Estimation of the marginal expected shortfall in the context of an infinite mean model

Armelle Guillou

with Juan-Juan Cai & Valérie Chavez-Demoulin

$Z \in \mathcal{DA}$(Fréchet):

$$\lim_{t \to \infty} \frac{\mathbb{P}(Z > tx)}{\mathbb{P}(Z > t)} = x^{-1/\gamma}$$

$\widehat{\gamma} > 1$: dangerous zone of infinite mean models

$\longrightarrow$ these results are unrealistic and useless

Solution:

- $W = Z^a$ with $a\gamma < 1$: Cai *et al.* (2015):

$$\mathbb{E}\left[W \middle| Y > U_Y\left(\frac{1}{p}\right)\right]$$

$\longrightarrow$ power-transformation sometimes does not make sense or is difficult to interpret

Introduction

Solution:

- $X := \log(Z)$: $X \in \mathcal{DA}(\text{Gumbel})$

On the basis of a random sample $\{(Z_1, Y_1), \ldots, (Z_n, Y_n)\}$

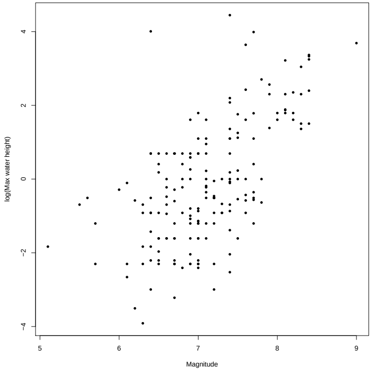
$$\theta_p = \mathbb{E}\left[X \middle| Y > U_Y\left(\frac{1}{p}\right)\right]$$

$\longrightarrow$ MES used to measure systemic risk of a financial institution

the logarithm is a natural transformation in many practical contexts

Motivation 1: Tsunami data

Two variables appear upper-tail dependent: the maximum water height and earthquake's magnitude



The magnitude of an earthquake is determined from the logarithm of the amplitude of waves recorded by seismographs

Motivation 2: Operational risk losses

Heavy-tailedness of the loss distribution and even infinite mean model

Consequences:

- Results are unthinkable (an infinite mean loss is not possible)
- Results are useless for practitioners in terms of regulatory capital calculation

This infinite mean finding typically appears when the data are contaminated by a few extremely high losses (such as Madoff's Ponzi scheme in 2008)

not outliers but real losses !

$\Rightarrow$ log-transformation of the data to avoid the zone of an infinite mean model

Cirillo and Taleb (2016); Chavez-Demoulin *et al.* (2015)

It is basically meant to thin the tail, allowing more reliable extreme value analytics

Estimator

(Z, Y) vector with continuous marginal df F_Z and F_Y

Interpretation of the MES: expected loss on the equity return of one financial institution given the occurrence of an extreme loss in the system
$\longrightarrow$ simultaneous high values of Z and Y
$\longrightarrow$ assumption imposed on the right upper-tail dependence of (Z, Y)

$$R_t(x,y) := t\,\mathbb{P}\left(1 - F_Z(Z) \leq \frac{x}{t}, 1 - F_Y(Y) \leq \frac{y}{t}\right) \xrightarrow{t\to\infty} R(x,y)$$

any continuous increasing transformation on Y will not change the MES

Tail dependence structure of (Z, Y) = Tail dependence structure of (X, Y)

Second order condition

Second Order Condition (SOC). *There exist* $\beta > 0$ *and* $\tau < 0$ *such that, as* $t \to \infty$

$$\sup_{0 < x \leq \infty, \frac{1}{2} \leq y \leq 2} \frac{|R_t(x,y) - R(x,y)|}{x^\beta \wedge 1} = O(t^\tau)$$

Theorem 1. *Assume* $Z \in \mathcal{DA}$*(Fréchet). Under the (SOC), we have*

$$\frac{\theta_p}{U_X(\frac{1}{p})} \xrightarrow{p \to 0} 1$$

First idea: wrong idea

$$\widehat{U}_X\left(\frac{1}{p}\right) = X_{n-k_1,n} + \widehat{\gamma}(k_1)\log\frac{k_1}{np}$$

$$\widehat{\gamma}(k_1) = \frac{1}{k_1}\sum_{i=1}^{k_1}(X_{n-i+1,n} - X_{n-k_1,n})$$

