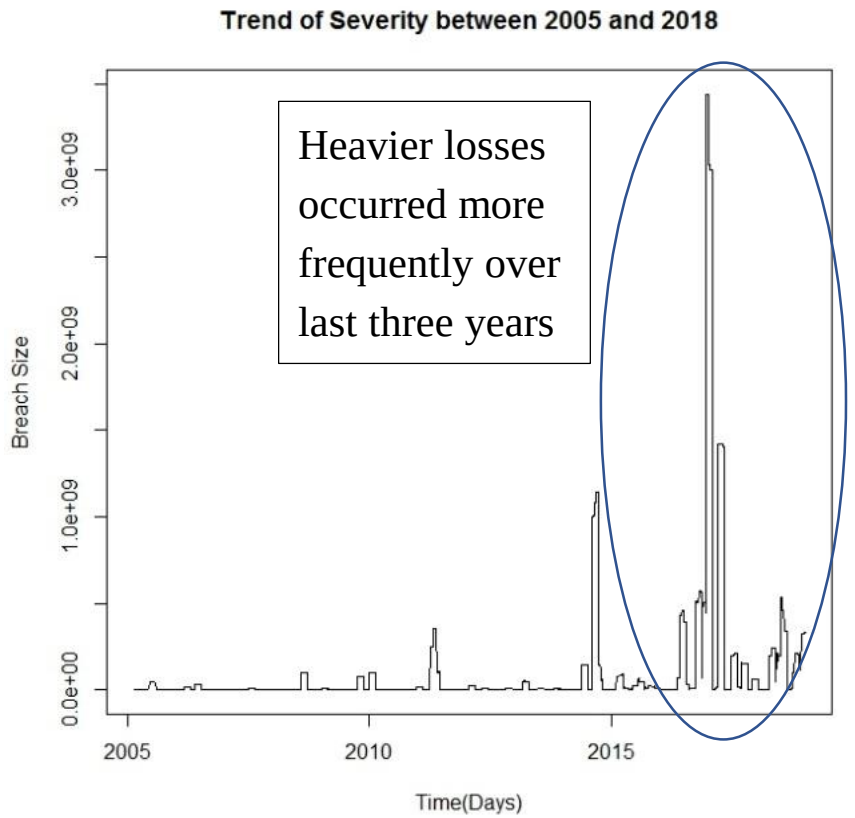
Source: Emsisoft Blog

Trends of loss frequency and severity

Frequency trend



Severity trend



Literature review on extreme cyber loss

	Maillart & Sornette (2010)	Edwards, Hofmeyr & Forrest (2016)	Wheatley, Maillart & Sornette (2016)	Eling & Jung (2018)	Eling & Wirfs (2019)	Hofmann, Wheatley & Sornette (2019)
Data period	2000-2008 (breach loss)	2005-2015 (breach loss)	2007-2015 (breach loss)	2005-2016 (breach loss)	1995-2014 (monetary loss)	2007-2017 (breach loss)
Methodology	Threshold-based	Lognormal	Threshold-based	Lognormal & threshold-based	Threshold-based	Threshold-based
Estimate of maximum loss	NA	130 million	300 million	1.1 billion (99.5%)	NA	NA

Dragon king beyond the estimation
(Sornette and Ouillon, 2012)

History of extreme loss events

Date	Breached entity	Risk type	Breached records
Dec 14, 2016	Yahoo	Hacking	3 billion
Mar 8, 2017	Multiple entities	Unintended disclosure	1.37 billion
Aug 5, 2014	Multiple entities	Hacking	1 billion
Sep 22, 2016	Yahoo	Hacking	0.50 billion
Nov 16, 2016	FriendFinder	Hacking	0.41 billion
May 31, 2016	MySpace	Hacking	0.36 billion
Jul 3, 2018	Exactis	Unintended disclosure	0.34 billion
Nov 30, 2018	Marriott International	Hacking	0.33 billion
Apr 2, 2011	Epsilon	Hacking	0.25 billion
Jun 19, 2017	DeepRootAnalytics	Unintended disclosure	0.20 billion
Dec 28, 2015	Multiple entities	Unintended disclosure	0.19 billion

