Estimations of Extreme CoVaR and CoES under Asymptotic Independence

Qingzhao Zhong[1,2]

[1]*School of Data Science, Fudan University*

[2]*School of Economics and Management, Tsinghua University*

Supplementary Material

The supplementary material contains all technical proofs with some theoretical clarifications, also provides additional simulations and empirical analyses.

S1 Auxiliary conclusions

In this section, we derive a fundamental theoretical tool, Proposition S1, which serves as the key ingredient for establishing the corresponding asymptotic normalities of our proposed estimations. We define, for any $x, y \in (0, \infty)^2$,

$$T_{n/k}(x,y) := \left(\frac{n}{k}\right)^{\frac{1}{\eta}} \frac{1}{n} \sum_{i=1}^{n} I\left(\overline{F}_X(X_i) < \frac{kx}{n}, \overline{F}_Y(Y_i) < \frac{ky}{n}\right), \qquad \text{(S1.1)}$$

as the pseudo-non-parametric estimator of $C_{n/k}(x,y)$, *i.e.*,

$$C_{n/k}(x,y) = \left(\frac{n}{k}\right)^{\frac{1}{\eta}} \mathbb{P}\left(\overline{F}_X(X) < \frac{kx}{n}, \overline{F}_Y(Y) < \frac{ky}{n}\right).$$

It is because the marginal distributions $F_X(\cdot)$, $F_Y(\cdot)$ and coefficient η are all unknown. The following proposition presents the asymptotic behaviour of this pseudo estimator. The limit process is characterized by the centered Gaussian process $W_C(\cdot,\cdot)$ with covariance structure,

$$E\left[W_C(x_1,y_1)W_C(x_2,y_2)\right] = C(x_1 \wedge y_1, x_2 \wedge y_2).$$

Proposition S1. *Suppose that condition* (1.4) *holds. For any* $\beta \in [0, \frac{1}{2})$ *and* T *positive, we have that, as* $n \to \infty$,

$$\sup_{x,y\in(0,T]} \left| \frac{\sqrt{k}\left(\frac{n}{k}\right)^{\frac{1}{2}-\frac{1}{2\eta}}\left(T_{n/k}(x,y) - C_{n/k}(x,y)\right) - W_C(x,y)}{x^{\beta}} \right| = o_{\mathbb{P}}(1),$$

provided that $k = O(n^{\iota})$ *with* $\iota \in \left(\frac{1-\eta}{1-2\beta\eta}, 1\right)$.

Remark 1. Note that the nontrivial lower bound of ι is primarily determined to ensure that the conditions of Theorem 2.11.9 in Van der Vaart and Wellner (1996) are satisfied. The existence of this lower bound also ensures that the convergence rate $\sqrt{k}\left(\frac{n}{k}\right)^{\frac{1}{2}-\frac{1}{2\eta}}$ diverges as n increases. Indeed, it follows that,

$$\sqrt{k}\left(\frac{n}{k}\right)^{\frac{1}{2}-\frac{1}{2\eta}} = O\left(n^{\frac{\iota}{2}+(1-\iota)\left(\frac{1}{2}-\frac{1}{2\eta}\right)}\right) = O\left(n^{\frac{1}{2}\left[\frac{\iota-1}{\eta}+1\right]}\right) \to \infty,$$

as $n \to \infty$, since $\frac{\iota-1}{\eta} + 1 > 0$ by noting that $0 < 1-\eta < \frac{1-\eta}{1-2\beta\eta} < \iota$.

Proof of Proposition S1. Define $U_i = \overline{F}_X(X_i)$, $V_i = \overline{F}_Y(Y_i)$, and

$$Z_{ni} = \frac{\sqrt{k}}{n}\left(\frac{n}{k}\right)^{\frac{1}{2}+\frac{1}{2\eta}} \delta\left((n/k)U_i, (n/k)V_i\right),$$

where $\delta(\cdot,\cdot)$ denotes Dirac measure. For all $(x, y) \in (0, T]^2$, define the function,

$$f_{x,y}(s,t) = \frac{I(0 \le s < x, 0 \le t < y)}{x^\beta},$$

and functional class $\mathcal{F} = \{f_{x,y},\ (x, y) \in (0, T]^2\}$. We equip $\mathcal{F}$ with the semi-metric d defined by

$$d(f_{x,y}, f_{u,v}) = \sqrt{E\left[\left(\frac{W_C(x,y)}{x^\beta} - \frac{W_C(u,v)}{u^\beta}\right)^2\right]}.$$

Let $Z_{ni}(f) = \int f\, dZ_{ni}$ and $\{Z_{ni}(f),\ f \in \mathcal{F}\}_{i=1}^n$ be independent stochastic process with finite second moments indexed by a totally bounded semi-metric space $(\mathcal{F}, d)$ with supremum norm $\|Z_{ni}\|_{\mathcal{F}} = \sup_{f\in\mathcal{F}} |Z_{ni}(f)|$. We aim to apply Theorem 2.11.9 in Van der Vaart and Wellner (1996) and prove that the five conditions of this theorem are met.

Part I. We briefly sketch the total boundedness of $(\mathcal{F}, d)$. We restrict ourselves to the case $x \ge y$, $u \ge v$ and $x \ge u$, $y \ge v$. For any $\delta > 0$, assuming $|x - u| \le \delta$ and $|y - v| \le \delta$, we have

$$d^2(f_{x,y}, f_{u,v}) = E\left[\left(\frac{W_C(x,y)}{x^\beta} - \frac{W_C(u,v)}{u^\beta}\right)^2\right] = \frac{C(x,y)}{x^{2\beta}} + \frac{C(u,v)}{u^{2\beta}} - \frac{2C(u,v)}{(xu)^\beta}.$$

On the one hand, if $u \le \delta$, then there exists a absolute constant K such that,

$$\begin{aligned} d^2(f_{x,y}, f_{u,v}) &\leq \frac{C(x,y)}{x^{2\beta}} + \frac{C(u,v)}{u^{2\beta}} + \frac{2C(u,v)}{u^{2\beta}} \\ &\leq C(1, y/x)x^{\frac{1}{\eta}-2\beta} + C(1, v/u)u^{\frac{1}{\eta}-2\beta} + 2C(1, v/u)u^{\frac{1}{\eta}-2\beta} \\ &\leq K\left((2\delta)^{\frac{1}{\eta}-2\beta} + \delta^{\frac{1}{\eta}-2\beta}\right) \\ &\leq K\delta^{\frac{1}{\eta}-2\beta}. \end{aligned}$$

Hence, since $\frac{1}{\eta} - 2\beta > 0$, we see that, for every $\varepsilon > 0$, we can find a $\delta > 0$ such that, for $|x-u| \leq \delta$ and $|y-v| \leq \delta$, it follows that $d^2(f_{x,y}, f_{u,v}) < \varepsilon$.

On the other hand, if $u > \delta$, then,

$$\begin{aligned} & C_{n/k}(x,y) - C_{n/k}(u,v) \\ \leq & \left(\frac{n}{k}\right)^{\frac{1}{\eta}} \mathbb{P}\left(\frac{ku}{n} \leq \overline{F}_X(X) < \frac{kx}{n}, \overline{F}_Y(Y) < \frac{ky}{n}\right) \\ & + \left(\frac{n}{k}\right)^{\frac{1}{\eta}} \mathbb{P}\left(\overline{F}_X(X) < \frac{kx}{n}, \frac{kv}{n} \leq \overline{F}_Y(Y) < \frac{ky}{n}\right) \\ \leq & \left(\frac{n}{k}\right)^{\frac{1}{\eta}-1}(x-u) + \left(\frac{n}{k}\right)^{\frac{1}{\eta}-1}(y-v) \\ \leq & 2\left(\frac{n}{k}\right)^{\frac{1}{\eta}-1}\delta. \end{aligned}$$

By taking limit on both of sides and choosing $\delta := \delta(n) = \left(\frac{k}{n}\right)^{\epsilon}$ with $\epsilon > \frac{1/\eta - 1}{1-2\beta}$, there exists a $\delta' = O\left(\left(\frac{n}{k}\right)^{\frac{1}{\eta}-1}\delta\right)$ such that $C(x,y) \leq C(u,v) + 2\delta'$. We then have,

$$\begin{aligned} d^2(f_{x,y}, f_{u,v}) &\leq \frac{C(u,v)(u^\beta - x^\beta)^2}{(xu)^{2\beta}} + \frac{2\delta'}{x^{2\beta}} \\ &\leq Ku^{\frac{1}{\eta}-4\beta}(u^\beta - x^\beta)^2 + 2\delta'\delta^{-2\beta} \end{aligned}$$

$$\leq K u^{\frac{1}{\eta}-4\beta} u^{2\beta-2}(u-x)^2 + O\left(\left(\frac{n}{k}\right)^{\frac{1}{\eta}-1} \delta^{1-2\beta}\right)$$
$$\leq K u^{\frac{1}{\eta}-2\beta-2}(u-x)^2 + O\left(\left(\frac{n}{k}\right)^{\frac{1}{\eta}-1} \left(\frac{k}{n}\right)^{\epsilon(1-2\beta)}\right)$$
$$\leq K \delta^{\frac{1}{\eta}-2\beta} + O\left(\left(\frac{k}{n}\right)^{\epsilon(1-2\beta)-(\frac{1}{\eta}-1)}\right).$$

Thus, since $\epsilon(1-2\beta)-(\frac{1}{\eta}-1) > 0$, we can say that, for every $\varepsilon > 0$, we can find a $\delta > 0$ such that, for $|x-u| \leq \delta$ and $|y-v| \leq \delta$, it follows $d^2(f_{x,y}, f_{u,v}) < \varepsilon$. Therefore, since $[0,T]^2$ is totally bounded with respect to the Euclidean metric, we obtain the total boundedness of $(\mathcal{F}, d)$.

Part II. We observe that

$$Z_{ni}(f_{x,y}) = \frac{\sqrt{k}}{n}\left(\frac{n}{k}\right)^{\frac{1}{2}+\frac{1}{2\eta}} \frac{I\left(U_i < (k/n)x, V_i < (k/n)y\right)}{x^\beta},$$

and

$$\sum_{i=1}^{n}(Z_{ni} - E[Z_{ni}])(f_{x,y}) = \sqrt{k}\left(\frac{n}{k}\right)^{\frac{1}{2}-\frac{1}{2\eta}} \frac{(T_{n/k}(x,y) - C_{n/k}(x,y))}{x^\beta},$$

and similarly for the marginal processes. We now show that, for every $\lambda > 0$,

$$\sum_{i=1}^{n} E\left[\|Z_{ni}\|_{\mathcal{F}} I\left(\|Z_{ni}\|_{\mathcal{F}} > \lambda\right)\right] \to 0, \tag{S1.2}$$

as $n \to \infty$. Note that,

$$\sup_{f_{x,y}\in\mathcal{F}} \frac{\sqrt{k}}{n}\left(\frac{n}{k}\right)^{\frac{1}{2}+\frac{1}{2\eta}} \frac{I\left(U_i < (k/n)x, V_i < (k/n)y\right)}{x^\beta} \leq \frac{\sqrt{k}}{n}\left(\frac{n}{k}\right)^{\frac{1}{2}+\frac{1}{2\eta}} \frac{1}{((n/k)(U_i))^\beta},$$

so for each $\lambda > 0$,

$$\begin{aligned}
&\sum_{i=1}^{n} E\left[\|Z_{ni}\|_{\mathcal{F}} I\left(\|Z_{ni}\|_{\mathcal{F}} > \lambda\right)\right] \\
\leq&\sqrt{k}\left(\frac{n}{k}\right)^{\frac{1}{2}+\frac{1}{2\eta}} E\left[\frac{1}{((n/k)U_1)^{\beta}} I\left((n/k)U_1 < \left(\sqrt{n}\lambda\,(k/n)^{\frac{1}{2\eta}}\right)^{-1/\beta}\right)\right] \\
=&\sqrt{k}\left(\frac{n}{k}\right)^{\frac{1}{2}+\frac{1}{2\eta}} \int_0^{\left(\sqrt{n}\lambda(k/n)^{\frac{1}{2\eta}}\right)^{-1/\beta}} x^{-\beta}\, d\mathbb{P}\left(U_1 < kx/n\right) \\
=&\sqrt{k}\left(\frac{n}{k}\right)^{\frac{1}{2}+\frac{1}{2\eta}} \left[\sqrt{n}\lambda\,(k/n)^{\frac{1}{2\eta}}\, \mathbb{P}\left(U_1 < \frac{k}{n}\left(\sqrt{n}\lambda\,(k/n)^{\frac{1}{2\eta}}\right)^{-1/\beta}\right)\right] \\
&+\beta\sqrt{k}\left(\frac{n}{k}\right)^{\frac{1}{2}+\frac{1}{2\eta}} \int_0^{\left(\sqrt{n}\lambda(k/n)^{\frac{1}{2\eta}}\right)^{-1/\beta}} \mathbb{P}\left(U_1 < kx/n\right) x^{-\beta-1}\, dx \\
\leq&\sqrt{k}\left(\frac{n}{k}\right)^{\frac{1}{2}+\frac{1}{2\eta}} \left[\sqrt{n}\lambda\,(k/n)^{\frac{1}{2\eta}}\, \frac{k}{n}\left(\sqrt{n}\lambda\,(k/n)^{\frac{1}{2\eta}}\right)^{-1/\beta}\right] \\
&+\beta\sqrt{k}\left(\frac{n}{k}\right)^{\frac{1}{2}+\frac{1}{2\eta}} \int_0^{\left(\sqrt{n}\lambda(k/n)^{\frac{1}{2\eta}}\right)^{-1/\beta}} \frac{k}{n} x^{-\beta}\, dx \\
=&\lambda^{1-1/\beta} n^{\iota\left(1-\frac{1}{2\beta\eta}\right)+\frac{1}{2\beta}\left(\frac{1}{\eta}-1\right)} + \frac{\beta\lambda^{1-1/\beta}}{1-\beta} n^{\iota\left(1-\frac{1}{2\beta\eta}\right)+\frac{1}{2\beta}\left(\frac{1}{\eta}-1\right)} \\
\to&0,
\end{aligned}$$

as $n \to \infty$, by recalling $\iota > \frac{1-\eta}{1-2\beta\eta}$ and $2\beta\eta < 1$, which implies the exponent

$$\iota\left(1-\frac{1}{2\beta\eta}\right)+\frac{1}{2\beta}\left(\frac{1}{\eta}-1\right) < \frac{1-\eta}{1-2\beta\eta}\left(1-\frac{1}{2\beta\eta}\right)+\frac{1}{2\beta}\left(\frac{1}{\eta}-1\right) = 0.$$

Hence, (S1.2) holds.