Rate of convergence

- $\widehat{U}_X\left(\frac{1}{p}\right)$ to $U_X\left(\frac{1}{p}\right)$: $\frac{\sqrt{k_1}}{\log\frac{k_1}{np}}$
- in Theorem 1: $\log\frac{1}{p}$

$$\frac{\sqrt{k_1}}{\log\frac{k_1}{np}}\left(\widehat{U}_X\left(\frac{1}{p}\right) - \theta_p\right)$$

does not converge to a Normal distribution

Decomposition of the MES

Let $s_{\frac{1}{p}}(x) := \frac{1}{p}\left(1 - F_X\left(xU_X\left(\frac{1}{p}\right)\right)\right)$

$$
\begin{aligned}
\theta_p &= \mathbb{E}\left[X \middle| Y > U_Y\left(\frac{1}{p}\right)\right] \\
&= U_X\left(\frac{1}{p}\right)\int_0^\infty R_{\frac{1}{p}}(s_{\frac{1}{p}}(x),1)dx \\
&= U_X\left(\frac{1}{p}\right) + U_X\left(\frac{1}{p}\right)\int_0^{a_n}\left[R_{\frac{1}{p}}(s_{\frac{1}{p}}(x),1) - 1\right]dx \\
&- U_X\left(\frac{1}{p}\right)\int_{a_n}^1\left[1 - R_{\frac{1}{p}}(s_{\frac{1}{p}}(x),1)\right]dx + U_X\left(\frac{1}{p}\right)\int_1^\infty R_{\frac{1}{p}}(s_{\frac{1}{p}}(x),1)dx
\end{aligned}
$$

where a_n is a sequence such that $a_n \in (0,1)$ and $a_n \to 1$ as $n \to \infty$

Let $s_{\frac{1}{p}}(x) := \frac{1}{p}\left(1 - F_X\left(xU_X\left(\frac{1}{p}\right)\right)\right)$

$$\begin{aligned}
\theta_p &= \mathbb{E}\left[X \middle| Y > U_Y\left(\frac{1}{p}\right)\right] \\
&= U_X\left(\frac{1}{p}\right)\int_0^\infty R_{\frac{1}{p}}(s_{\frac{1}{p}}(x),1)dx \\
&= U_X\left(\frac{1}{p}\right) + U_X\left(\frac{1}{p}\right)\int_0^{a_n}\left[R_{\frac{1}{p}}(s_{\frac{1}{p}}(x),1) - 1\right]dx \\
&\quad - U_X\left(\frac{1}{p}\right)\int_{a_n}^1\left[1 - R_{\frac{1}{p}}(s_{\frac{1}{p}}(x),1)\right]dx + U_X\left(\frac{1}{p}\right)\int_1^\infty R_{\frac{1}{p}}(s_{\frac{1}{p}}(x),1)dx
\end{aligned}$$

where a_n is a sequence such that $a_n \in (0,1)$ and $a_n \to 1$ as $n \to \infty$

Decomposition of the MES

Let $s_{\frac{1}{p}}(x) := \frac{1}{p}\left(1 - F_X\left(xU_X\left(\frac{1}{p}\right)\right)\right)$

$$
\begin{aligned}
\theta_p &= \mathbb{E}\left[X \middle| Y > U_Y\left(\frac{1}{p}\right)\right] \\
&= U_X\left(\frac{1}{p}\right)\int_0^\infty R_{\frac{1}{p}}(s_{\frac{1}{p}}(x),1)dx \\
&= U_X\left(\frac{1}{p}\right) + U_X\left(\frac{1}{p}\right)\int_0^{a_n}\left[R_{\frac{1}{p}}(s_{\frac{1}{p}}(x),1) - 1\right]dx \\
&\quad - U_X\left(\frac{1}{p}\right)\int_{a_n}^1\left[1 - R_{\frac{1}{p}}(s_{\frac{1}{p}}(x),1)\right]dx + U_X\left(\frac{1}{p}\right)\int_1^\infty R_{\frac{1}{p}}(s_{\frac{1}{p}}(x),1)dx
\end{aligned}
$$

where a_n is a sequence such that $a_n \in (0,1)$ and $a_n \to 1$ as $n \to \infty$

