Estimation of the marginal expected shortfall: the mean when a related variable is extreme

Introduction

- Denote the loss of the equity return of a financial institution and that of the entire market as X and Y respectively.
- The MES is equal to $E(X|Y > t)$, where t is a high threshold such that $p = P(Y > t)$ is extremely small.
- The CTE in a univariate context is the same as that of the tail value at risk.
- In other words, the MES at probability level p is defined as

$$MES(p) = E(X|Y > Q_Y(1 - p)),$$

where Q_Y is the quantile function of Y.

Notation

- Let (X, Y) be a random vector with a continuous distribution function F.
- Define
$$U_j = \left(\frac{1}{1-F_j}\right)^{\leftarrow}, j = 1, 2.$$
- Then the MES at a probability level p can be written as
$$\theta_p := E\{X|Y > U_2(1/p)\}.$$
- $p = p(n) \to 0.$

Tail Dependence

Suppose that, for all $(x, y) \in [0, \infty]^2 \backslash \{(\infty, \infty)\}$, the following limit exists:

$$\lim_{t \to \infty} tP\{1 - F_1(X) \leqslant x/t, 1 - F_2(Y) \leqslant y/t\} =: R(x, y).$$

The function R completely determines the so-called stable tail dependence function l as, for all $x, y \geq 0$,

$$l(x, y) = x + y - R(x, y).$$

We assume X follows a distribution with a heavy right-hand tail: there exists $\gamma_1 > 0$ such that, for $x > 0$,

$$\lim_{t \to \infty} U_1(tx)/U_1(t) = x^{\gamma_1}.$$

Brief Review of multivariate extremes

Suppose $(X_1, Y_1), (X_2, Y_2), \ldots$ are independent and identically distributed random variables with distribution function F. Suppose that there exist sequences of constants $a_n, c_n > 0$, b_n and d_n real and a distribution function G with nondegenerate marginals such that for all continuity points (x, y) of G,

$$\lim_{n\to\infty} P\left(\frac{\max(X_1, \ldots, X_n) - b_n}{a_n} \le x, \frac{\max(Y_1, \ldots, Y_n) - d_n}{c_n} \le y\right) = G(x, y)$$

Any limit distribution function G with nondegenerate marginals is called a multivariate extreme value function.

Brief Review of multivariate Extremes

This implies convergence of the one-dimensional two marginal distributions, we have

$$\lim_{n\to\infty} P\left(\frac{\max\left(X_1, X_2, \dots, X_n\right) - b_n}{a_n} \leq x\right) = G(x, \infty),$$

and

$$\lim_{n\to\infty} P\left(\frac{\max\left(Y_1, Y_2, \dots, Y_n\right) - d_n}{c_n} \leq y\right) = G(\infty, y).$$

Now we choose the constants a_n, c_n, b_n and d_n such that for $\gamma_1, \gamma_2 \in \mathbb{R}$,

$$G(x, \infty) = \exp\left(-\left(1 + \gamma_1 x\right)^{-1/\gamma_1}\right),$$

and

$$G(\infty, y) = \exp\left(-\left(1 + \gamma_2 y\right)^{-1/\gamma_2}\right).$$

Brief Review of multivariate Extremes

$$\lim_{n\to\infty} F^n\left(a_n x + b_n, c_n y + d_n\right) = G(x, y),$$

Note that if $x_n \to u, y_n \to v$, then by the continuity of G and the monotonicity of F

$$\lim_{n\to\infty} F^n\left(a_n x_n + b_n, c_n y_n + d_n\right) = G(u, v).$$

We apply this with

$$x_n := \frac{U_1(nx) - b_n}{a}, y_n := \frac{U_2(ny) - d_n}{c_n}.$$

And we get

$$\lim_{n\to\infty} F^n\left(U_1(nx), U_2(ny)\right) = G\left(\frac{x^{\gamma_1} - 1}{\gamma_1}, \frac{y^{\gamma_2} - 1}{\gamma_2}\right) := G_0(x, y).$$