Part III. We now show that, for every $\delta \downarrow 0$,

$$\sup_{d(f_{x,y}, f_{u,v}) < \delta} \sum_{i=1}^{n} E\left[\left(Z_{ni}(f_{x,y}) - Z_{ni}(f_{u,v})\right)^2\right] \to 0. \tag{S1.3}$$

Note that $d(f_{x,y}, f_{u,v}) < \delta$ means

$$\begin{aligned} d^2(f_{x,y}, f_{u,v}) &= E\left[\left(\frac{W_C(x,y)}{x^\beta} - \frac{W_C(u,v)}{u^\beta}\right)^2\right] \\ &= \frac{C(x,y)}{x^{2\beta}} + \frac{C(u,v)}{u^{2\beta}} - \frac{2C(x\wedge u, y\wedge v)}{x^\beta u^\beta} \\ &\leq \delta^2. \end{aligned}$$

We hence get,

$$\begin{aligned} &\sum_{i=1}^n E\left[(Z_{ni}(f_{x,y}) - Z_{ni}(f_{u,v}))^2\right] \\ =&nE\left[(Z_{n1}(f_{x,y}) - Z_{n1}(f_{u,v}))^2\right] \\ =&n\frac{k}{n^2}\left(\frac{n}{k}\right)^{1+\frac{1}{\eta}} E\left[\frac{I\left(U_i < (k/n)x, V_i < (k/n)y\right)}{x^{2\beta}} + \frac{I\left(U_i < (k/n)u, V_i < (k/n)v\right)}{u^{2\beta}}\right. \\ &\left.-\frac{2I\left(U_i < (k/n)x, V_i < (k/n)y\right) I\left(U_i < (k/n)u, V_i < (k/n)v\right)}{x^\beta u^\beta}\right] \\ =&\left(\frac{n}{k}\right)^{\frac{1}{\eta}} E\left[\frac{I\left(U_i < (k/n)x, V_i < (k/n)y\right)}{x^{2\beta}} + \frac{I\left(U_i < (k/n)u, V_i < (k/n)v\right)}{u^{2\beta}}\right. \\ &\left.-\frac{2I\left(U_i < (k/n)(x\wedge u), V_i < (k/n)(y\wedge v)\right)}{x^\beta u^\beta}\right] \\ \to&\frac{C(x,y)}{x^{2\beta}} + \frac{C(u,v)}{u^{2\beta}} - \frac{2C(x\wedge u, y\wedge v)}{x^\beta u^\beta} \\ \leq&\delta^2. \end{aligned}$$

Hence, (S1.3) holds.

Part IV. For any $\varepsilon > 0$, the bracketing number $N_{[\,]}(\varepsilon, \mathcal{F}, L_2^n)$ is the minimal number of sets N_ε in a partition $\mathcal{F} = \bigcup_{j=1}^{N_\varepsilon} \mathcal{F}_{\varepsilon j}^n$ of the index set into sets $\mathcal{F}_{\varepsilon j}^n$ such that, for every partitioning set $\mathcal{F}_{\varepsilon j}^n$,

$$\sum_{i=1}^n E\left[\sup_{f,g\in\mathcal{F}_{\varepsilon j}^n} |Z_{ni}(f)-Z_{ni}(g)|^2\right] \le \varepsilon^2. \tag{S1.4}$$

Next, we want to prove, for every $\delta_n \downarrow 0$,

$$\int_0^{\delta_n} \sqrt{\log N_{[\ ]}(\varepsilon, \mathcal{F}, L_2^n)}\, d\varepsilon \to 0. \tag{S1.5}$$

For notational convenience, we choose $T=1$; for general $T>0$, the proof goes the same. Let $\varepsilon > 0$ be a small number, define $a_n = \left(\frac{n}{k}\right)^{\frac{1-1/\eta}{1-2\beta}} \varepsilon^{3/(1-2\beta)}$ and $\theta_n = 1 - \left(\frac{n}{k}\right)^{1-1/\eta} \varepsilon^3$. Define

$$\mathcal{F}_n(a_n) = \{f_{x,y} \in \mathcal{F} : x \wedge y \le a_n\},$$

$$\mathcal{F}_n(l,m) = \{f_{x,y} \in \mathcal{F} : \theta_n^{l+1} \le x \le \theta_n^l, \theta_n^{m+1} \le y \le \theta_n^m\}.$$

Then, for any n, it follows that,

$$\mathcal{F} = \mathcal{F}_n(a_n) \bigcup \left(\bigcup_{m=0}^{[\log a_n/\log\theta_n]} \bigcup_{l=0}^{[\log a_n/\log\theta_n]} \mathcal{F}_n(l,m) \right).$$

We first check (S1.4) for $\mathcal{F}_n(a_n)$,

$$\begin{aligned}
&\sum_{i=1}^n E\left[\sup_{f,g\in\mathcal{F}_n(a_n)} (Z_{ni}(f)-Z_{ni}(g))^2\right] \\
=&nE\left[\sup_{f,g\in\mathcal{F}_n(a_n)} (Z_{n1}(f)-Z_{n1}(g))^2\right] \\
\le&4nE\left[\sup_{f\in\mathcal{F}_n(a_n)} Z_{n1}^2(f)\right] \\
=&4\left(\frac{n}{k}\right)^{\frac{1}{\eta}} E\left[\sup_{x,y>0, x\wedge y\le a_n} \frac{I(U_1<kx/n, V_1<ky/n)}{x^{2\beta}}\right]
\end{aligned}$$

$$
\begin{aligned}
&\leq 4\left(\frac{n}{k}\right)^{\frac{1}{\eta}} E\left[\left(\frac{n}{k}U_1\right)^{-2\beta} I((n/k)U_1 < a_n)\right]\\
&=4\left(\frac{n}{k}\right)^{\frac{1}{\eta}} \int_0^{a_n k/n} \left(\frac{n}{k}x\right)^{-2\beta} dx\\
&=\frac{4}{1-2\beta}\left(\frac{n}{k}\right)^{\frac{1}{\eta}-1} a_n^{1-2\beta}\\
&\leq \varepsilon^2.
\end{aligned}
$$

We next check (S1.4) for $\mathcal{F}_n(l,m)$, without loss of generality, we take $l \leq m$,

$$
\begin{aligned}
&\sum_{i=1}^n E\left[\sup_{f,g\in\mathcal{F}_n(l.m)} (Z_{ni}(f)-Z_{ni}(g))^2\right]\\
\leq& nE\left[\left(\sup_{f\in\mathcal{F}_n(l,m)} Z_{n1}(f) - \inf_{f\in\mathcal{F}_n(l,m)} Z_{n1}(f)\right)^2\right]\\
\leq& \left(\frac{n}{k}\right)^{\frac{1}{\eta}} E\left[\left(\frac{I(U_1<(k/n)\theta_n^l, V_1<(k/n)\theta_n^m)}{(\theta_n^{l+1})^\beta} - \frac{I(U_1<(k/n)\theta_n^{l+1}, V_1<(k/n)\theta_n^{m+1})}{(\theta_n^l)^\beta}\right)^2\right]\\
\leq& \left(\frac{n}{k}\right)^{\frac{1}{\eta}} E\left[\left(I(U_1<(k/n)\theta_n^l, V_1<(k/n)\theta_n^m)\left(\frac{1}{\theta_n^{\beta(l+1)}} - \frac{1}{\theta_n^{\beta l}}\right)\right.\right.\\
&\left.\left.+\ \left(I(U_1<(k/n)\theta_n^l, V_1<(k/n)\theta_n^m) - I(U_1<(k/n)\theta_n^{l+1}, V_1<(k/n)\theta_n^{m+1})\right)\frac{1}{\theta_n^{\beta l}}\right)^2\right]\\
\leq& 2\left(\frac{n}{k}\right)^{\frac{1}{\eta}}\left\{\mathbb{P}\left(U_1<\frac{k}{n}\theta_n^l, V_1<\frac{k}{n}\theta_n^m\right)\frac{1}{\theta_n^{2\beta l}}\left(\frac{1}{\theta_n^\beta}-1\right)^2\right.\\
&\left.+\left[\mathbb{P}\left(U_1<\frac{k}{n}\theta_n^l, V_1<\frac{k}{n}\theta_n^m\right) - \mathbb{P}\left(U_1<\frac{k}{n}\theta_n^{l+1}, V_1<\frac{k}{n}\theta_n^{m+1}\right)\right]\frac{1}{\theta_n^{2\beta l}}\right\}\\
\leq& 2\left(\frac{n}{k}\right)^{\frac{1}{\eta}}\left\{\frac{k}{n}\frac{\theta_n^l}{\theta_n^{2\beta l}}\left(\frac{1}{\theta_n^\beta}-1\right)^2 + \frac{2k}{n}\frac{\theta_n^l(1-\theta_n)}{\theta_n^{2\beta l}}\right\}\\
\leq& 2\left(\frac{k}{n}\right)^{1-\frac{1}{\eta}}\left\{\left(\frac{1}{\theta_n^{1/2}}-1\right)^2 + 2(1-\theta_n)\right\}\\
\leq& 2\left(\frac{k}{n}\right)^{1-\frac{1}{\eta}}(1-\theta_n)^2 + 4\left(\frac{k}{n}\right)^{1-\frac{1}{\eta}}(1-\theta_n)
\end{aligned}
$$

$$=2\left(\frac{n}{k}\right)^{1-\frac{1}{\eta}}\varepsilon^6+4\varepsilon^3$$
$$\leq\varepsilon^2.$$

Then, for any ε, we can find n such that $\varepsilon < (n/k)^{1-1/\eta}$, which implies that $a_n > \varepsilon^{\frac{4}{1-2\beta}} =: a$ and $\theta_n < 1-\varepsilon^4 =: \theta$. Therefore, the bracketing number satisfies $N_{[\,]}(\varepsilon,\mathcal{F},L_2^n) < \left(\frac{\log a}{\log\theta}+1\right)^2+1$, which can be bounded by ε^{-10}, meeting with (S1.5).

Part V. It remains to prove the marginal converge, that is, for each $M\in\mathbb{N}$ and for each $f_{x_1,y_1},...,f_{x_M,y_M}\in\mathcal{F}$, the random vector

$$\left(\sum_{i=1}^n\left(Z_{ni}(f_{x_1,y_1})-E[Z_{ni}(f_{x_1,y_1})]\right),...,\sum_{i=1}^n\left(Z_{ni}(f_{x_M,y_M})-E[Z_{ni}(f_{x_M,y_M})]\right)\right),$$

converges to a multivariate normal distribution. It suffices to show that, for each $a_1,...,a_M\in\mathbb{R}$, we have,

$$\sum_{j=1}^M a_j\left[\sum_{i=1}^n\left(Z_{ni}(f_{x_j,y_j})-E[Z_{ni}(f_{x_j,y_j})]\right)\right]=:\sum_{i=1}^n\left(N_{ni}-E[N_{ni}]\right),$$

converges to a normal distribution, where $N_{ni}=\sum_{j=1}^M a_jZ_{ni}(f_{x_j,y_j})$. This will follow from the Lindeberg-Feller central limit theorem (for example, see Van der Vaart (2000)). We need to show that, for each $\varepsilon>0$,

$$\sum_{i=1}^n E\left[|N_{ni}|^2I(|N_{ni}|>\varepsilon)\right]\to 0, \tag{S1.6}$$

and

$$\sum_{i=1}^n \mathrm{Var}\left(N_{ni}\right)\to\sigma^2. \tag{S1.7}$$

First, we observe that there exists a absolute constant K, such that,

$$\begin{aligned}
&\sum_{i=1}^{n} E\left[|N_{ni}|^2 I(|N_{ni}|>\varepsilon)\right] \\
=& nE\left[|N_{n1}|^2 I(|N_{n1}|>\varepsilon)\right] \\
\leq& \frac{nE[|N_{n1}|^{2+\delta}]}{\varepsilon^{\delta}} \\
\leq& Kn\sum_{j=1}^{M}|a_j|^{2+\delta}E\left[|Z_{ni}(f_{x_j,y_j})|^{2+\delta}\right] \\
\leq& Kn\left(\frac{\sqrt{k}}{n}\left(\frac{n}{k}\right)^{\frac{1}{2}+\frac{1}{2\eta}}\right)^{2+\delta}\sum_{j=1}^{M}\frac{|a_j|^{2+\delta}}{x_j^{\beta(2+\delta)}}\mathbb{P}(U_1<kx_j/n, V_1<ky_j/n) \\
=& K\left(\frac{\sqrt{k}}{n}\left(\frac{n}{k}\right)^{\frac{1}{2}+\frac{1}{2\eta}}\right)^{\delta}\sum_{j=1}^{M}\frac{|a_j|^{2+\delta}}{x_j^{\beta(2+\delta)}}\left(\frac{n}{k}\right)^{\frac{1}{\eta}}\mathbb{P}(U_1<kx_j/n, V_1<ky_j/n) \\
=& K\left(n^{\frac{1}{2\eta}-\frac{1}{2}-\frac{\iota}{2\eta}}\right)^{\delta}\sum_{j=1}^{M}\frac{|a_j|^{2+\delta}}{x_j^{\beta(2+\delta)}}\left(\frac{n}{k}\right)^{\frac{1}{\eta}}\mathbb{P}(U_1<kx_j/n, V_1<ky_j/n) \\
\to& 0,
\end{aligned}$$