Decomposition of the MES

Let $s_{\frac{1}{p}}(x) := \frac{1}{p}\left(1 - F_X\left(xU_X\left(\frac{1}{p}\right)\right)\right)$

$$
\begin{aligned}
\theta_p &= \mathbb{E}\left[X \middle| Y > U_Y\left(\frac{1}{p}\right)\right] \\
&= U_X\left(\frac{1}{p}\right)\int_0^\infty R_{\frac{1}{p}}(s_{\frac{1}{p}}(x),1)dx \\
&= U_X\left(\frac{1}{p}\right) + U_X\left(\frac{1}{p}\right)\int_0^{a_n}\left[R_{\frac{1}{p}}(s_{\frac{1}{p}}(x),1) - 1\right]dx \\
&\quad \underbrace{-U_X\left(\frac{1}{p}\right)\int_{a_n}^1\left[1 - R_{\frac{1}{p}}(s_{\frac{1}{p}}(x),1)\right]dx}_{\Theta_1} \underbrace{+U_X\left(\frac{1}{p}\right)\int_1^\infty R_{\frac{1}{p}}(s_{\frac{1}{p}}(x),1)dx}_{\Theta_2}
\end{aligned}
$$

where a_n is a sequence such that $a_n \in (0,1)$ and $a_n \to 1$ as $n \to \infty$

Decomposition of the MES

Using a non-parametric estimator of $R(x,1)$

$$\Theta_1 \approx -\frac{1}{k_2}\sum_{i=1}^{n} \mathbb{1}_{\{R_i^X<n-k_2+1, R_i^Y\geq n-k_2+1\}}\left\{\left[(1-a_n)U_X\left(\frac{1}{p}\right)\right]\wedge\left[\gamma\log\frac{n-R_i^X}{k_2}\right]\right\}$$

$$\Theta_2 \approx \frac{\gamma}{k_1}\sum_{i=1}^{n} \mathbb{1}_{\{R_i^X>n-k_1+1, R_i^Y>n-k_1+1\}}\log\frac{k_1}{n-R_i^X+1}$$

$$U_X\left(\frac{1}{p}\right) \longrightarrow \widehat{U}_X\left(\frac{1}{p}\right)$$

$$\gamma \longrightarrow \widehat{\gamma}(k_1)$$

Final estimator

$$\widehat{\theta}_p = \widehat{U}_X\left(\frac{1}{p}\right)$$

$$-\frac{1}{k_2}\sum_{i=1}^{n}\mathbb{1}_{\{R_i^X<n-k_2+1,R_i^Y\geq n-k_2+1\}}\left\{\left[(1-a_n)\widehat{U}_X\left(\frac{1}{p}\right)\right]\wedge\left[\widehat{\gamma}(k_1)\log\frac{n-R_i^X}{k_2}\right]\right\}$$

$$+\frac{\widehat{\gamma}(k_1)}{k_1}\sum_{i=1}^{n}\mathbb{1}_{\{R_i^X>n-k_1+1,R_i^Y>n-k_1+1\}}\log\frac{k_1}{n-R_i^X+1}$$

Assumption: Hall-type model

$$1-F_Z(t)=Ct^{-\frac{1}{\gamma}}\left(1+D\,t^{\frac{\rho}{\gamma}}(1+o(1))\right)$$

Asymptotic result

- (SOC) + Hall model
- $R_t(x,y) \xrightarrow{t\to\infty} R(x,y)$
- $r_1(x,y) = \frac{\partial}{\partial x} R(x,y)$ and $r_2(x,y) = \frac{\partial}{\partial y} R(x,y)$ continuous
- $k_1 = n^a$ and $k_2 = n^b$ for positive a and b:

$$\begin{cases} a \in \left(0, \min\left(\frac{-2\tau}{1-2\tau}, \frac{-2\rho_1}{1-2\rho_1} b, -2\tau(1-b), 2b-1\right)\right) \\ np = o(k_1) \quad \text{and} \quad \log(np) = o(\sqrt{k_1}) \\ a_n \to 1 \quad \text{and} \quad \sqrt{k_1}(1-a_n) \to \infty \\ p^{a_n-1} = O\left(n^{1-b}\right) \quad \text{and} \quad n^{\frac{a}{2}} p^{\tau(a_n-1)} = o(1) \end{cases}$$

Then

$$\frac{\sqrt{k_1}}{\log \frac{k_1}{np}} \left(\widehat{\theta}_p - \theta_p\right) \xrightarrow{d} \mathcal{N}(0, \gamma^2)$$

Simulation study

• $B = 1000$ samples from a bivariate standard Cauchy distribution on $\mathbb{R}^2$ with density $(1/2\pi)(1+x^2+y^2)^{-3/2}$