Brief Review of multivariate Extremes

For any (x, y) for which $0 < G_0(x, y) < 1$,

$$\lim_{l \to \infty} n \left\{1 - F\left(U_1(nx), U_2(ny)\right)\right\} = -\log G_0(x, y) := L(x, y).$$

$$\begin{aligned}
R(x, y) &= \lim_{t \to \infty} tP\left\{1 - F_1(X) \leqslant x/t, 1 - F_2(Y) \leqslant y/t\right\} \\
&= \lim_{t \to \infty} tP\left\{X \geq U_1(\frac{t}{x}), Y \geq U_2(\frac{t}{y})\right\} \\
&= \lim_{t \to \infty} t\left(1 - P\{X \leq U_1(t/x)\} - P\{Y \leq U_2(t/y)\} + P\{X \leq U_1(t/x\right. \\
&= \lim_{t \to \infty} t\left(1 - P\{X \leq U_1(t/x)\}\right) + \lim_{t \to \infty} t\left(1 - P\{X \leq U_2(t/y)\}\right) \\
&\quad - \lim_{t \to \infty} t\left(1 - P\{X \leq U_1(t/x), Y \leq U_2(t/y)\}\right) \\
&= x + y - L(x, y).
\end{aligned}$$

MES

$$\theta_p = E(X|Y > U_2(1/p))$$

- X Positive
- X real

X Positive

Assume that X takes values in $(0, \infty)$.

Proposition 1. Suppose $\gamma_1 \in (0, 1)$. Then,

$$\lim_{p \downarrow 0} \frac{\theta_p}{U_1(1/p)} = \int_0^\infty R\left(x^{-1/\gamma_1}, 1\right) \mathrm{d}x.$$

we construct an estimator of θ_p by two-stage approach.

- Firstly, we consider the estimation $\theta_{k/n}, k/n \to 0$.
- Secondly, we use an extrapolation method

$$\theta_p \sim \frac{U_1(1/p)}{U_1(n/k)} \theta_{k/n} \sim \left(\frac{k}{np}\right)^{\gamma_1} \theta_{k/n}.$$

A Sketch of Proof of Proposition 1

Recall that, for a non-negative random variable Z

$$E(Z) = \int_0^\infty P(Z > x)\mathrm{d}x$$

Hence,

$$\begin{aligned}\frac{\theta_p}{U_1(1/p)} &= \int_0^\infty \frac{1}{p} P\left\{X > x, Y > U_2(1/p)\right\} \frac{\mathrm{d}x}{U_1(1/p)} \\ &= \int_0^\infty \frac{1}{p} P\left\{X > U_1(1/p)x, Y > U_2(1/p)\right\} \mathrm{d}x\end{aligned}$$

The limit relationships (1) and (2) imply that

$$\lim_{p \downarrow 0} \frac{1}{p} P\left\{X > U_1\left(\frac{1}{p}\right) x, Y > U_2\left(\frac{1}{p}\right)\right\} = R\left(x^{-1/\gamma_1}, 1\right)$$

Use DCT to prove the integral and limit can be interchanged.

Non-parametric Estimator

- $$\hat{\theta}_{k/n} = \frac{1}{k} \sum_{i=1}^{n} X_i I\left(Y_i > Y_{n-k,n}\right)$$

- We estimate γ_1 with the Hill (1975) estimator

$$\hat{\gamma}_1 = \frac{1}{k_1} \sum_{i=1}^{k_1} \log\left(X_{n-i+1,n}\right) - \log\left(X_{n-k_1,n}\right).$$

- we estimate θ_p by

$$\hat{\theta}_p = \left(\frac{k}{np}\right)^{\hat{\gamma}_1} \hat{\theta}_{k/n}.$$

Some Conditions

(a) There exist $\beta > \gamma_1$ and $\tau < 0$ such that, as $t \to \infty$,

$$\sup_{\substack{0<x<\infty \\ 1/2 \leqslant y \leqslant 2}} \frac{\left| tP\left\{1-F_1(X) < x/t, 1-F_2(Y) < y/t\right\} - R(x,y)\right|}{x^\beta \wedge 1} = O\left(t^\tau\right).$$

(b) There exist $\rho_1 < 0$ and an eventually positive or negative function A_1 such that, as $t \to \infty$, $A_1(tx)/A_1(t) \to x^{\rho_1}$ for all $x > 0$ and

$$\sup_{x>1} \left| x^{-\gamma_1} \frac{U_1(tx)}{U_1(t)} - 1 \right| = O\left\{A_1(t)\right\}.$$

(c) As $n \to \infty$, $\sqrt{k_1} A_1(n/k_1) \to 0$.