as $n\to\infty$, by recalling $\iota>1-\eta$ and hence $\frac{1}{2\eta}-\frac{1}{2}-\frac{\iota}{2\eta}<0$. Second, we note that,

$$\sqrt{n}E[Z_{n1}(f_{x_j,y_j})]=\left(\frac{n}{k}\right)^{-\frac{1}{2\eta}}\left(\frac{n}{k}\right)^{\frac{1}{\eta}}\frac{\mathbb{P}(U_1<kx_j/n, V_1<kx_j/n)}{x_j^{\beta}}\to 0,$$

as $n\to\infty$. Therefore, it follows that,

$$\begin{aligned}
&\sum_{i=1}^{n}\mathrm{Var}\left(N_{ni}\right) \\
=& n\left\{E\left[\left(\sum_{j=1}^{M}a_j Z_{n1}(f_{x_j,y_j})\right)^2\right]-\left(E\left[\sum_{j=1}^{m}a_j Z_{ni}(f_{x_j,y_j})\right]\right)^2\right\}
\end{aligned}$$

$$
\begin{aligned}
&= nE\left[\sum_{j,k=1}^{M} a_j a_k Z_{n1}(f_{x_j,y_j}) Z_{n1}(f_{x_k,y_k})\right] - \left(\sum_{j=1}^{M} a_j \sqrt{n} E[Z_{n1}(f_{x_j,y_j})]\right)^2 \\
&= \sum_{j,k=1}^{M} a_j a_k nE\left[Z_{n1}(f_{x_j,y_j}) Z_{n1}(f_{x_k,y_k})\right] + o(1) \\
&= \sum_{j,k=1}^{M} a_j a_k \left(\frac{n}{k}\right)^{\frac{1}{\eta}} \frac{\mathbb{P}(U_1 < (k/n)(x_j \wedge x_k), V_1 < (k/n)(y_j \wedge y_k)))}{x_j^\beta x_k^\beta} + o(1) \\
&\to \sum_{j,k=1}^{M} a_j a_k \frac{C(x_j \wedge x_k, y_j \wedge y_k)}{x_j^\beta x_k^\beta} \\
&=: \sigma^2,
\end{aligned}
$$

as $n \to \infty$.

We have now verified the five conditions required by Theorem 2.11.9 of Van der Vaart and Wellner (1996), which leads to the conclusion that $\sum_{i=1}^{n}(Z_{ni} - E[Z_{ni}])(f_{x,y})$ is asymptotically tight in $\ell^\infty(\mathcal{F})$ and converges in distribution to a Gaussian process. Finally, we compute the covariance structure of the limit process. For each $f_{x_1,y_1}, f_{x_2,y_2} \in \mathcal{F}$, denote

$$
\widetilde{Z}_{ni}(f_{x,y}) = x^\beta Z_{ni}(f_{x,y}) = \frac{\sqrt{k}}{n}\left(\frac{n}{k}\right)^{\frac{1}{2}+\frac{1}{2\eta}} I\left(U_i < (k/n)x, V_i < (k/n)y\right),
$$

we then have,

$$
\begin{aligned}
& E[W_C(x_1,y_1)W_C(x_2,y_2)] \\
=& \lim_{n\to\infty} \operatorname{Cov}\Bigg(\sqrt{k}\left(\frac{n}{k}\right)^{\frac{1}{2}-\frac{1}{2\eta}} (T_{n/k}(x_1,y_1) - C_{n/k}(x_1,y_1)), \\
& \sqrt{k}\left(\frac{n}{k}\right)^{\frac{1}{2}-\frac{1}{2\eta}} (T_{n/k}(x_2,y_2) - C_{n/k}(x_2,y_2))\Bigg)
\end{aligned}
$$

$$
\begin{aligned}
&= \lim_{n\to\infty} \operatorname{Cov}\left(\sum_{i=1}^{n}(\widetilde{Z}_{ni}(f_{x_1,y_1}) - E[\widetilde{Z}_{ni}(f_{x_1,y_1}]), \sum_{i=1}^{n}(\widetilde{Z}_{ni}(f_{x_2,y_2}) - E[\widetilde{Z}_{ni}(f_{x_2,y_2}])\right)\\
&= \lim_{n\to\infty} \operatorname{Cov}\left(\sum_{i=1}^{n}\widetilde{Z}_{ni}(f_{x_1,y_1}), \sum_{i=1}^{n}\widetilde{Z}_{ni}(f_{x_2,y_2})\right)\\
&= \lim_{n\to\infty} n\operatorname{Cov}\left(\widetilde{Z}_{n1}(f_{x_1,y_1}), \widetilde{Z}_{n1}(f_{x_2,y_2})\right)\\
&= \lim_{n\to\infty} \left(nE\left[\widetilde{Z}_{n1}(f_{x_1,y_1})\widetilde{Z}_{n1}(f_{x_2,y_2})\right] - nE[\widetilde{Z}_{n1}(f_{x_1,y_1})]E[\widetilde{Z}_{n1}(f_{x_2,y_2})]\right)\\
&=C(x_1 \wedge x_2, y_1 \wedge y_2).
\end{aligned}
$$

This proof is therefore complete. □

The following corollary summarizes the one-sided case.

Corollary 1. *Under the conditions of Proposition S1, we have that as* $n \to \infty$,

$$
\sup_{s\in(0,T]}\left|\frac{\sqrt{k}\left(\frac{n}{k}\right)^{\frac{1}{2}-\frac{1}{2\eta}}(T_{n/k}((k/n)^{\frac{1}{\eta}-1}s,\infty) - s) - W(s)}{s^{\beta}}\right| = o_{\mathbb{P}}(1), \tag{S1.8}
$$

where $W(\cdot)$ *is a centered univariate Gaussian process with covariance structure* $E[W(s_1)W(s_2)] = s_1 \wedge s_2$.

Remark 2. Note that, for any finite s, it is problematic to directly regard the one-sided limit process as $W_C(s,\infty)$. It is because the variance of $W_C(s,\infty)$ characterized by limit (1.4) is always divergent. Therefore, one needs to multiply s by a suitable rate $(k/n)^{\frac{1}{\eta}-1}$ to ensure that the limit process becomes a non-degenerate Gaussian process.

The following lemma shows the boundedness of the $W_C(\cdot,\cdot)$ with proper weighting function.

Lemma S1. *For any $T>0$ and $\mu\in(0,1/2]$, with probability 1,*

$$\sup_{x,y\in(0,T]}\frac{|W_C(x,y)|}{x^{\mu}}<\infty. \tag{S1.9}$$

Proof of Lemma S1. Define

$$\mathcal{B}_m=\left\{(x,y):\frac{T}{2^{m+1}}\leq x\leq\frac{T}{2^m},\frac{T}{2^{m+1}}\leq y\leq T\right\},$$

for $m\in\mathbb{N}_0=\{0,1,2,...\}$. Denote Z as a standard normal random variable, it follows that, for a $\lambda>0$,

$$\begin{aligned}
&\mathbb{P}\left(\sup_{x,y\in(0,T]}\frac{|W_C(x,y)|}{x^{\mu}}\geq\lambda\right)\\
=&\mathbb{P}\left(\sup_{m\in\mathbb{N}_0}\sup_{(x,y)\in\mathcal{B}_m}\frac{|W_C(x,y)|}{x^{\mu}}\geq\lambda\right)\\
\leq&\sum_{m=0}^{\infty}\mathbb{P}\left(\sup_{(x,y)\in\mathcal{B}_m}\frac{|W_C(x,y)|}{x^{\mu}}\geq\lambda\right)\\
\leq&\sum_{m=0}^{\infty}\mathbb{P}\left(\sup_{(x,y)\in\mathcal{B}_m}|W_C(x,y)|\geq\lambda\left(\frac{T}{2^{m+1}}\right)^{\mu}\right)\\
\leq&4\sum_{m=0}^{\infty}\mathbb{P}\left(\left|W_C\left(\frac{T}{2^m},T\right)\right|\geq\lambda\left(\frac{T}{2^{m+1}}\right)^{\mu}\right)\\
\leq&4\sum_{m=0}^{\infty}\mathbb{P}\left(|Z|\geq\lambda\frac{T^{\mu-\frac{1}{2\eta}}}{2^{\mu(m+1)}}\frac{1}{\left(C(2^{-m},1)\right)^{1/2}}\right)\\
\leq&8\sum_{m=0}^{\infty}\frac{2^{\mu(m+1)}}{T^{\mu-\frac{1}{2\eta}}}\frac{\left(C(2^{-m},1)\right)^{1/2}}{\lambda}\exp\left\{-\lambda^2\frac{T^{2\mu-\frac{1}{\eta}}}{4^{\mu+1}}\frac{1}{C\left(2^{m(2\mu\eta-1)},2^{2m\mu\eta}\right)}\right\},
\end{aligned}$$

where the third inequality follows from the Lemma 1.2 in Orey and Pruitt (1973) and the last inequality follows from the Mill's inequality. It is readily to check that $\mathbb{P}\left(\sup_{x,y\in(0,T]}\frac{|W_C(x,y)|}{x^\mu}\geq\lambda\right)\to 0$ by letting $\lambda\to\infty$ and $\mu<\frac{1}{2\eta}$.

□

Define, for any $\tau\in(0,1)$,

$$\begin{cases}\hat{\theta}_X(\tau):=\frac{1}{1-\tau}\overline{F}_X(\widehat{\mathrm{VaR}}_X(\tau))=\frac{\overline{F}_X(\widehat{\mathrm{VaR}}_X(\tau))}{\overline{F}_X(\mathrm{VaR}_X(\tau))},\\ \hat{\theta}_Y(\tau):=\frac{1}{1-\tau}\overline{F}_Y(\widehat{\mathrm{VaR}}_Y(\tau))=\frac{\overline{F}_Y(\widehat{\mathrm{VaR}}_Y(\tau))}{\overline{F}_Y(\mathrm{VaR}_Y(\tau))}.\end{cases}\tag{S1.10}$$

When deriving the asymptotic theories below, we will frequently refer the asymptotic properties of $\hat{\theta}_X(\cdot)$. So we present them below as a separate lemma for completeness. Note that these asymptotic results are derived within the framework of Proposition S1. The following lemma only provides the asymptotic properties of $\hat{\theta}_X(\cdot)$. The corresponding conclusions for $\hat{\theta}_Y(\cdot)$ can be derived in an analogous manner and are therefore omitted.

Lemma S2. *Suppose that condition* (1.4) *holds. As $n\to\infty$, letting $k=O(n^\iota)$ with $\iota\in(1-\eta,1)$, we have that, for any $x>0$ such that $1-(k/n)^{2-\frac{1}{\eta}}x\in(0,1)$,*

$$\sqrt{k}\left(\frac{n}{k}\right)^{\frac{1}{2\eta}-\frac{1}{2}}\left(\hat{\theta}_X(1-(k/n)^{2-\frac{1}{\eta}}x)-1\right)x\xrightarrow{d}W(x).\tag{S1.11}$$

Furthermore, one corollary gives that,

$$\sqrt{k}\left(\frac{k}{n}\right)^{1-\frac{1}{2\eta}}\left(\hat{\theta}_X(1-(k/n)^{3-\frac{1}{\eta}})-1\right)\xrightarrow{d}W(1),\tag{S1.12}$$

in which case the range of ι is divided into two parts according to η,

$$\begin{cases} \iota \in (1-\eta, 1), & \text{if } \eta \in (1/2/2/3], \\ \iota \in \left(1 - \frac{\eta}{3\eta-1}, 1\right), & \text{if } \eta \in (2/3, 1); \end{cases}$$

the other one gives that,

$$\sqrt{k}\left(\hat{\theta}_X(1-k/n) - 1\right) \xrightarrow{d} W(1). \tag{S1.13}$$

Proof of Lemma S2. Applying (S1.8) with $s = \left(\frac{n}{k}\right)^{\frac{2}{\eta}-2} x$, we have that,

$$\sqrt{k}\left(\frac{n}{k}\right)^{\frac{1}{2\eta}-\frac{1}{2}} \left(\left(\frac{k}{n}\right)^{\frac{2}{\eta}-2} T_{n/k}((n/k)^{\frac{1}{\eta}-1}x, \infty) - x\right) \xrightarrow{d} W(x). \tag{S1.14}$$

Recalling the definition,

$$\widehat{\mathrm{VaR}}_X(1-(k/n)^{2-\frac{1}{\eta}}x) = \inf\left\{s \in (0,\infty),\ \widehat{F}_X(s) \geq 1-(k/n)^{2-\frac{1}{\eta}}x\right\},$$

which implies that,

$$\begin{aligned} \hat{\theta}_X(1-(k/n)^{2-\frac{1}{\eta}}x)x &= \frac{\overline{F}_X(\widehat{\mathrm{VaR}}_X(1-(k/n)^{2-\frac{1}{\eta}}x))}{\overline{F}_X(\mathrm{VaR}_X(1-(k/n)^{2-\frac{1}{\eta}}x))}x \\ &= \sup\left\{s \in (0,\infty),\ \left(\frac{n}{k}\right)^{2-\frac{2}{\eta}} T_{n/k}\left((k/n)^{1-\frac{1}{\eta}}s, \infty\right) < x\right\}, \end{aligned}$$

i.e., $\hat{\theta}_X(1-(k/n)^{2-\frac{1}{\eta}}x)x$ can be regarded as a generalized inverse function of $\left(\frac{n}{k}\right)^{2-\frac{2}{\eta}} T_{n/k}\left((k/n)^{1-\frac{1}{\eta}}x, \infty\right)$, which is a non-decreasing function on x. Then, using Vervaat's Lemma (see Lemma A.0.2 in de Haan and Ferreira (2006)), we have

$$\sqrt{k}\left(\frac{n}{k}\right)^{\frac{1}{2\eta}-\frac{1}{2}} \left(\hat{\theta}_X(1-(k/n)^{2-\frac{1}{\eta}}x)x - x\right) \xrightarrow{d} W(x),$$

by noting the convergence rate $\sqrt{k}\left(\frac{n}{k}\right)^{\frac{1}{2\eta}-\frac{1}{2}}$ is divergent as n increases.