• first component transformed with the logarithm

• k_2 has less impact than $k_1 \longrightarrow b = 0.75$, $a = 0.2$ or 0.3

$$\begin{aligned}(n,p) \quad = \quad & (50, 0.05); (50, 0.01); (50, 0.001) \\ & (200, 0.01); (200, 0.001) \\ & (1000, 0.001); (1000, 0.0005)\end{aligned}$$

Simulation study: $n = 50, p = 0.05$

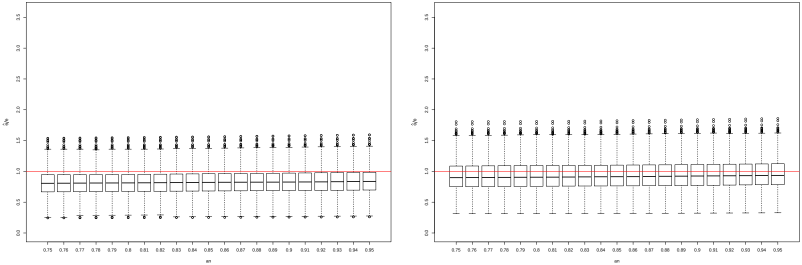


Figure 1: Boxplots of $\frac{\widehat{\theta}_p}{\theta_p}$. Left: $a = 0.2$; Right: $a = 0.3$

Simulation study: $n = 50, p = 0.01$

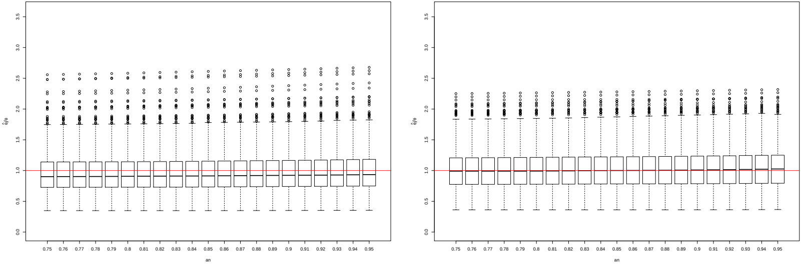


Figure 2: Boxplots of $\frac{\widehat{\theta}_p}{\theta_p}$. Left: $a = 0.2$; Right: $a = 0.3$

Simulation study: $n = 50, p = 0.001$

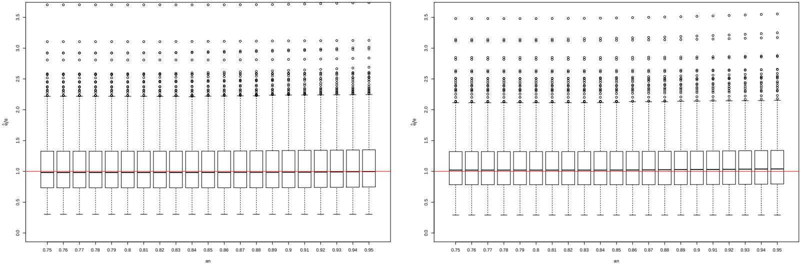


Figure 3: Boxplots of $\frac{\widehat{\theta}_p}{\theta_p}$. Left: $a = 0.2$; Right: $a = 0.3$

$np = 2.5 \longrightarrow$ Empirical estimator $\widehat{\theta}_{emp} = \frac{1}{\lfloor np \rfloor} \sum_1^n \log(Z_i) \mathbb{1}_{\{Y_i > Y_{n-\lfloor np \rfloor, n}\}}$

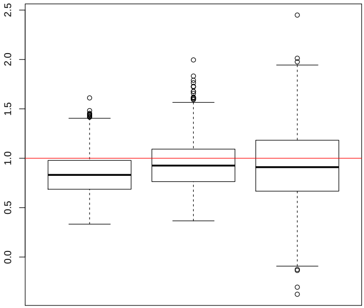


Figure 4: Boxplots of $\frac{\widehat{\theta}_p}{\theta_p}$: Left: $a = 0.2$; Middle: $a = 0.3$; Boxplots of $\frac{\widehat{\theta}_{emp}}{\theta_p}$: Right: empirical

Simulation study: $n = 200, a = 0.2$

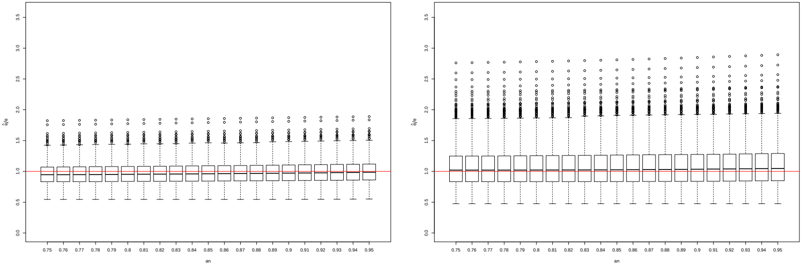


Figure 5: Boxplots of $\frac{\widehat{\theta}_p}{\theta_p}$. Left: $p = 0.01$; Right: $p = 0.001$