(d) As $n \to \infty, k = O(n^\alpha)$, for some $\alpha < \min\left\{-2\tau/(-2\tau+1), 2\gamma_1\rho_1/\left(2\gamma_1\rho_1 + \rho_1 - 1\right)\right\}$.

Limit Distribution

We define a mean 0 Gaussian process W_R on $(x, y) \in [0, \infty]^2 \backslash \{(\infty, \infty)\}$ with covariance structure

$$E\left\{W_R\left(x_1, y_1\right) W_R\left(x_2, y_2\right)\right\} = R\left(x_1 \wedge x_2, y_1 \wedge y_2\right).$$

Set

$$\Theta = \left(\gamma_1 - 1\right) W_R(\infty, 1) + \left\{\int_0^\infty R(s, 1) \mathrm{d}s^{-\gamma_1}\right\}^{-1} \int_0^\infty W_R(s, 1) \mathrm{d}s^{-\gamma_1},$$

and

$$\Gamma = \gamma_1 \left\{-W_R(1, \infty) + \int_0^1 s^{-1} W_R(s, \infty) \mathrm{d}s\right\}.$$

Theorem 1

Suppose that conditions (a)–(d) hold and $\gamma \in (0, 1/2)$. Assume that $d_n = k/(np) \geq 1$ and

$r := \lim_{n\to\infty} \sqrt{k} \log(d_n) / \sqrt{k_1} \in [0, \infty]$. If $\lim_{n\to\infty} \log(d_n) / \sqrt{k}_1 = 0$,then, as $n \to \infty$,

$$\min\left\{\sqrt{k}, \frac{\sqrt{k_1}}{\log(d_n)}\right\}\left(\frac{\hat{\theta}_p}{\theta_p} - 1\right) \xrightarrow{\mathrm{d}} \begin{cases} \Theta + r\Gamma, & \text{if } r \leqslant 1 \\ (1/r)\Theta + \Gamma, & \text{if } r > 1 \end{cases}$$

$$\operatorname{var}(\Theta) = \gamma_1^2 - 1 - b^2 \int_0^\infty R(s,1)\mathrm{d}s^{-2\gamma_1}, \quad \operatorname{var}(\Gamma) = \gamma_1^2$$

$$\begin{aligned}\operatorname{cov}(\Gamma, \Theta) &= \gamma_1 (1 - \gamma_1 + b) R(1,1) \\ &\quad - \gamma_1 \int_0^1 \left[(1-\gamma_1) + bs^{-\gamma_1}\{1 - \gamma_1 - \gamma_1 \ln(s)\}\right] R(s,1) s^{-1} \mathrm{d}s\end{aligned}$$

A sketch of Proof of Theorem 1

$$\frac{\widehat{\theta}_p}{\theta_p} = \frac{d_n^{\widehat{\gamma}_1}}{d_n^{\gamma_1}} \frac{\widehat{\theta}_{k/n}}{\theta_{k/n}} \frac{d_n^{\gamma_1} \theta_{k/n}}{\theta_p} =: L_1 L_2 L_3$$

- L_1 is easy to deal.

$$\frac{\sqrt{k_1}}{\log d_n}(L_1 - 1) \xrightarrow{p} \Gamma$$

- L_2

$$\sqrt{k}\left(\frac{\widehat{\theta}_{k/n}}{\theta_{k/n}} - 1\right) \xrightarrow{\mathrm{d}} \Theta.$$

- $L_3 = 1 + o(1/\sqrt{k})$.