Next, we check the two special conclusions ((S1.12) and (S1.13)) by choosing special values for x. On the one hand, we take $x = \frac{k}{n}$. Noting that $W(k/n)$ converges to 0 in probability at this point, we hence need to multiply both sides of (S1.11) by an appropriate rate $(n/k)^{1/2}$ in order to ensure that the limiting distribution is a trivial Gaussian process, *i.e.*,

$$\sqrt{k}\left(\frac{k}{n}\right)^{1-\frac{1}{2\eta}}\left(\hat{\theta}_X(1-(k/n)^{3-\frac{1}{\eta}})-1\right) \xrightarrow{d} W(1).$$

Besides, it is important to notice that the level $1-(k/n)^{3-\frac{1}{\eta}}$ is always intermediate when $\iota > 1-\frac{\eta}{3\eta-1}$ and so it is no need to employ any extrapolation for estimating $\mathrm{VaR}_X(1-(k/n)^{3-\frac{1}{\eta}})$ in this case. It is directly to check that, when $\eta \in (1/2, 2/3]$, it follows that $1-\eta \geq 1-\frac{\eta}{3\eta-1}$ and hence (S1.12) holds with original range $\iota \in (1-\eta, 1)$; when $\eta \in (2/3, 1)$, it follows that $1-\frac{\eta}{3\eta-1} > 1-\eta$ and then (S1.12) holds with a narrower range $\iota \in \left(1-\frac{\eta}{3\eta-1}, 1\right)$.

On the other hand, we take $x = (k/n)^{\frac{1}{\eta}-1}$. Note also that $W((k/n)^{\frac{1}{\eta}-1})$ converges to 0 in probability, we hence need to multiply both sides of (S1.11) by an appropriate rate $(k/n)^{\frac{1}{2}-\frac{1}{2\eta}}$ in order to ensure the normality, *i.e.*,

$$\sqrt{k}\left(\hat{\theta}_X(1-k/n)-1\right) \xrightarrow{d} W(1).$$

This aligns with the typical result. This proof is complete.

□

Remark 3. Note that Lemma S2 follows directly from an application of Proposition S1 (or Corollary 1) with Vervaat's Lemma by selecting $\beta = 0$.

We also provide an additional explanation for (S1.12) on its convergence rate herein. It is readily to check that the rate $\sqrt{k}\left(\frac{k}{n}\right)^{1-\frac{1}{2\eta}} \to \infty$ as $n \to \infty$ only when $\iota > 1 - \frac{\eta}{3\eta-1}$ by recalling $k = O(n^\iota)$. When $1/2 < \eta \leq 2/3$, it satisfies $1 - \eta \geq 1 - \frac{\eta}{3\eta-1}$, implying $\sqrt{k}\left(\frac{k}{n}\right)^{1-\frac{1}{2\eta}} \to \infty$ by just assuming $\iota > 1 - \eta$. On the other hand, when $2/3 < \eta < 1$, it satisfies $1 - \frac{\eta}{3\eta-1} > 1 - \eta$, implying $\sqrt{k}\left(\frac{k}{n}\right)^{1-\frac{1}{2\eta}}$ diverges to infinity only if ι is constrained to $\left(1 - \frac{\eta}{3\eta-1}, 1\right)$. This argument coincides exactly with the partitioning of ι in (S1.12). In fact, (S1.12) indeed reports the weak convergence of an intermediate-level estimation, *i.e.*, if we denote $\tau_n := 1 - (k/n)^{3-\frac{1}{\eta}}$ (an intermediate level), then a more intuitive form can be written as

$$\sqrt{n(1-\tau_n)}\left(\hat{\theta}_X(\tau_n) - 1\right) \xrightarrow{d} W(1).$$

The following lemma states that, under rate $\sqrt{k}$, the difference between $\widehat{C}_{n/k}\left((k/n)^{2-\frac{1}{\hat{\eta}}}, 1\right)$ and $T_{n/k}\left((k/n)^{2-\frac{1}{\hat{\eta}}}, 1\right)$, evaluated at the estimated argument $(k/n)^{2-\frac{1}{\hat{\eta}}}$, can not be treated as negligible.

Lemma S3. *Suppose that Assumption 1 (c) and Assumption 2 (b), (f) hold with* $\eta \in \left(\frac{7+\sqrt{17}}{16}, 1\right)$. *As* $n \to \infty$, *letting* $k = O(n^\iota)$ *with* $\iota \in \left(\frac{2}{3}, 1 - \frac{1}{1-2\alpha}\right)$, *we have that,*

$$\left|\sqrt{k}\left(\widehat{C}_{n/k}\left((k/n)^{2-\frac{1}{\hat{\eta}}},1\right)-T_{n/k}\left((k/n)^{2-\frac{1}{\hat{\eta}}},1\right)\right)-C_2\left((k/n)^{2-\frac{1}{\eta}},1\right)W(1)\right|=o_{\mathbb{P}}(1).$$

Proof of Lemma S3. We first rewrite $\widehat{C}_{n/k}\left((k/n)^{2-\frac{1}{\hat{\eta}}},1\right)$ as,

$$\begin{aligned}
&\widehat{C}_{n/k}\left((k/n)^{2-\frac{1}{\hat{\eta}}},1\right)\\
&=\left(\frac{n}{k}\right)^{\frac{1}{\hat{\eta}}}\frac{1}{n}\sum_{i=1}^{n}I\left(1-\widehat{F}_X(X_i)\leq\left(\frac{k}{n}\right)^{3-\frac{1}{\hat{\eta}}},1-\widehat{F}_Y(Y_i)\leq\frac{k}{n}\right)\\
&=\left(\frac{n}{k}\right)^{\frac{1}{\hat{\eta}}}\frac{1}{n}\sum_{i=1}^{n}I\left(\widehat{F}_X(X_i)\geq1-\left(\frac{k}{n}\right)^{3-\frac{1}{\hat{\eta}}},\widehat{F}_Y(Y_i)\geq1-\frac{k}{n}\right)\\
&=\left(\frac{n}{k}\right)^{\frac{1}{\hat{\eta}}}\frac{1}{n}\sum_{i=1}^{n}I\left(X_i\geq\widehat{\mathrm{VaR}}_X(1-(k/n)^{3-\frac{1}{\hat{\eta}}}),Y_i\geq\widehat{\mathrm{VaR}}_Y(1-k/n)\right)\\
&=\left(\frac{n}{k}\right)^{\frac{1}{\hat{\eta}}}\frac{1}{n}\sum_{i=1}^{n}I\left(\overline{F}_X(X_i)\leq\frac{\overline{F}_X(\widehat{\mathrm{VaR}}_X(1-(k/n)^{3-\frac{1}{\hat{\eta}}}))}{\overline{F}_X(\mathrm{VaR}_X(1-(k/n)^{3-\frac{1}{\hat{\eta}}}))}\left(\frac{k}{n}\right)^{3-\frac{1}{\hat{\eta}}},\right.\\
&\qquad\left.\overline{F}_Y(Y_i)\leq\frac{\overline{F}_Y(\widehat{\mathrm{VaR}}_Y(1-k/n))}{\overline{F}_Y(\mathrm{VaR}_Y(1-k/n))}\frac{k}{n}\right)\\
&=\left(\frac{n}{k}\right)^{\frac{1}{\hat{\eta}}-\frac{1}{\eta}}T_{n/k}\left(\left(\frac{k}{n}\right)^{2-\frac{1}{\hat{\eta}}}\frac{\overline{F}_X(\widehat{\mathrm{VaR}}_X(1-(k/n)^{3-\frac{1}{\hat{\eta}}}))}{\overline{F}_X(\mathrm{VaR}_X(1-(k/n)^{3-\frac{1}{\hat{\eta}}}))},\frac{\overline{F}_Y(\widehat{\mathrm{VaR}}_Y(1-k/n))}{\overline{F}_Y(\mathrm{VaR}_Y(1-k/n))}\right)\\
&\qquad+o(1)\\
&=\left(\frac{n}{k}\right)^{\frac{1}{\hat{\eta}}-\frac{1}{\eta}}T_{n/k}\left((k/n)^{2-\frac{1}{\hat{\eta}}}\hat{\theta}_X(1-(k/n)^{3-\frac{1}{\hat{\eta}}}),\hat{\theta}_Y(1-k/n)\right)+o(1),
\end{aligned}$$

where $\hat{\theta}_X(1-(k/n)^{3-\frac{1}{\hat{\eta}}})$ and $\hat{\theta}_Y(1-k/n)$ are defined in (S1.10). It is also worthy noting that $\left(\frac{n}{k}\right)^{\frac{1}{\hat{\eta}}-\frac{1}{\eta}}\xrightarrow{\mathbb{P}}1$ by using the consistency of $\hat{\eta}$ (see (2.19)).

Then, it follows the following decomposition,

$$
\begin{aligned}
&\widehat{C}_{n/k}\left((k/n)^{2-\frac{1}{\hat{\eta}}},1\right)-T_{n/k}\left((k/n)^{2-\frac{1}{\hat{\eta}}},1\right)\\
=&\left\{T_{n/k}\left((k/n)^{2-\frac{1}{\hat{\eta}}}\,\hat{\theta}_X(1-(k/n)^{3-\frac{1}{\hat{\eta}}}),\hat{\theta}_Y(1-k/n)\right)\right.\\
&\left.-C_{n/k}\left((k/n)^{2-\frac{1}{\hat{\eta}}}\,\hat{\theta}_X(1-(k/n)^{3-\frac{1}{\hat{\eta}}}),\hat{\theta}_Y(1-k/n)\right)\right\}\\
&+\left\{C_{n/k}\left((k/n)^{2-\frac{1}{\hat{\eta}}}\,\hat{\theta}_X(1-(k/n)^{3-\frac{1}{\hat{\eta}}}),\hat{\theta}_Y(1-k/n)\right)\right.\\
&\left.-C\left((k/n)^{2-\frac{1}{\hat{\eta}}}\,\hat{\theta}_X(1-(k/n)^{3-\frac{1}{\hat{\eta}}}),\hat{\theta}_Y(1-k/n)\right)\right\}\\
&+C\left((k/n)^{2-\frac{1}{\hat{\eta}}}\,\hat{\theta}_X(1-(k/n)^{3-\frac{1}{\hat{\eta}}}),\hat{\theta}_Y(1-k/n)\right)-C\left((k/n)^{2-\frac{1}{\hat{\eta}}},1\right)\\
&+C\left((k/n)^{2-\frac{1}{\hat{\eta}}},1\right)-C_{n/k}\left((k/n)^{2-\frac{1}{\hat{\eta}}},1\right)\\
&+C_{n/k}\left((k/n)^{2-\frac{1}{\hat{\eta}}},1\right)-T_{n/k}\left((k/n)^{2-\frac{1}{\hat{\eta}}},1\right)\\
&+o_{\mathbb{P}}(1)\\
=:&I_1+I_2+I_3+I_4+I_5+o_{\mathbb{P}}(1).
\end{aligned}
$$

We aim to apply Proposition S1 and Lemma S1 to deal with I_1 and I_5 first. For convenience, we simplify $\hat{\theta}_X(1-(k/n)^{3-\frac{1}{\hat{\eta}}})$ and $\hat{\theta}_Y(1-k/n)$ to $\hat{\theta}_X$

and $\hat{\theta}_Y$. Then, it follows that,

$$\begin{aligned}
&\left|\sqrt{k} I_1\right| \\
=&\left|\sqrt{k}\left(\frac{n}{k}\right)^{\frac{1}{2}-\frac{1}{2\eta}}\left(T_{n/k}\left((k/n)^{2-\frac{1}{\hat{\eta}}}\hat{\theta}_X, \hat{\theta}_Y\right) - C_{n/k}\left((k/n)^{2-\frac{1}{\hat{\eta}}}\hat{\theta}_X, \hat{\theta}_Y\right)\right)\right|\left(\frac{k}{n}\right)^{\frac{1}{2}-\frac{1}{2\eta}} \\
\leq&\left|\frac{\sqrt{k}\left(\frac{n}{k}\right)^{\frac{1}{2}-\frac{1}{2\eta}}\left(T_{n/k}\left(\left(\frac{k}{n}\right)^{2-\frac{1}{\hat{\eta}}}\hat{\theta}_X, \hat{\theta}_Y\right) - C_{n/k}\left(\left(\frac{k}{n}\right)^{2-\frac{1}{\hat{\eta}}}\hat{\theta}_X, \hat{\theta}_Y\right)\right) - W_C\left(\left(\frac{k}{n}\right)^{2-\frac{1}{\hat{\eta}}}\hat{\theta}_X, \hat{\theta}_Y\right)}{\left((k/n)^{2-\frac{1}{\hat{\eta}}}\hat{\theta}_X\right)^{\beta}}\right| \\
&\times\left(\frac{k}{n}\right)^{\frac{1}{2}-\frac{1}{2\eta}+\beta\left(2-\frac{1}{\hat{\eta}}\right)} \\
&+\left|\frac{W_C\left((k/n)^{2-\frac{1}{\hat{\eta}}}\hat{\theta}_X, \hat{\theta}_Y\right)}{\left((k/n)^{2-\frac{1}{\hat{\eta}}}\hat{\theta}_X\right)^{\beta}}\right|\left(\frac{k}{n}\right)^{\frac{1}{2}-\frac{1}{2\eta}+\beta\left(2-\frac{1}{\hat{\eta}}\right)} \\
=&\left(I_{11}+I_{12}\right)(k/n)^{\frac{1}{2}-\frac{1}{2\eta}+\beta\left(2-\frac{1}{\hat{\eta}}\right)}.
\end{aligned}$$