Objectives of the study

Research questions

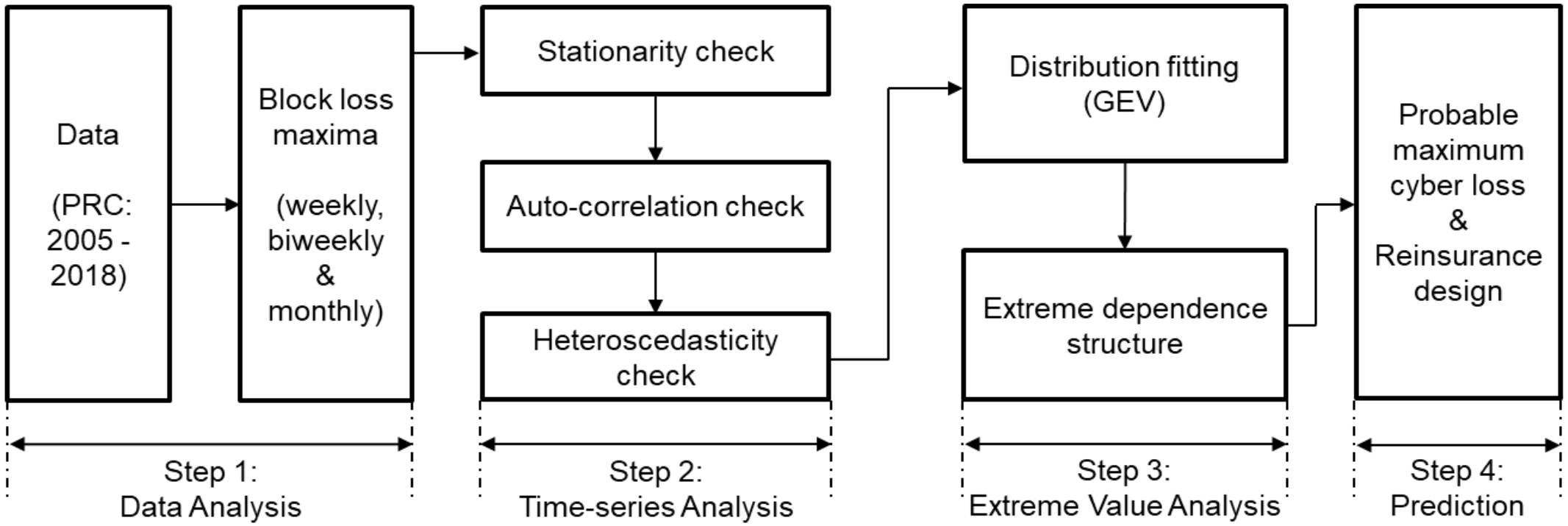
1) Can one statistically estimate **the size of cyber dragon king**?

2) If one can estimate the size of cyber dragon king, how can she apply this to the current insurance market and what could be a solution to manage a catastrophe cyber loss?

Objectives and Contributions

Aim 1:	Aim 2:	Aim 3:
The provision of an **alternative approach** to modeling extreme cyber loss	The provision of a **definition** on probable maximum loss for cyber risk	The provision of an **empirical benchmark** on reinsurance with **public-private partnership** (PPP)

Overview of modeling



Data

Data source: Privacy Rights Clearinghouse (PRC)

Period: Jan 1st, 2005 – Dec 31st, 2018

# of obs: 6,047 in total

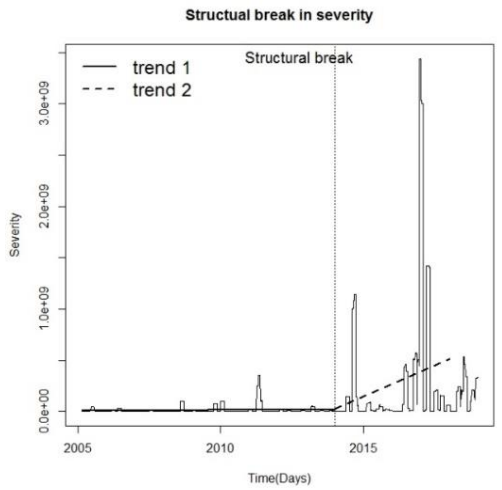
Risk classification (Edwards et al., 2016):

Risk type	Variable	Explanation
Malicious	Hacking (HACK)	Hacking attack by outsiders or infection by malware
	Insider (INSD)	Breached by an insider (e.g., employee or contractor)
	Payment card fraud (CARD)	Fraud involving debit and credit cards
Negligent	Portable device (PORT)	Lost, discarded or stolen portable devices
	Stationary device (STAT)	Lost stationary computers
	Unintended disclosure (DISC)	Privacy information disclosed unintentionally
	Physical loss (PHYS)	Lost, discarded or stolen non-electronic information

ACTUARIAL COLLOQUIUM PARIS 2020

Empirical estimation

Step 1: Find a break point



Break point	OLS-CUSUM	Rec-CUSUM	Chow
Jan, 2014	5.89***	3.85***	73.06***

Split the dataset into two periods: **pre-2014** and **post-2014**

Step 2: Time series analysis

Stationarity			
	Week	Biweek	Month
ADF	-8.76***	-4.86***	-4.89***
PP	-749.9***	-377.9***	-184.1***
KPSS	0.109	0.105	0.114
Temporal dependency			
Model	AR(12)	AR(6)	AR(3)
ARCH effect			
Lag=4	0.712	0.307	0.172
Lag=8	0.726	0.656	0.266
Lag=12	1.554	1.432	0.329
Lag=16	1.578	1.554	0.394
Lag=20	1.594	1.574	0.445
Lag=24	3.031	1.609	0.483

Stationary series, short-range temporal dependency and homoscedasticity

Step 3: GEV fitting and extreme dependency

Type I (Gumbel): $\gamma = 0$ (shape parameter)