X Real

In this section, X takes values in $\mathbb{R}$. Define $X^+ = \max(X, 0)$ and $X^- = X - X^+$. We require two more conditions.

- (e) $E\left|X^-\right|^{1/\gamma} < \infty$.
- (f) As $n \to \infty, k = o\left(p^{2\tau(1-\gamma_1)}\right)$.

We estimate θ_p with

$$\hat{\theta}_p = \left(\frac{k}{np}\right)^{\hat{\gamma}_1} \frac{1}{k} \sum_{i=1}^{n} X_i I\left(X_i > 0, Y_i > Y_{n-k,n}\right).$$

Theorem 2 Under the conditions of theorem 1 and conditions (e) and (f), as $n \to \infty$,

$$\min\left\{\sqrt{k}, \frac{\sqrt{k_1}}{\log\left(d_n\right)}\right\}\left(\frac{\hat{\theta}_p}{\theta_p} - 1\right) \xrightarrow{\mathrm{d}} \begin{cases} \Theta + r\Gamma, & \text{if } r \leqslant 1 \\ (1/r)\Theta + \Gamma, & \text{if } r > 1 \end{cases}$$

A Sketch of Theorem 2

Write $\theta_p^+ := E\{X^+|Y > U_2(1/p)\}$. Then,

$$\frac{\hat{\theta}_p}{\theta_p} = \frac{\hat{\theta}_p}{\theta_p^+}\frac{\theta_p^+}{\theta_p}$$

- Prove that $\hat{\theta}_p/\theta_p^+$ follows the asymptotic normality stated in theorem 1. Verify the conditions.
- $\dfrac{\theta_p^+}{\theta_p} = 1 + o(1/\sqrt{k})$.

Simulation study

- We first generate data from three bivariate distributions.
- Throughout this section, (Z_1, Z_2) denotes a standard Cauchy distribution on $\mathbb{R}^2$ with density $1/(2\pi)(1+x^2+y^2)^{-3/2}$.
- We draw 500 samples from each distribution with sample sizes $n = 500$ and $n = 200$.
- On the basis of each sample, we estimate θ_p for p equal to $1/500, 1/5000$ and $1/10000$.

Simulation study

- The first distribution is a transformed Cauchy distribution on $(0, \infty)^2$ defined as
$$(X, Y) = (|Z_1|^{2/5}, |Z_2|).$$
It follows that $\gamma_1 = 2/5$ and $R(x, y) = x + y - \sqrt{(x^2 + y^2)}, x, y \geq 0$.

- The second distribution is a Student t_3 -distribution on $(0, \infty)^2$ with density
$$f(x, y) = \frac{2}{\pi}\left(1 + \frac{x^2 + y^2}{3}\right)^{-5/2}, \quad x, y > 0.$$
We have $\gamma_1 = 1/3, R(x, y) =$
$x + y - (x^{4/3} + 1/2x^{2/3}y^{2/3} + y^{4/3})/\sqrt{(x^{2/3} + y^{2/3})}$.

- The third distribution is a transformed Cauchy distribution on the whole $\mathbb{R}^2$ defined as
$$(X, Y) = \left(Z_1^{2/5} I\left(Z_1 \geqslant 0\right) + Z_1^{1/5} I\left(Z_1 < 0\right), Z_2 I\left(Z_1 \geqslant 0\right) + Z_2^{1/3} I\left(Z_1\right.\right.$$
We have $\gamma_1 = 2/5$, $R(x, y) = x/2 + y - \sqrt{(x^2/4 + y)}$

Simulation Study

Besides the estimator we propose, we construct two other estimators.