Simulation study: $n = 1000, a = 0.2$

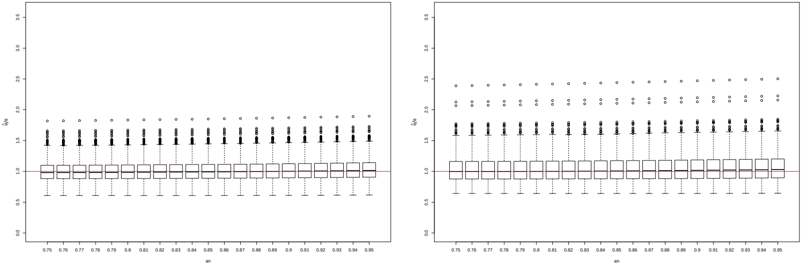


Figure 6: Boxplots of $\frac{\widehat{\theta}_p}{\theta_p}$. Left: $p = 0.001$; Right: $p = 0.0005$

Applications

Application to two different contexts, the features of which are infinite mean models

The utility of the methodology:

- comparing estimates $\widehat{\theta}_p$ in time (environmental ex.)
- comparing estimates $\widehat{\theta}_p$ between components (financial ex.)

Application: Tsunami data

US National Oceanic and Atmospheric Administration
on Japanese tsunamis from 1400 to 2011: 188 time points values

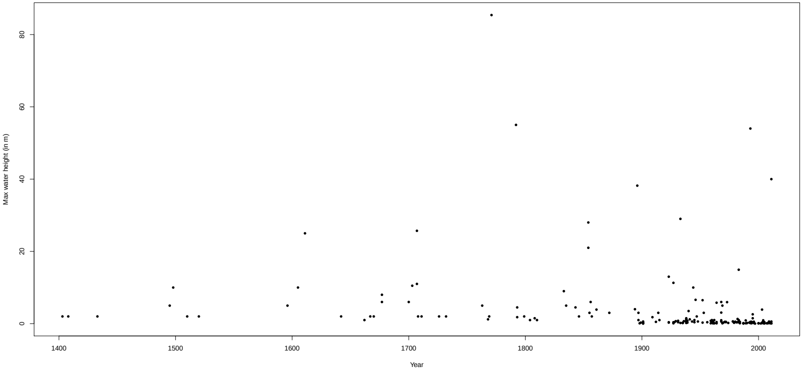


Figure 7: Maximum water height in meters over year (1400 to 2011)

Application: Tsunami data

• Maximum water height of 85.4m appeared in 1771, due to a earthquake of magnitude 7.4 in Ryukyu Islands $\longrightarrow$ 13,500 deaths

• In 1993 the 54m water height in Sea of Japan which succeeded an earthquake of magnitude 7.7 $\longrightarrow$ 208 deaths

• The 2011 event in Honshu preceding earthquake of magnitude 9 $\longrightarrow$ 15,550 deaths

the two variables show an upper-tail dependence structure

Application: Tsunami data

Two sub-datasets each of size 50:

• the most recent 50 observations: 1982 to 2011 $\longrightarrow$ data2011

• 50 observations from 1769 to 1936 $\longrightarrow$ data1936

$\longrightarrow$ the two datasets have no common observations

$\longrightarrow$ the number of years included in each dataset is not the same, but we use the same size $n = 50$ to obtain for each estimation the same amount of information

Application: Tsunami data

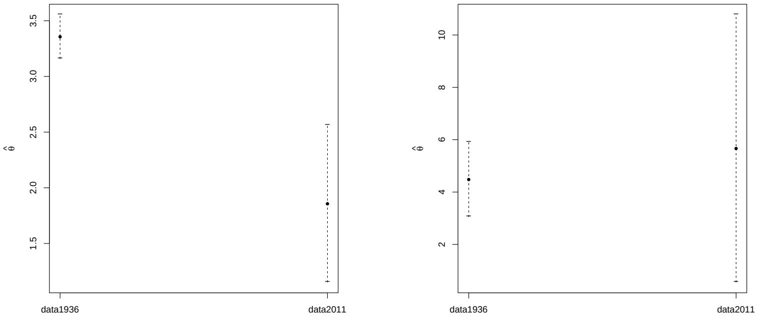
Considering a threshold of 0.02 which is the minimum observed values of water height and justified by the fact that any earthquake is an extreme event, we fit a generalized Pareto distribution (GPD) model on the maximum water height on the two datasets