To apply Proposition S1, the parameter ι must be constrained to the interval $\left(\frac{1-\eta}{1-2\beta\eta}, 1\right)$. In other words, this requires the inclusion $\left(\frac{2}{3}, 1-\frac{1}{1-2\alpha}\right) \subset \left(\frac{1-\eta}{1-2\beta\eta}, 1\right)$ to hold. Since the upper bound is automatically satisfied, this condition reduces to the inequality $\frac{2}{3} > \frac{1-\eta}{1-2\beta\eta}$ on the lower bound, which is equivalent to $\beta < \frac{1}{4}\left(3-\frac{1}{\eta}\right)$. On the other hand, if we further choose a β such that $\frac{1}{4}\left(\frac{1}{2\eta-1}-1\right) < \beta$, then it follows that $\left(\frac{k}{n}\right)^{\left(2-\frac{1}{\hat{\eta}}\right)\beta+\left(\frac{1}{2}-\frac{1}{2\eta}\right)} \to 0$ as $n \to \infty$ under Assumption 2 (f), because the exponent $\left(2-\frac{1}{\hat{\eta}}\right)\beta + \left(\frac{1}{2}-\frac{1}{2\eta}\right) > 0$ when n becomes large sufficiently. Moreover, by solving the order

$$0 < \frac{1}{4}\left(\frac{1}{2\eta-1}-1\right) < \frac{1}{4}\left(3-\frac{1}{\eta}\right) < \frac{1}{2},$$

we obtain the range of η, *i.e.*, $\eta \in \left(\frac{7+\sqrt{17}}{16}, 1\right)$. Thus, this indicates that there always exists a $\beta \in (0, 1/2)$ such that

$$\frac{1}{4}\left(\frac{1}{2\eta-1}-1\right) < \beta < \frac{1}{4}\left(3-\frac{1}{\eta}\right).$$

By choosing such β, applying Proposition S1 and Lemma S1, we find that $I_{11} = o_{\mathbb{P}}(1)$, $I_{12} = O_{\mathbb{P}}(1)$ and $\left(\frac{k}{n}\right)^{\left(2-\frac{1}{\hat{\eta}}\right)\beta+\left(\frac{1}{2}-\frac{1}{2\eta}\right)} = o(1)$. This implies that $\sqrt{k}I_1 = o_{\mathbb{P}}(1)$. A similar result for I_5 gives $\sqrt{k}I_5 = o_{\mathbb{P}}(1)$.

Second, we work on I_2 and I_4, using Assumption 2 (b), we have that

$$\begin{aligned}
\sqrt{k}I_2 &= \sqrt{k}I_2\left(\left(\frac{k}{n}\right)^{2-\frac{1}{\hat{\eta}}}\hat{\theta}_X\right)^{-\beta_1}\left(\left(\frac{k}{n}\right)^{2-\frac{1}{\hat{\eta}}}\hat{\theta}_X\right)^{\beta_1} \\
&= O\left(\sqrt{k}\left(\frac{n}{k}\right)^{\alpha}\left(\frac{k}{n}\right)^{\left(2-\frac{1}{\hat{\eta}}\right)\beta_1}\right) \\
&= O\left(n^{\frac{\iota}{2}+(\iota-1)\left[\left(2-\frac{1}{\hat{\eta}}\right)\beta_1-\alpha\right]}\right) \\
&= O\left(n^{\left(2-\frac{1}{\hat{\eta}}\right)(\iota-1)\left[\beta_1-\frac{\iota/[2(1-\iota)]+\alpha}{2-1/\hat{\eta}}\right]}\right).
\end{aligned}$$

Note that $\iota < 1 - \frac{1}{1-2\alpha}$ implies $\alpha + \frac{\iota}{2(1-\iota)} < 0$ and hence $\left(2-\frac{1}{\hat{\eta}}\right)(\iota - 1)\left[\beta_1 - \frac{\iota/[2(1-\iota)]+\alpha}{2-1/\hat{\eta}}\right] < 0$. This indicates $\sqrt{k}I_2 = o(1)$ as $n \to \infty$. A similar result for I_4 gives $\sqrt{k}I_4 = o(1)$.

Finally, I_3, we apply intermediate value theorem,

$$\begin{aligned}I_3 &= C\left((k/n)^{2-\frac{1}{\hat\eta}}\,\hat\theta_X, \hat\theta_Y\right) - C\left((k/n)^{2-\frac{1}{\hat\eta}}, \hat\theta_Y\right)\\ &\quad + C\left((k/n)^{2-\frac{1}{\hat\eta}}, \hat\theta_Y\right) - C\left((k/n)^{2-\frac{1}{\hat\eta}}, 1\right)\\ &= C_1\left((k/n)^{2-\frac{1}{\hat\eta}}\,\zeta_1, \hat\theta_Y\right)(k/n)^{2-\frac{1}{\hat\eta}}\left(\hat\theta_X - 1\right)\\ &\quad + C_2\left((k/n)^{2-\frac{1}{\hat\eta}}, \zeta_2\right)\left(\hat\theta_Y - 1\right),\end{aligned}$$

where $\zeta_1 \in \left(\hat\theta_X, 1\right)$ tending to 1 and $\zeta_2 \in \left(\hat\theta_Y, 1\right)$ tending to 1 as well. Then, we have that,

$$\begin{aligned}\sqrt{k}I_3 &= C_1\left((k/n)^{2-\frac{1}{\hat\eta}}\,\zeta_1, \hat\theta_Y\right)(k/n)^{1-\frac{1}{2\hat\eta}}\left[\sqrt{k}\left(\frac{k}{n}\right)^{1-\frac{1}{2\hat\eta}}\left(\hat\theta_X - 1\right)\right]\\ &\quad + C_2\left((k/n)^{2-\frac{1}{\hat\eta}}, \zeta_2\right)\sqrt{k}\left(\hat\theta_Y - 1\right).\end{aligned}$$

It is directly to check that $\eta \in \left(\frac{7+\sqrt{17}}{16}, 1\right) \subset \left(\frac{2}{3}, 1\right)$ and $\iota \in \left(\frac{2}{3}, 1 - \frac{1}{1-2\alpha}\right) \subset \left(1 - \frac{\eta}{3\eta-1}, 1\right)$ by noting $\frac{2}{3} > 1 - \frac{\eta}{3\eta-1}$. It thus implies that, by using (S1.12) of Lemma S2 and the consistency of $\hat\eta$ given in (2.19) under Assumption 2 (f),

$$\left|\sqrt{k}I_3 - C_2\left((k/n)^{2-\frac{1}{\eta}}, 1\right)W(1)\right| = o_{\mathbb{P}}(1).$$

Therefore, combining the above arguments for I_1, I_2, I_3, I_4, I_4 yields that,

$$\left|\sqrt{k}\left(\widehat{C}_{n/k}\left((k/n)^{2-\frac{1}{\hat\eta}}, 1\right) - T_{n/k}\left((k/n)^{2-\frac{1}{\hat\eta}}, 1\right)\right) - C_2\left((k/n)^{2-\frac{1}{\eta}}, 1\right)W(1)\right| = o_{\mathbb{P}}(1).$$

This proof is therefore complete. □

S2 Technical proofs

In this section, we provide the full theoretical details underlying the conclusions presented in the main article.

S2.1 Proofs of Lemma 1 and Proposition 1

Proof of Lemma 1. This proof is an analogue to that of Lemmas S1.1 and S1.2 of Nolde et al. (2022). We first show that $\xi_\tau^* \to 0$ as $\tau \uparrow 1$. If otherwise, there exists a constant $d > 0$ such that $\xi_\tau^* \to d$ as $\tau \uparrow 1$. Then, we have $C(\xi_\tau^*, 1) \to C(d, 1) > 0$ by using the continuity of $C(x, 1)$ on x. However, this contradicts with $C(\xi_\tau^*, 1) = (1-\tau)^{2-\frac{1}{\eta}} \to 0$ as $\tau \uparrow 1$. Hence, we conclude that $\xi_\tau^* \to 0$ as $\tau \uparrow 1$. Moreover, using the intermediate value theorem, there exists a constant $\tilde{\xi}_\tau \in [0, \xi_\tau^*]$, tending to 0, such that,

$$(1-\tau)^{2-\frac{1}{\eta}} = C(\xi_\tau^*, 1) = C(0, 1) + \xi_\tau^* C_1(\tilde{\xi}_\tau, 1).$$

Hence, we get that, as $\tau \uparrow 1$,

$$\frac{(1-\tau)^{2-\frac{1}{\eta}}}{\xi_\tau^*} = C_1(\tilde{\xi}_\tau, 1) \to C_1(0, 1),$$

which follows from the continuity of $x \mapsto C_1(x, 1)$ assumed by Assumption 1 (c).

Below, we prove the second statement, *i.e.*, $\xi_\tau/\xi_\tau^* \to 1$. We begin with showing that $\xi_\tau \to 0$ as $\tau \uparrow 1$. If otherwise, there exists a constant $d' > 0$

such that $\xi_\tau \to d'$ as $\tau \uparrow 1$. Then, by taking $\tau \uparrow 1$ on both sides of (2.9), we obtain that $C(d', 1) = 0$, which contradicts with the fact $C(x, y) > 0$ with $(x, y) \in (0, \infty)^2$. Hence, we conclude that $\xi_\tau \to 0$ as $\tau \uparrow 1$. Next, we show that

$$\limsup_{\tau \uparrow 1} \frac{\xi_\tau}{\xi_\tau^*} \leq 1,$$

by contradiction. If assuming otherwise, there exists a constant $d'' > 1$ such that, as $\tau \uparrow 1$,

$$\frac{\xi_\tau}{\xi_\tau^*} \to d'' > 1.$$

Therefore, for any $1 < \tilde{d}'' < d''$, there exists a $\tau_0 \in (0, 1)$ such that for $\tau > \tau_0$, $\xi_\tau / \xi_\tau^* > \tilde{d}''$, and hence $\xi_\tau > \tilde{d}'' \xi_\tau^* > \xi_\tau^*$. Using intermediate value theorem again, there exists $\bar{\xi}_\tau \in (\xi_\tau^*, \xi_\tau)$ such that

$$C(\xi_\tau, 1) - C(\xi_\tau^*, 1) = C_1(\bar{\xi}_\tau, 1)(\xi_\tau - \xi_\tau^*).$$

Note that $\bar{\xi}_\tau \to 0$ since both ξ_τ and ξ_τ^* tend to 0 as $\tau \uparrow 1$. Note also that $\xi_\tau - \xi_\tau^* > (\tilde{d}'' - 1)\xi_\tau^*$. Then, applying the continuity of $C_1(x, y)$ at $(0, 1)$, it follows that

$$\begin{aligned}
\liminf_{\tau \uparrow 1} \frac{C(\xi_\tau, 1) - (1-\tau)^{2-\frac{1}{\eta}}}{(1-\tau)^{2-\frac{1}{\eta}}} &= \liminf_{\tau \uparrow 1} \frac{C(\xi_\tau, 1) - C(\xi_\tau^*, 1)}{(1-\tau)^{2-\frac{1}{\eta}}} \\
&= \liminf_{\tau \uparrow 1} \frac{C(\xi_\tau, 1) - C(\xi_\tau^*, 1)}{\xi_\tau^*} \times \frac{\xi_\tau^*}{(1-\tau)^{2-\frac{1}{\eta}}} \\
&= \liminf_{\tau \uparrow 1} \frac{C_1(\bar{\xi}_\tau, 1)(\xi_\tau - \xi_\tau^*)}{\xi_\tau^*} \times \frac{\xi_\tau^*}{(1-\tau)^{2-\frac{1}{\eta}}}
\end{aligned}$$

$$\geq C_1(0,1)(\tilde{d}''-1)\times\frac{1}{C_1(0,1)}$$
$$=\tilde{d}''-1$$
$$>0.$$

One the other hand, using Assumption 2 (b), we get,

$$\begin{aligned}
&\lim_{\tau\uparrow 1}\frac{C(\xi_\tau,1)-(1-\tau)^{2-\frac{1}{\eta}}}{(1-\tau)^{2-\frac{1}{\eta}}}\\
=&\lim_{\tau\uparrow 1}(1-\tau)^{\frac{1}{\eta}-2}\left(C(\xi_\tau,1)-(1-\tau)^{-\frac{1}{\eta}}\,\mathbb{P}(\overline{F}_X(X)\leq(1-\tau)\xi_\tau,\overline{F}_F(Y)\leq 1-\tau)\right)\\
=&\lim_{\tau\uparrow 1}(1-\tau)^{\frac{1}{\eta}-2}\left(C(\xi_\tau,1)-(1-\tau)^{-\frac{1}{\eta}}\,\mathbb{P}(\overline{F}_X(X)\leq(1-\tau)\xi_\tau,\overline{F}_F(Y)\leq 1-\tau)\right)\xi_\tau^{-\beta_1}\xi_\tau^{\beta_1}\\
\leq&\lim_{\tau\uparrow 1}(1-\tau)^{\frac{1}{\eta}-2}\left(C(\xi_\tau,1)-(1-\tau)^{-\frac{1}{\eta}}\,\mathbb{P}(\overline{F}_X(X)\leq(1-\tau)\xi_\tau,\overline{F}_F(Y)\leq 1-\tau)\right)\xi_\tau^{-\beta_1}\\
=&\lim_{\tau\uparrow 1}O\left((1-\tau)^{\frac{1}{\eta}-2-\alpha}\right)\\
=&0,
\end{aligned}$$

by recalling $\frac{1}{\eta}-2>-1>\alpha$. Therefore, the above two limit relations contradict each other. Therefore, we conclude that $\limsup_{\tau\uparrow 1}\frac{\xi_\tau}{\xi_\tau^*}\leq 1$. One can show a lower bound for $\frac{\xi_\tau}{\xi_\tau^*}$ via imposing a similar argument, *i.e.*, $\liminf_{\tau\uparrow 1}\frac{\xi_\tau}{\xi_\tau^*}\geq 1$. This proof is complete. □