- for $np \geq 1$, an empirical counterpart of θ_p, given by

$$\hat{\theta}_{\text{emp}} = \frac{1}{\lfloor np \rfloor} \sum_{i=1}^{n} X_i I\left(Y_i > Y_{n-\lfloor np \rfloor, n}\right).$$

- Secondly, exploiting the relationship in proposition 1 and using the empirical estimator of R given by

$$\hat{R}(x,y) = \frac{1}{k} \sum_{i=1}^{n} I\left(X_i > X_{n-\lfloor kx \rfloor, n}, Y_i > Y_{n-\lfloor ky \rfloor, n}\right), \quad x, y \geqslant 0.$$

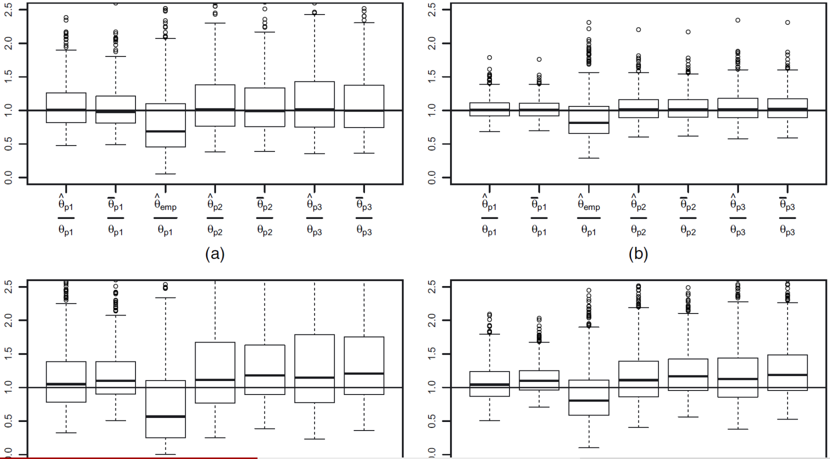
  and the Weissman (1978) estimator of $U_1(1/p)$ given by $\hat{U}_1(1/p) = d_n^{\hat{\gamma}_1} X_{n-k,n}$, we define an alternative EVT estimator as

$$\bar{\theta}_p = -\hat{U}_1\left(\frac{1}{p}\right) \int_0^{\infty} \hat{R}(x,1) \mathrm{d}x^{-\hat{\gamma}_1}$$

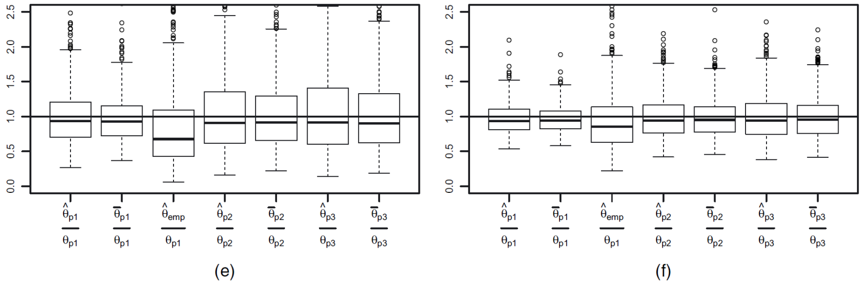
Simulation Study

$p_1 = 500, p_2 = 5000, p_3 = 10000$

transformed Cauchy distribution 1, $n = 500, k = 75, k_1 = 75$



Simulation Study



Simulation Study

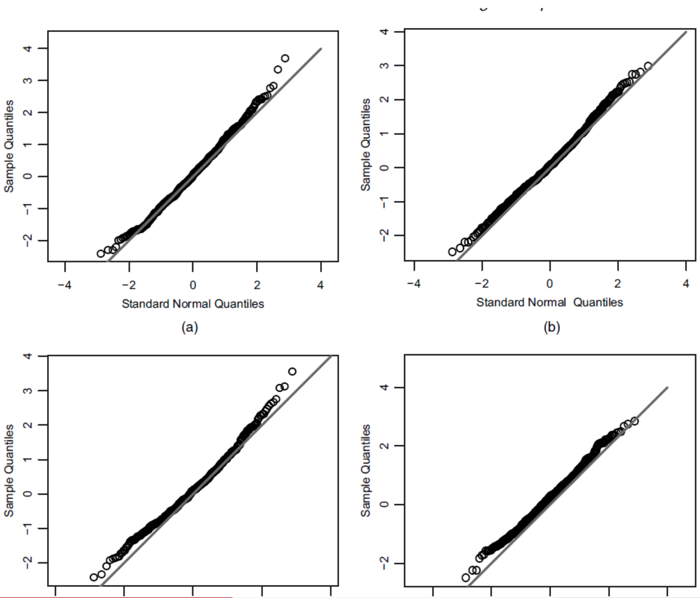
We write $\sigma_p^2 = \frac{1}{k}\mathrm{var}(\Theta + r\Gamma)$.