Type II (Fréchet): $\gamma > 0$

Type III (Weibull): $\gamma < 0$

Fitting Generalized Extreme Value			
	Week	Biweek	Month
AIC	19,435.2	10,762.7	5,464.7
K-S GoF	0.030	0.035	0.058
A-D GoF	0.802	0.567	0.667
Shape parameter	2.272	2.115	1.661
Pickands test and extreme copula			
Test result	0.338***	0.027*	0.025
Copula	Clayton	Clayton	Tawn

Fréchet distribution and extreme dependency in monthly maxima series

Probable maximum loss and applications

Probable maximum cyber loss

$$\mathrm{P}\left[\widetilde{M}_n \le \xi_p\right] = 1 - p,$$

$$\xi_p = G_{\widetilde{M}_n}^{\theta\ -1}(1-p)$$

for some small $p \in [0,1]$, where
$\widetilde{M}_n$: a series of the cyber loss maxima
ξ_p: the probable maximum cyber loss
$G_{\widetilde{M}_n}^{\theta}$: the probability function of the maxima series with the parameter of θ.

Panel A: PMCL estimates					(million breach)
		Composite	Malicious	Negligent	Dependence
Next 3 yr	Entire	692.2	1,539.9	52.5	2,241.7
	Pre-2014	50.7	227.0	15.2	284.5
	Post-2014	62,693.3	20,533.2	313.1	33,004.6
Next 5 yr	Entire	2,053.2	5,987.1	140.8	8,723.7
	Pre-2014	117.6	784.8	32.6	876.3
	Post-2014	371,964.4	98,198.5	1,179.4	132,992.7

Panel B: Estimates of the recent literature			(million breach)
	Edwards et al. (2016) (Lognormal)	**Wheatley et al. (2016) (Pareto)**	**Eling and Jung (2018) (Correlated risk)**
Data source	PRC	Open Security Foundation & PRC	PRC
Maximum loss	130.00	300.00	1,053.11
Time prediction	Next 3 yr	Next 5 yr	1 out of 200 cases (99.5%)

Reinsurance with public intervention

Government (Cyber-deposit):

$$X_{Gov} = \begin{cases} 0, & L < U \\ L - U, & U < L \end{cases}$$

Cover limit (=U)

Reinsurer (Quota share):

$$X_{Re} = \begin{cases} L \times (1-q), & L < U \\ U \times (1-q), & U < L \end{cases}$$

Insurer:

$$X_{Ins} = \begin{cases} L \times q, & L < U \\ U \times q, & U < L \end{cases}$$

1 - q | q

Aggregate premium size (on a monthly basis) & The average size of loss per event

		Expectation principle		
($ million)		Comp	Mal	Neg
Reinsurer	Entire	**627.01**	768.20	112.73
Insurer	Entire	**209.00**	256.07	37.58
Government	Entire	**2,091.21**	1,959.31	1,285.63

- The estimated annual gross premium (=$10.03 bn) is almost **60% larger than** the predicted premium size of cyber insurance worldwide for 2019 (=$6.2 bn) by the industry (PwC, 2016).
- The government needs to take up on average **$ 2.1 billion loss per event** beyond our PMCL estimate.

Findings and implications

Research questions	Findings
	✓ **Stationary, but short-range temporal dependent maxima series** are identified (weekly, bi-weekly)
1) Estimation of the size of cyber dragon king?	✓ **Fréchet type** of GEV distribution is found to be optimal for cyber loss maxima series
	✓ **Seven times larger** than the one with a widely used Pareto-based model
2) How to apply the estimation to the insurance market and how to manage catastrophe cyber loss?	✓ Reinsurance design with the public intervention → **higher cover limit set-up**

Limitations of this study

Pont 1:
Lack of method to translate the breached records to the monetary cost.

Point 2:
Dataset covering mainly data breach risk, but not the entire set of cyber risk.

Thank you for your attention

ACTUARIAL COLLOQUIUM PARIS 2020

Contact details :

Kwangmin Jung

Drake University,
2507 University Ave.,
Des Moines, IA 50311,USA

kwangmin.jung@drake.edu (kwangmin.jung@unisg.ch)

https://www.actuarialcolloquium2020.com/

ACTUARIAL COLLOQUIUM PARIS 2020

Disclaimer:

The views or opinions expressed in this presentation are those of the authors and do not necessarily reflect official policies or positions of the Institut des Actuaires (IA), the International Actuarial Association (IAA) and its Sections.

While every effort has been made to ensure the accuracy and completeness of the material, the IA, IAA and authors give no warranty in that regard and reject any responsibility or liability for any loss or damage incurred through the use of, or reliance upon, the information contained therein. Reproduction and translations are permitted with mention of the source.

Permission is granted to make brief excerpts of the presentation for a published review. Permission is also granted to make limited numbers of copies of items in this presentation for personal, internal, classroom or other instructional use, on condition that the foregoing copyright notice is used so as to give reasonable notice of the author, the IA and the IAA's copyrights. This consent for free limited copying without prior consent of the author, IA or the IAA does not extend to making copies for general distribution, for advertising or promotional purposes, for inclusion in new collective works or for resale.