Proof of Proposition 1. Note first that the quantity $XI(X\geq \mathrm{CoVaR}_{X|Y}(\tau),Y\geq \mathrm{VaR}_Y(\tau))$ is always positive as τ approaches to 1. It is because $\mathrm{CoVaR}_{X|Y}(\tau)\to\infty$ as $\tau\uparrow 1$ due to the heavy-tailed assumption. Then, using Fubini's theorem and change of variable, we have,

$$
\begin{aligned}
&\frac{\text{CoES}_{X|Y}(\tau)}{\text{CoVaR}_{X|Y}(\tau)}\\
=&\frac{E[X|X \geq \text{CoVaR}_{X|Y}(\tau), Y \geq \text{VaR}_Y(\tau)]}{\text{CoVaR}_{X|Y}(\tau)}\\
=&\frac{E[XI(X \geq \text{CoVaR}_{X|Y}(\tau), Y \geq \text{VaR}_Y(\tau))]}{(1-\tau)^2\text{CoVaR}_{X|Y}(\tau)}\\
=&\frac{1}{(1-\tau)^2}\int_0^\infty \frac{\mathbb{P}(X \geq x \vee \text{CoVaR}_{X|Y}(\tau), Y \geq \text{VaR}_Y(\tau))}{\text{CoVaR}_{X|Y}(\tau)}\,dx\\
=&\frac{1}{(1-\tau)^2}\int_0^{\text{CoVaR}_{X|Y}(\tau)} \frac{\mathbb{P}(X \geq \text{CoVaR}_{X|Y}(\tau), Y \geq \text{VaR}_Y(\tau))}{\text{CoVaR}_{X|Y}(\tau)}\,dx\\
&+\frac{1}{(1-\tau)^2}\int_{\text{CoVaR}_{X|Y}(\tau)}^\infty \frac{\mathbb{P}(X \geq x, Y \geq \text{VaR}_Y(\tau))}{\text{CoVaR}_{X|Y}(\tau)}\,dx\\
=&1+\frac{1}{(1-\tau)^2}\int_1^\infty \mathbb{P}(X \geq x\text{CoVaR}_{X|Y}(\tau), Y \geq \text{VaR}_Y(\tau))\,dx\\
=&1+\int_1^\infty \frac{\mathbb{P}(X \geq x\text{CoVaR}_{X|Y}(\tau), Y \geq \text{VaR}_Y(\tau))}{\mathbb{P}(X \geq \text{CoVaR}_{X|Y}(\tau), Y \geq \text{VaR}_Y(\tau))}\,dx\\
=&1+\int_1^\infty \frac{\mathbb{P}(\overline{F}_X(X) \leq \overline{F}_X(x\text{VaR}_X(1-(1-\tau)\xi_\tau)), \overline{F}_Y(Y) \leq 1-\tau)}{\mathbb{P}(\overline{F}_X(X) \leq (1-\tau)\xi_\tau, \overline{F}_Y(Y) \leq 1-\tau)}\,dx,
\end{aligned}
$$

where the last step follows identity (2.7). Using (1.4), regular variation (2.5) twice and L'Hopital rule, it follows that

$$
\begin{aligned}
&\lim_{\tau\uparrow 1}\frac{\mathbb{P}(\overline{F}_X(X) \leq \overline{F}_X(x\text{CoVaR}_{X|Y}(\tau)), \overline{F}_Y(Y) \leq 1-\tau)}{\mathbb{P}(\overline{F}_X(X) \leq (1-\tau)\xi_\tau, \overline{F}_Y(Y) \leq 1-\tau)}\\
=&\lim_{\tau\uparrow 1}\frac{(1-\tau)^{-\frac{1}{\eta}}\mathbb{P}(\overline{F}_X(X) \leq \overline{F}_X(x\text{CoVaR}_{X|Y}(\tau)), \overline{F}_Y(Y) \leq 1-\tau)}{(1-\tau)^{-\frac{1}{\eta}}\mathbb{P}(\overline{F}_X(X) \leq (1-\tau)\xi_\tau, \overline{F}_Y(Y) \leq 1-\tau)}\\
=&x^{-1/\gamma_1},
\end{aligned}
$$

by noting that, as $\tau \uparrow 1$,

$$\overline{F}_X(x\text{CoVaR}_{X|Y}(\tau)) = \frac{\overline{F}_X(x\text{CoVaR}_{X|Y}(\tau))}{\overline{F}_X(\text{CoVaR}_{X|Y}(\tau))}(1-\tau)\xi_\tau \sim x^{-\frac{1}{\gamma_1}}(1-\tau)\xi_\tau.$$

This also implies that, for a $0 < \varepsilon < 1/\gamma_1 - 1$, there exists a $\tau(\varepsilon)$ such that, for all $\tau > \tau(\varepsilon)$ and $x > 1$,

$$\frac{\mathbb{P}(\overline{F}_X(X) \leq \overline{F}_X(x\text{CoVaR}_{X|Y}(\tau)), \overline{F}_Y(Y) \leq 1-\tau)}{\mathbb{P}(\overline{F}_X(X) \leq (1-\tau)\eta_\tau, \overline{F}_Y(Y) \leq 1-\tau)} \leq x^{-1/\gamma_1+\varepsilon},$$

where the bound $x^{-1/\gamma_1+\varepsilon}$ is integrable since $\int_1^\infty x^{-1/\gamma_1+\varepsilon}\,dx < \infty$. Then, using dominated convergence theorem, we have

$$\begin{aligned}
&\lim_{\tau\uparrow 1} \frac{\text{CoES}_{X|Y}(\tau)}{\text{CoVaR}_{X|Y}(\tau)} \\
=&1 + \lim_{\tau\uparrow 1} \int_1^\infty \frac{\mathbb{P}(\overline{F}_X(X) \leq \overline{F}_X(x\text{CoVaR}_{X|Y}(\tau)), \overline{F}_Y(Y) \leq 1-\tau)}{\mathbb{P}(\overline{F}_X(X) \leq (1-\tau)\eta_\tau, \overline{F}_Y(Y) \leq 1-\tau)}\,dx \\
=&1 + \int_1^\infty x^{-1/\gamma_1}\,dx \\
=&\frac{1}{1-\gamma_1}.
\end{aligned}$$

□

S2.2 Proofs of Proposition 2 and Theorem 1

Proof of Proposition 2. We follow the decomposition,

$$\begin{aligned}
&\widehat{C}_{n/k}\left((k/n)^{2-\frac{1}{\eta}}, 1\right) - C\left((k/n)^{2-\frac{1}{\eta}}, 1\right) \\
=&\widehat{C}_{n/k}\left((k/n)^{2-\frac{1}{\eta}}, 1\right) - T_{n/k}\left((k/n)^{2-\frac{1}{\eta}}, 1\right)
\end{aligned}$$

$$
\begin{aligned}
&+ T_{n/k}\left((k/n)^{2-\frac{1}{\hat\eta}},1\right) - C_{n/k}\left((k/n)^{2-\frac{1}{\hat\eta}},1\right) \\
&+ C_{n/k}\left((k/n)^{2-\frac{1}{\hat\eta}},1\right) - C\left((k/n)^{2-\frac{1}{\hat\eta}},1\right) \\
&+ C\left((k/n)^{2-\frac{1}{\hat\eta}},1\right) - C\left((k/n)^{2-\frac{1}{\eta}},1\right) \\
=:& I_1 + I_2 + I_3 + I_4.
\end{aligned}
$$

Note that $\sqrt{k}(\hat\eta - \eta) = O_{\mathbb{P}}(1)$, indicating that, for any $\varepsilon > 0$, there exists a N such that $\eta - \varepsilon < \hat\eta < \eta + \varepsilon$ and hence $\hat\eta \in (1/2, 1)$ for all $n > N$. Thus, the quantity $(k/n)^{2-\frac{1}{\hat\eta}}$ tends to 0 as $n \to \infty$.

The asymptotic result for I_1 has been reported in Lemma S3. Moreover, once $\hat\theta_X$ and $\hat\theta_Y$ are replaced with 1, terms I_2 and I_3 have also already been studied in the proof of Lemma S3, *i.e.*, $\sqrt{k}I_2 = o_{\mathbb{P}}(1)$ and $\sqrt{k}I_3 = o(1)$. For I_4, we have, using intermediate value theorem,

$$
\sqrt{k}I_4 = C_1\left((k/n)^{2-\frac{1}{\tilde\eta}},1\right)\left(\frac{k}{n}\right)^{2-\frac{1}{\tilde\eta}} \log\frac{k}{n}\sqrt{k}\left(\frac{1}{\hat\eta} - \frac{1}{\eta}\right) = o_{\mathbb{P}}(1),
$$

by noting $\left(\frac{k}{n}\right)^{2-\frac{1}{\tilde\eta}} \log\frac{k}{n} \to 0$ as $n \to \infty$, where $\tilde\eta$ is between $\hat\eta$ and η and tends to η as $n \to \infty$.

Then, it follows that

$$
\left|\sqrt{k}\left(\widehat{C}_{n/k}\left((k/n)^{2-\frac{1}{\hat\eta}},1\right) - C\left((k/n)^{2-\frac{1}{\eta}},1\right)\right) - C_2\left((k/n)^{2-\frac{1}{\eta}},1\right)W(1)\right| = o_{\mathbb{P}}(1).
$$

Notice that both $\widehat{C}_{n/k}\left(x^{2-\frac{1}{\hat\eta}},1\right)$ and $C\left(x^{2-\frac{1}{\eta}},1\right)$ are non-decreasing functions on $x \in (0,1)$, and moreover, the derivative $\frac{\partial}{\partial x}C\left(x^{2-\frac{1}{\eta}},1\right) =$

$C_1\left(x^{2-\frac{1}{\eta}},1\right)\left(2-\frac{1}{\eta}\right)x^{1-\frac{1}{\eta}}$ is positive on $(0,1)$. Now, a direct application of Vervaat's Lemma (see Lemma A.0.2 in de Haan and Ferreira (2006)) yields that,

$$\sqrt{k}\left(\frac{k}{n}\right)^{\frac{1}{\eta}-1}\left(\hat{\xi}_{1-k/n}-\xi_{1-k/n}\right)\xrightarrow{d}\left(\frac{1}{\eta}-2\right)\frac{C_2(0,1)}{C_1(0,1)}W(1),$$

by noting that $\xi^*_{1-k/n}/\xi_{1-k/n}\to 1$ and

$$\frac{\partial \xi^*_{1-x}}{\partial x}=\frac{1}{C_1(\xi^*_{1-x},1)}\left(2-\frac{1}{\eta}\right)x^{1-\frac{1}{\eta}}.$$

The final result follows from the convergence $\frac{\xi_{1-k/n}}{(k/n)^{2-\frac{1}{\eta}}}\to 1/C_1(0,1)$ as $n\to\infty$ established in Lemma 1. □

Proof of Theorem 1. Note that

$$\begin{aligned}
&\left(\frac{\widetilde{\mathrm{CoVaR}}^{(1)}_{X|Y}(\tau_n')}{\mathrm{CoVaR}_{X|Y}(\tau_n')}-1,\frac{\widetilde{\mathrm{CoES}}^{(1)}_{X|Y}(\tau_n')}{\mathrm{CoES}_{X|Y}(\tau_n')}-1\right)^\top\\
=&\left(\frac{\widetilde{\mathrm{CoVaR}}^{(1)}_{X|Y}(\tau_n')}{\mathrm{CoVaR}_{X|Y}(\tau_n')}-1,\frac{1}{1-\hat{\gamma}_1}\frac{\mathrm{CoVaR}_{X|Y}(\tau_n')}{\mathrm{CoES}_{X|Y}(\tau_n')}\left(\frac{\widetilde{\mathrm{CoVaR}}^{(1)}_{X|Y}(\tau_n')}{\mathrm{CoVaR}_{X|Y}(\tau_n')}-1\right)\right.\\
&\left.+\frac{1}{1-\hat{\gamma}_1}\frac{\mathrm{CoVaR}_{X|Y}(\tau_n')}{\mathrm{CoES}_{X|Y}(\tau_n')}-1\right)^\top\\
=&\left(1,\frac{1}{1-\hat{\gamma}_1}\frac{\mathrm{CoVaR}_{X|Y}(\tau_n')}{\mathrm{CoES}_{X|Y}(\tau_n')}\right)^\top\left(\frac{\widetilde{\mathrm{CoVaR}}^{(1)}_{X|Y}(\tau_n')}{\mathrm{CoVaR}_{X|Y}(\tau_n')}-1\right)\\
&+\left(0,\frac{1}{1-\hat{\gamma}_1}\frac{\mathrm{CoVaR}_{X|Y}(\tau_n')}{\mathrm{CoES}_{X|Y}(\tau_n')}-1\right)^\top.
\end{aligned}\qquad\text{(S2.15)}$$

First, using (2.19) (under Assumption 2 (f)) on Hill estimator $\hat{\gamma}_1$, it follows that,

$$\sqrt{k}\left(\frac{1}{1-\hat{\gamma}_1}-\frac{1}{1-\gamma_1}\right)=O_{\mathbb{P}}(1), \tag{S2.16}$$

and hence

$$\begin{aligned}\frac{1}{1-\hat{\gamma}_1}\frac{\mathrm{CoVaR}_{X|Y}(\tau_n')}{\mathrm{CoES}_{X|Y}(\tau_n')}&=\left(\frac{1}{1-\gamma_1}+O_{\mathbb{P}}\left(\frac{1}{\sqrt{k}}\right)\right)((1-\gamma_1)+o(1))\\&=1+O_{\mathbb{P}}\left(\frac{1}{\sqrt{k}}\right).\end{aligned} \tag{S2.17}$$