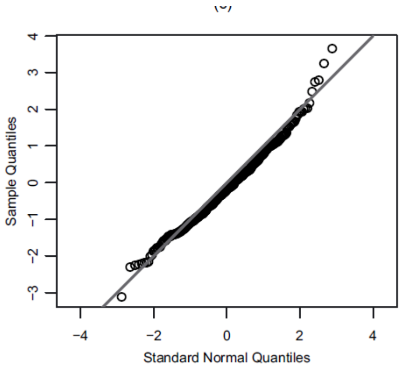
$$r = \frac{\sqrt{k}\log\{k/(np)\}}{\sqrt{k_1}}.$$

We compare the distribution of $(1/\sigma_p)\log(\hat{\theta}_p/\theta_p)$ with the limit distribution $N(0,1)$.

We use QQ plots to detect this.

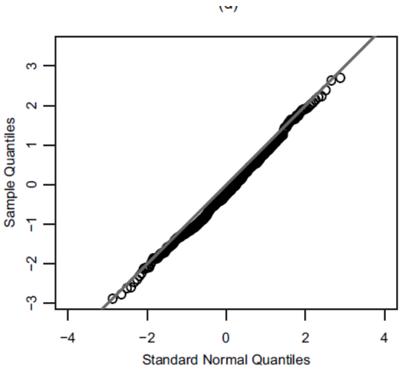
Simulation Study





Sample Quantiles
4
3
2
1
0
−1
−2
−3
−4
−2
0
2
4
Standard Normal Quantiles

(e)



Sample Quantiles
3
2
1
0
−1
−2
−3
−4
−2
0
2
4
Standard Normal Quantiles

(f)

Simulation Study

We also investigate the performance of our estimator when our assumptions are partially violated.

- The transformed Cauchy distribution 3 is defined as

$$(X, Y) = \left(|Z_1|^{0.7}, |Z_2|\right).$$

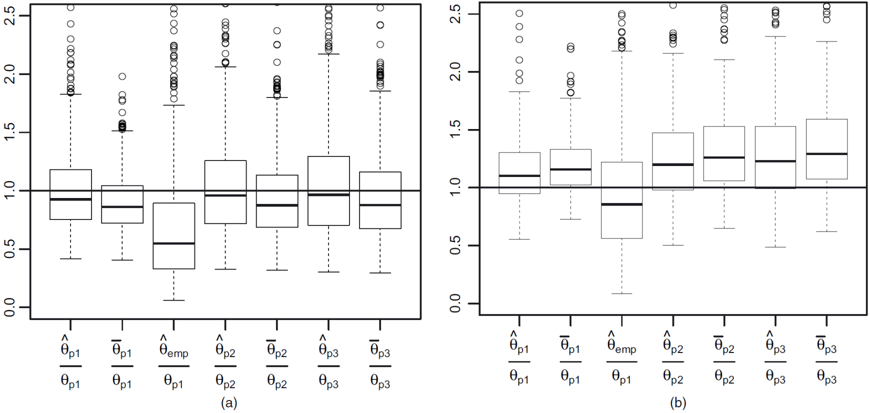
  The dependence structure between X and Y is the same as that of the transformed Cauchy distribution 1. $\gamma_1 = 0.7 > 1/2$.