Time	data1936	data2011
$\widehat{\gamma}(se)$	1.12(0.30)	1.08(0.26)

Entire dataset (with 188 values): 1.06 (0.15)

These unrealistic situations of infinite mean models (although with confidence intervals containing values below 1) typically arise when some values of the dataset are extremely large, contaminating the distribution

Application: Tsunami data, left $p = 0.05$, right $p = 0.01$



Earthquakes of very large magnitude (above the 95%-quantile) may have had a less important impact on maximum wave height during more recent earthquakes compared to their impact during the 1930 whereas extremely large magnitudes (above the 99%-quantile) have slightly larger impact than observed previously on the wave heights

This is possibly due to the fact that globally the wave height has a decreasing trend whereas the level of earthquake magnitude changes less

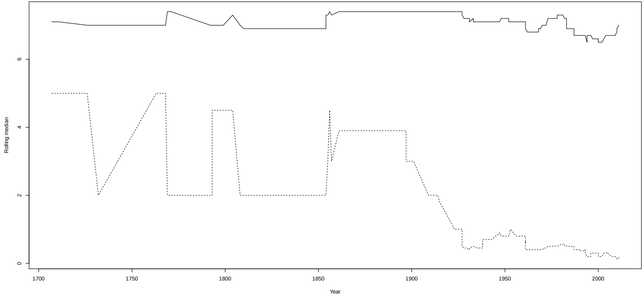


Figure 8: Rolling median over the wave height (dashed line) and over the earthquake magnitude (straight line) using a window size of 19 time points

Application: Operational risk data

OpRisk data collected from public media by Willis Professional Risks

The data correspond to the net loss amount in GBP classified in 10 business lines dictated by the Basel Committee:

- Agency Services (AS)
- Asset Management (AM)
- Commercial Banking (CB)
- Corporate Finance (CF)
- Insurance (I)
- Payment and Settlement (PS)
- Retail Banking (RBa)
- Retail Brokerage (RBr)
- Trading and Sales (TS)
- unallocated business line (UBL)

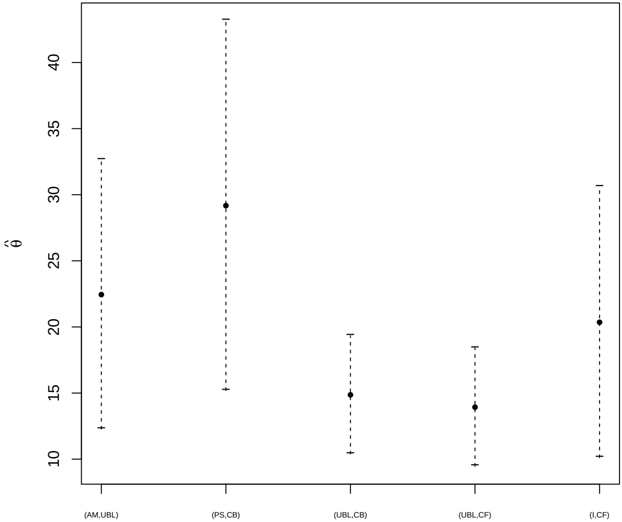
Application: Operational risk data

• We aggregate yearly losses for each business line so that we get pairs of losses which arrival times coincide

Based on the $n = 34$ yearly losses, we fit a GPD model for each business line and get estimate $\widehat{\gamma}$ above 1 apart for the business lines AS, CF and TS, but the high uncertainty due to small data size leads to high standard errors

Application: Operational risk data, $p = 0.001$

Z corresponds to the first business line of the pair and Y to the second



Application: Operational risk data, $p = 0.001$

Interpretation:

• the risk measure provide the information that the business line PS would be more greatly affected than the business line UBL if the business line CB experienced severe losses

• the pairs (AM, UBL) and (PS, CB) not surprinsingly get a high value of $\widehat{\theta}_{0.001}$ with large confidence interval

- AM contains the largest loss: 40 819 M GBP due to Madoff's Ponsi scheme in 2008

- PS contains the second largest loss: corresponding to Parmalat (related to dubious transactions with funds on Cayman Islands)