Therefore, it remains to study the limit distributions of $\frac{\widetilde{\mathrm{CoVaR}}^{(1)}_{X|Y}(\tau_n')}{\mathrm{CoVaR}_{X|Y}(\tau_n')}$, which gives

$$\begin{aligned}&\frac{\widetilde{\mathrm{CoVaR}}^{(1)}_{X|Y}(\tau_n')}{\mathrm{CoVaR}_{X|Y}(\tau_n')}\\
&=\left(\frac{k}{n(1-\tau_n')}\right)^{\hat{\gamma}_1\left(3-\frac{1}{\hat{\eta}}\right)}\hat{\xi}_{1-k/n}^{-\hat{\gamma}_1}\frac{\widehat{\mathrm{VaR}}_X(1-k/n)}{\mathrm{VaR}_X(1-k/n)}\frac{\mathrm{VaR}_X(1-k/n)}{\mathrm{VaR}_X(1-(1-\tau_n')\xi_{\tau_n'})}\\
&=\left(\frac{k}{n(1-\tau_n')}\right)^{\hat{\gamma}_1\left(3-\frac{1}{\hat{\eta}}\right)}\hat{\xi}_{1-k/n}^{-\hat{\gamma}_1}\frac{\widehat{\mathrm{VaR}}_X(1-k/n)}{\mathrm{VaR}_X(1-k/n)}\left(\frac{n(1-\tau_n')}{k}\xi_{\tau_n'}\right)^{\gamma_1}\\
&\quad\times\left(\frac{n(1-\tau_n')}{k}\xi_{\tau_n'}\right)^{-\gamma_1}\frac{U_1(n/k)}{U_1\left(1/((1-\tau_n')\xi_{\tau_n'})\right)}\\
&=\left(\frac{k}{n(1-\tau_n')}\right)^{\hat{\gamma}_1\left(3-\frac{1}{\hat{\eta}}\right)-\gamma_1\left(3-\frac{1}{\eta}\right)}\left(\frac{\hat{\xi}_{1-k/n}}{\xi_{1-k/n}}\right)^{-\hat{\gamma}_1}\xi_{1-k/n}^{\gamma_1-\hat{\gamma}_1}\frac{\widehat{\mathrm{VaR}}_X(1-k/n)}{\mathrm{VaR}_X(1-k/n)}\\
&\quad\times(1+O(A_1(1/(1-\tau_n')))),\end{aligned}$$

where the last step follows (2.14) and Assumption 2 (a),

$$\left(\frac{n(1-\tau_n')}{k}\xi_{\tau_n'}\right)^{-\gamma_1}\frac{U_1(n/k)}{U_1\left(1/((1-\tau_n')\xi_{\tau_n'})\right)}=1+O(A_1(1/(1-\tau_n'))).$$

Then, we take logarithm,

$$\begin{aligned}\log \frac{\widetilde{\mathrm{CoVaR}}^{(1)}_{X|Y}(\tau_n')}{\mathrm{CoVaR}_{X|Y}(\tau_n')} &= \left[\hat{\gamma}_1\left(3-\frac{1}{\hat{\eta}}\right)-\gamma_1\left(3-\frac{1}{\eta}\right)\right]\log d_n \\ &\quad - \hat{\gamma}_1 \log \frac{\hat{\xi}_{1-k/n}}{\xi_{1-k/n}} \\ &\quad + (\gamma_1-\hat{\gamma}_1)\log \xi_{1-k/n} \\ &\quad + \log \frac{\widehat{\mathrm{VaR}}_X(1-k/n)}{\mathrm{VaR}_X(1-k/n)} \\ &\quad + \log\left[1+O(A_1(1/(1-\tau_n')))\right] \\ &=: I_1+I_2+I_3+I_4+I_5.\end{aligned}$$

For the term I_1, using (2.19) (under Assumption 2 (f)), there exists a absolute constant K such that,

$$\max\left\{\left|3-\frac{1}{\hat{\eta}}\right|,\left|3-\frac{1}{\eta}\right|\right\}\leq K,$$

and therefore,

$$\left|\frac{k^{2/3}}{n}\left[\hat{\gamma}_1\left(3-\frac{1}{\hat{\eta}}\right)-\gamma_1\left(3-\frac{1}{\eta}\right)\right]\log d_n\right| \leq K\left|\sqrt{k}(\hat{\gamma}_1-\gamma_1)\right|\frac{k}{n}\log d_n = o_{\mathbb{P}}(1), \tag{S2.18}$$

which follows from $\frac{k}{n}\log d_n = o(1)$, because, when $n(1-\tau_n')\to c>0$, it can be verified directly that, as $n\to\infty$,

$$\frac{k}{n}\log d_n = \frac{k}{n}\log\frac{k}{n(1-\tau_n')} \sim \frac{k}{n}(\log k - \log c) = O(n^{\iota-1}\log n) = o(1),$$

when $n(1-\tau_n')\to 0$, it holds by assumed condition.

For the term I_2, using Proposition 2 and the consistency of $\hat{\gamma}_1$,

$$\frac{k^{2/3}}{n}\hat{\gamma}_1 \log \frac{\hat{\xi}_{1-k/n}}{\xi_{1-k/n}} \xrightarrow{d} \gamma_1 \left(\frac{1}{\eta} - 2\right) C_2(0,1)W(1). \tag{S2.19}$$

For the term I_3, we also have that

$$\frac{k^{2/3}}{n}(\gamma_1 - \hat{\gamma}_1) \log \xi_{1-k/n} = \sqrt{k}(\gamma_1 - \hat{\gamma}_1)\frac{k}{n} \log \xi_{1-k/n} = o(1), \tag{S2.20}$$

by noting $\xi_{1-k/n} \sim \left(\frac{k}{n}\right)^{2-\frac{1}{\eta}}$, $\frac{k}{n}\log\frac{k}{n} \to 0$ as $n \to \infty$, and hence $\frac{k}{n}\log \xi_{1-k/n} \to 0$, as $n \to \infty$.

The fourth term I_4 gives

$$\frac{k^{2/3}}{n} \log \frac{\widehat{\mathrm{VaR}}_X(1-k/n)}{\mathrm{VaR}_X(1-k/n)} = o_{\mathbb{P}}(1), \tag{S2.21}$$

by recalling that $\sqrt{k}\log\frac{\widehat{\mathrm{VaR}}_X(1-k/n)}{\mathrm{VaR}_X(1-k/n)} = O_{\mathbb{P}}(1)$.

The final term I_5 gives

$$\frac{k^{2/3}}{n} \log\left[1 + O(A_1(1/(1-\tau_n')))\right] \leq \frac{k}{n}\left|O\left(\sqrt{k}A_1(n/k)\right)\right| = o_{\mathbb{P}}(1), \tag{S2.22}$$

as $n \to \infty$.

Therefore, the final results follows the combination of (S2.15), (S2.16), (S2.17), (S2.18), (S2.19), (S2.20), (S2.21), and (S2.22), which completes this proof.

□

S2.3 Proofs of Proposition 3 and Theorem 2

To prove the asymptotic properties of $\widehat{\mathrm{CoES}}_{X|Y}(1-k/n)$, we need an auxiliary lemma. Before that, we define some new notations for convenience,

$$e_{1n}(x^{-\gamma_1}) := \frac{n}{k}\overline{F}_X(x^{-\gamma_1}\mathrm{CoVaR}_{X|Y}(1-k/n)),$$

$$e_{2n}(x^{-\gamma_1}) := \frac{\overline{F}_X(x^{-\gamma_1}\widehat{\mathrm{CoVaR}}_{X|Y}(1-k/n))}{\overline{F}_X(x^{-\gamma_1}\mathrm{CoVaR}_{X|Y}(1-k/n))},$$

for $x>0$. From the regular variation (2.5), equality (2.7) and consistency for $\widehat{\mathrm{CoVaR}}_{X|Y}(1-k/n)$ established in (S2.29), we obtain that, as $n\to\infty$,

$$\left|e_{1n}(x^{-\gamma_1}) - x\xi_{1-k/n}\right| = o(1), \tag{S2.23}$$

and

$$\left|e_{2n}(x^{-\gamma_1}) - 1\right| = O_{\mathbb{P}}\left(\frac{n}{k^{3/2}}\right). \tag{S2.24}$$

Let $s_n(x^{-\gamma_1}) := e_{1n}(x^{-\gamma_1})e_{2n}(x^{-\gamma_1})$, then, from (S2.23) and (S2.24), it follows,

$$\left|s_n(x^{-\gamma_1}) - x\xi_{1-k/n}\right| = O_{\mathbb{P}}\left(\frac{n}{k^{3/2}}\right). \tag{S2.25}$$

The following lemma shows that, $s_n(x)$ can be substituted by $x\xi_{1-k/n}$ in the limit when handling proper integrals. For convenience, we directly assume (S2.25) holds without restating the conditions below, which can be guaranteed under the conditions in Proposition 3.

Lemma S4. *Suppose that the convergence in* (S2.25) *holds with* $\iota > \frac{2}{3}$*. Denote* $g(\cdot,\cdot)$ *as a bounded and continuous function on* $(0,T] \times [a,b]$ *with* $0 \le a < b < \infty$*. Moreover, suppose that there exist* $\mu_1 > \gamma_1$ *and* $M > 0$ *such that,*

$$\sup_{x\in(0,T],y\in[a,b]} \frac{|g(x,y)|}{x^{\mu_1}} \le M.$$

Then, we have that,

$$\lim_{n\to\infty} \sup_{a\le y\le b} \left| \int_0^1 g(s_n(x^{-\gamma_1}), y) - g(x\xi_{1-k/n}, y)\, dx^{-\gamma_1} \right| = 0. \tag{S2.26}$$

Proof of Lemma S4. Notice first that, using (S2.25), there exists a positive integer N and a $\delta > 0$ such that, for all $x \in (0,1]$ and $n > N$,

$$s_n(x^{-\gamma_1}) < x\xi_{1-k/n}(1+\delta).$$

Then, as both $s_n(x^{-\gamma_1})$ and $x\xi_{1-k/n}$ are positive and bounded for any $x \in (0,1]$, it follows that, when $n > N$,

$$\begin{aligned}
&\sup_{a\le y\le b} \left| \int_0^1 g(s_n(x^{-\gamma_1}), y) - g(x\xi_{1-k/n}, y)\, dx^{-\gamma_1} \right| \\
\le &\sup_{a\le y\le b} \left| \int_0^1 \frac{|g(s_n(x^{-\gamma_1}), y)|}{(s_n(x^{-\gamma_1}))^{\mu_1}} (s_n(x^{-\gamma_1}))^{\mu_1} + \frac{|g(x\xi_{1-k/n}, y)|}{(x\xi_{1-k/n})^{\mu_1}} (x\xi_{1-k/n})^{\mu_1}\, dx^{-\gamma_1} \right| \\
\le &M \sup_{a\le y\le b} \left| \int_0^1 (x\xi_{1-k/n}(1+\delta))^{\mu_1} + (x\xi_{1-k/n})^{\mu_1}\, dx^{-\gamma_1} \right| \\
= &M((1+\delta)^{\mu_1} + 1)\xi_{1-k/n}^{\mu_1} \left| \int_0^1 x^{\mu_1}\, dx^{-\gamma_1} \right| \\
= &\frac{M((1+\delta)^{\mu_1} + 1)\gamma_1}{\mu_1 - \gamma_1} \xi_{1-k/n}^{\mu_1} \\
\to &0,
\end{aligned}$$

as $n \to \infty$, by recalling $\xi_{1-k/n} \to 0$. This proof is complete. $\square$

Proof of Proposition 3. The key step in this proof still remains the utilization of the Vervaat's Lemma (Lemma A.0.2 in de Haan and Ferreira (2006)). Before that, it is necessary to redefine $\widehat{\text{CoVaR}}_{X|Y}(1-k/n)$, since the Vervaat's Lemma is only applicable to nondecreasing functions. Note that,

$$
\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n} I\left(X_i \geq s, Y_i \geq \widehat{\text{VaR}}_Y(1-k/n)\right) \\
=&\frac{1}{n}\sum_{i=1}^{n} I\left(\overline{F}_X(X_i) \leq \frac{\overline{F}_X(s)}{\overline{F}_X(\text{CoVaR}_{X|Y}(1-k/n))}\frac{n}{k}\overline{F}_X(\text{CoVaR}_{X|Y}(1-k/n))\frac{k}{n},\right.\\
&\left.\overline{F}_Y(Y_i) \leq \frac{\overline{F}_Y(\widehat{\text{VaR}}_Y(1-k/n))}{\overline{F}_Y(\text{VaR}_Y(1-k/n))}\frac{k}{n}\right)\\
=&\frac{1}{n}\sum_{i=1}^{n} I\left(\overline{F}_X(X_i) \leq \frac{\overline{F}_X(s)}{\overline{F}_X(\text{CoVaR}_{X|Y}(1-k/n))}\xi_{1-k/n}\frac{k}{n},\right.\\
&\left.\overline{F}_Y(Y_i) \leq \frac{\overline{F}_Y(\widehat{\text{VaR}}_Y(1-k/n))}{\overline{F}_Y(\text{VaR}_Y(1-k/n))}\frac{k}{n}\right)\\
=&\left(\frac{k}{n}\right)^{\frac{1}{\eta}} T_{n/k}\left(\frac{\overline{F}_X(s)}{\overline{F}_X(\text{CoVaR}_{X|Y}(1-k/n))}\xi_{1-k/n}, \frac{\overline{F}_Y(\widehat{\text{VaR}}_Y(1-k/n))}{\overline{F}_Y(\text{VaR}_Y(1-k/n))}\right)\\
=&\left(\frac{k}{n}\right)^{\frac{1}{\eta}} T_{n/k}\left(\frac{\overline{F}_X(s)}{\overline{F}_X(\text{CoVaR}_{X|Y}(1-k/n))}\xi_{1-k/n}, \hat{\theta}_Y(1-k/n)\right),
\end{aligned}
$$

with noting, by (2.7),

$$
\frac{n}{k}\overline{F}_X(\text{CoVaR}_{X|Y}(1-k/n)) = \frac{n}{k}\overline{F}_X(\text{VaR}_X(1-k\xi_{1-k/n}/n)) = \xi_{1-k/n}.
$$