- The second distribution is an asymptotically independent distribution defined as

$$(X, Y) = (V_1 + W_1, V_2 + W_2),$$

  where (V_1, V_2) follows the Student t_3-distribution with density (9) and W_1 and W_2 are Pareto distributed with density $(25/2)(1+5x)^{-7/2}, x > 0$. Moreover, $(V_1, V_2), W_1$ and W_2 are independent. This does not satisfy condition (a).

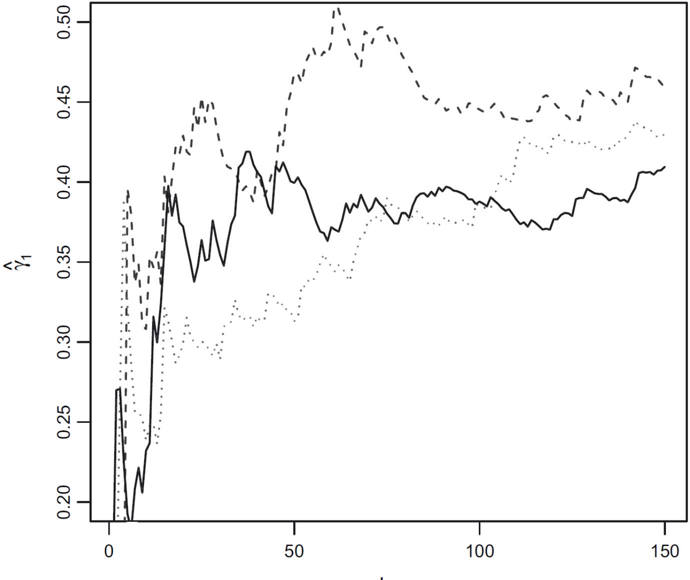
Simulation Study



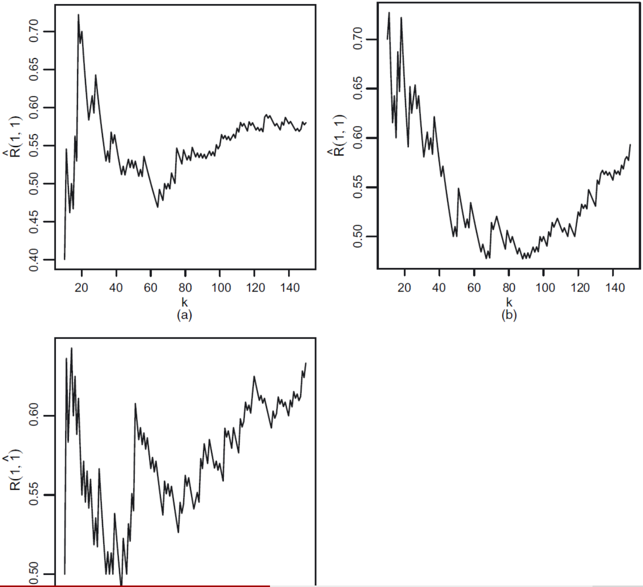
Application

- We apply our estimation method to estimate the MES for financial institutions
- We consider three large investment banks in the USA, namely Goldman Sachs, Morgan Stanley and T. Rowe Price.
- $p = 1/2513$, which corresponds to a once-per-decade systemic event.

Application



Application



Application

Table 1. MES of the three investment banks†

Bank	*Daily loss*		*Weekly loss*	
	$\hat{\gamma}_1$	$\hat{\theta}_p$	$\hat{\gamma}_1$	$\hat{\theta}_p$
Goldman Sachs	0.388	0.308	0.417	0.346
Morgan Stanley	0.465	0.608	0.483	0.654
T. Rowe Price	0.378	0.316	0.347	0.339

†The second and third columns report the results based on *daily* loss returns ($n = 2513$ and $p = 1/n$). The estimates $\hat{\gamma}_1$ are computed by taking the average for $k_1 \in [70, 100]$. The estimates of the MES are based on these values of $\hat{\gamma}_1$. We report the average of the MES estimates $\hat{\theta}_p$ for $k \in [70, 100]$. The last two columns report the results based on *weekly* loss returns from the same sample period ($n = 522$ and $p = 1/n$), where both k_1 and k are from the interval [20, 30].