Hence, we can redefine the $\widehat{\text{CoVaR}}_{X|Y}(1-k/n)$ as follows,

$$
\widehat{\text{CoVaR}}_{X|Y}(1-k/n)
$$

$$= \sup\left\{ s \in (0,\infty),\ T_{n/k}\left(\frac{\overline{F}_X(s)}{\overline{F}_X(\mathrm{CoVaR}_{X|Y}(1-k/n))}\xi_{1-k/n}, \hat{\theta}_Y(1-k/n) \right) \geq \left(\frac{k}{n}\right)^{2-\frac{1}{\eta}} \right\},$$

which is equivalent to,

$$\begin{aligned}
&\frac{\overline{F}_X(\widehat{\mathrm{CoVaR}}_{X|Y}(1-k/n))}{\overline{F}_X(\mathrm{CoVaR}_{X|Y}(1-k/n))}\xi_{1-k/n} \\
&= \inf\left\{ s \in (0,\infty),\ T_{n/k}\left(s, \hat{\theta}_Y(1-k/n)\right) \geq \left(\frac{k}{n}\right)^{2-\frac{1}{\eta}} \right\}.
\end{aligned}$$

Next, we consider the following decomposition,

$$\begin{aligned}
&T_{n/k}\left((k/n)^{2-\frac{1}{\eta}}, \hat{\theta}_Y(1-k/n)\right) - C\left((k/n)^{2-\frac{1}{\eta}}, 1\right) \\
=&T_{n/k}\left((k/n)^{2-\frac{1}{\eta}}, \hat{\theta}_Y(1-k/n)\right) - C_{n/k}\left((k/n)^{2-\frac{1}{\eta}}, \hat{\theta}_Y(1-k/n)\right) \\
&+ C_{n/k}\left((k/n)^{2-\frac{1}{\eta}}, \hat{\theta}_Y(1-k/n)\right) - C\left((k/n)^{2-\frac{1}{\eta}}, \hat{\theta}_Y(1-k/n)\right) \\
&+ C\left((k/n)^{2-\frac{1}{\eta}}, \hat{\theta}_Y(1-k/n)\right) - C\left((k/n)^{2-\frac{1}{\eta}}, 1\right) \\
=:&I_1 + I_2 + I_3.
\end{aligned}$$

It is readily to check that $\sqrt{k}I_1 = o_{\mathbb{P}}(1)$ by following a similar argument in the proof of Lemma S3. Secondly, using Assumption 2 (b), it follows that, as $n \to \infty$,

$$\begin{aligned}
\sqrt{k}I_2 &= \sqrt{k}I_2\left((k/n)^{2-\frac{1}{\eta}}\right)^{-\beta_1}\left((k/n)^{2-\frac{1}{\eta}}\right)^{\beta_1} \\
&= O\left(\sqrt{k}\left(\frac{n}{k}\right)^{\alpha}\left((k/n)^{2-\frac{1}{\eta}}\right)^{\beta_1}\right) \\
&= O\left(n^{\left(2-\frac{1}{\eta}\right)(\iota-1)\left[\beta_1 - \frac{\iota/(2(1-\iota))+\alpha}{2-1/\eta}\right]}\right) \qquad \text{(S2.27)} \\
&\to 0,
\end{aligned}$$

by recalling $\beta_1 > 0$ and $\iota < 1 - \frac{1}{1-2\alpha}$, which implies $\beta_1 - \frac{\iota/(2(1-\iota))+\alpha}{2-1/\eta} > 0$.

Thirdly, for I_3, applying intermediate value theorem, we obtain,

$$\sqrt{k} I_3 = C_2\left((k/n)^{2-\frac{1}{\eta}}, \tilde{\theta}\right)\sqrt{k}\left(\hat{\theta}_Y(1-k/n) - 1\right), \tag{S2.28}$$

where $\tilde{\theta}$ is between $\hat{\theta}_Y(1-k/n)$ and 1, tending to 1 as n increases.

Therefore, combining (S2.27), and (S2.28), we find,

$$\left|\sqrt{k}\left(T_{n/k}\left((k/n)^{2-\frac{1}{\eta}}, \hat{\theta}_Y(1-k/n)\right) - C\left((k/n)^{2-\frac{1}{\eta}}, 1\right)\right)\right.$$
$$\left. -C_2\left((k/n)^{2-\frac{1}{\eta}}, 1\right)W(1)\right| = o_{\mathbb{P}}(1),$$

by recalling the asymptotic distribution of $\hat{\theta}_Y(1-k/n)$ given in (S1.13) for Y's version.

Using Vervaat's Lemma and regular variation again, it follows

$$\frac{k^{3/2}}{n}\left(\frac{\widehat{\mathrm{CoVaR}}_{X|Y}(1-k/n)}{\mathrm{CoVaR}_{X|Y}(1-k/n)} - 1\right) \xrightarrow{d} \left(2 - \frac{1}{\eta}\right)\gamma_1 C_2(0,1)W(1), \tag{S2.29}$$

by noting $\xi_{1-k/n}/\xi^*_{1-k/n} \to 0$ and $\frac{(k/n)^{2-\frac{1}{\eta}}}{\xi^*_{1-k/n}} \to C_1(0,1)$ as $n \to \infty$.

Now, we focus on the asymptotic behaviors of $\widehat{\mathrm{CoES}}_{X|Y}(1-k/n)$. Due to $\frac{\mathrm{CoES}_{X|Y}(1-k/n)}{\mathrm{CoVaR}_{X|Y}(1-k/n)} \to \frac{1}{1-\gamma_1}$ as $n \to \infty$, and

$$\begin{aligned}
&\frac{\widehat{\mathrm{CoES}}_{X|Y}(1-k/n)}{\mathrm{CoES}_{X|Y}(1-k/n)} - 1 \\
=&\frac{\mathrm{CoVaR}_{X|Y}(1-k/n)}{\mathrm{CoES}_{X|Y}(1-k/n)}\left[\left(\frac{\widehat{\mathrm{CoVaR}}_{X|Y}(1-k/n)}{\mathrm{CoVaR}_{X|Y}(1-k/n)} - 1\right)\frac{\widehat{\mathrm{CoES}}_{X|Y}(1-k/n)}{\widehat{\mathrm{CoVaR}}_{X|Y}(1-k/n)}\right. \\
&\left.+\frac{\widehat{\mathrm{CoES}}_{X|Y}(1-k/n)}{\widehat{\mathrm{CoVaR}}_{X|Y}(1-k/n)} - \frac{\mathrm{CoES}_{X|Y}(1-k/n)}{\mathrm{CoVaR}_{X|Y}(1-k/n)}\right],
\end{aligned}$$

it hence suffices to find the asymptotic property of $\left(\frac{\widehat{\text{CoES}}_{X|Y}(1-k/n)}{\widehat{\text{CoVaR}}_{X|Y}(1-k/n)} - \frac{\text{CoES}_{X|Y}(1-k/n)}{\text{CoVaR}_{X|Y}(1-k/n)}\right)$ first. Following the decomposition in the proof of Proposition 1, we have that,

$$
\begin{aligned}
&\frac{\text{CoES}_{X|Y}(1-k/n)}{\text{CoVaR}_{X|Y}(1-k/n)} \\
=&1+\frac{n^2}{k^2}\int_1^\infty \mathbb{P}\left(X \geq x\text{CoVaR}_{X|Y}(1-k/n), Y \geq \text{VaR}_Y(1-k/n)\right)\,dx \\
=&1+\frac{n^2}{k^2}\int_1^\infty \mathbb{P}\left(\overline{F}_X(X) \leq \overline{F}_X(x\text{CoVaR}_{X|Y}(1-k/n)), \overline{F}_Y(Y) \leq \frac{k}{n}\right)\,dx \\
=&1+\left(\frac{n}{k}\right)^{2-\frac{1}{\eta}}\int_1^\infty \left(\frac{n}{k}\right)^{\frac{1}{\eta}} \mathbb{P}\left(\overline{F}_X(X) \leq \frac{n}{k}\overline{F}_X(x\text{CoVaR}_{X|Y}(1-k/n))\frac{k}{n},\right. \\
&\left.\overline{F}_Y(Y) \leq \frac{k}{n}\right)\,dx \\
=&1+\left(\frac{n}{k}\right)^{2-\frac{1}{\eta}}\int_1^\infty C_{n/k}\left(\frac{n}{k}\overline{F}_X(x\text{CoVaR}_{X|Y}(1-k/n)),1\right)\,dx \\
=&1+\left(\frac{n}{k}\right)^{2-\frac{1}{\eta}}\int_0^1 C_{n/k}\left(\frac{n}{k}\overline{F}_X(x^{-\gamma_1}\text{CoVaR}_{X|Y}(1-k/n)),1\right)\,dx^{-\gamma_1} \\
=&1+\left(\frac{n}{k}\right)^{2-\frac{1}{\eta}}\int_0^1 C_{n/k}\left(e_{1n}(x^{-\gamma_1}),1\right)\,dx^{-\gamma_1},
\end{aligned}
\tag{S2.30}
$$

where the last step follows a change of variable. Similarly, it also follows the empirical version,

$$
\begin{aligned}
&\frac{\widehat{\mathrm{CoES}}_{X|Y}(1-k/n)}{\widehat{\mathrm{CoVaR}}_{X|Y}(1-k/n)} \\
=&1+\left(\frac{n}{k}\right)^{2-\frac{1}{\eta}}\int_1^\infty T_{n/k}\left(\frac{n}{k}\overline{F}_X(x\widehat{\mathrm{CoVaR}}_{X|Y}(1-k/n)),\frac{n}{k}\overline{F}_Y(\widehat{\mathrm{VaR}}_Y(1-k/n))\right)\,dx \\
=&1+\left(\frac{n}{k}\right)^{2-\frac{1}{\eta}}\int_0^1 T_{n/k}\left(\frac{n}{k}\overline{F}_X(x^{-\gamma_1}\widehat{\mathrm{CoVaR}}_{X|Y}(1-k/n)),\right. \\
&\left.\frac{n}{k}\overline{F}_Y(\widehat{\mathrm{VaR}}_Y(1-k/n))\right)\,dx^{-\gamma_1} \\
=&1+\left(\frac{n}{k}\right)^{2-\frac{1}{\eta}}\int_0^1 T_{n/k}\left(e_{1n}(x^{-\gamma_1})e_{2n}(x^{-\gamma_1}),\hat{\theta}_Y(1-k/n)\right)\,dx^{-\gamma_1} \\
=&1+\left(\frac{n}{k}\right)^{2-\frac{1}{\eta}}\int_0^1 T_{n/k}\left(s_n(x^{-\gamma_1}),\hat{\theta}_Y(1-k/n)\right)\,dx^{-\gamma_1}.
\end{aligned}
\tag{S2.31}
$$

Then, we have that,

$$
\begin{aligned}
&\frac{k^{3/2}}{n}\left(\frac{\widehat{\mathrm{CoES}}_{X|Y}(1-k/n)}{\widehat{\mathrm{CoVaR}}_{X|Y}(1-k/n)}-\frac{\mathrm{CoES}_{X|Y}(1-k/n)}{\mathrm{CoVaR}_{X|Y}(1-k/n)}\right) \\
=&\sqrt{k}\left(\frac{n}{k}\right)^{1-\frac{1}{\eta}}\int_0^1 T_{n/k}\left(s_n(x^{-\gamma_1}),\hat{\theta}_Y(1-k/n)\right)-C_{n/k}\left(e_{1n}(x^{-\gamma_1}),1\right)\,dx^{-\gamma_1} \\
=&\left(\frac{n}{k}\right)^{\frac{1}{2}-\frac{1}{2\eta}}\int_0^1\sqrt{k}\left(\frac{n}{k}\right)^{\frac{1}{2}-\frac{1}{2\eta}}\left\{T_{n/k}\left(s_n(x^{-\gamma_1}),\hat{\theta}_Y(1-k/n)\right)\right. \\
&\quad\left.-C_{n/k}\left(s_n(x^{-\gamma_1}),\hat{\theta}_Y(1-k/n)\right)\right\} \\
&\quad-W_C(x\xi_{1-k/n},1)\,dx^{-\gamma_1} \\
&+\left(\frac{n}{k}\right)^{\frac{1}{2}-\frac{1}{2\eta}}\int_0^1 W_C(x\xi_{1-k/n},1)\,dx^{-\gamma_1} \\
&+\sqrt{k}\left(\frac{n}{k}\right)^{1-\frac{1}{\eta}}\int_0^1 C_{n/k}\left(s_n(x^{-\gamma_1}),\hat{\theta}_Y(1-k/n)\right)-C_{n/k}\left(e_{1n}(x^{-\gamma_1}),1\right)\,dx^{-\gamma_1} \\
=&I_1'+I_2'+I_3'.
\end{aligned}
$$

For term I_2', applying Lemma S1 with $\mu \in (\gamma_1, 1/2]$, it follows that,

$$\begin{aligned}
I_2' &= \left(\frac{n}{k}\right)^{\frac{1}{2}-\frac{1}{2\eta}} \int_0^1 W_C(x\xi_{1-k/n}, 1)\, dx^{-\gamma_1} \\
&\le \left(\frac{n}{k}\right)^{\frac{1}{2}-\frac{1}{2\eta}} \sup_{0<x\le 1} \frac{|W_C(x\xi_{1-k/n}, 1)|}{(x\xi_{1-k/n})^{\mu}} \xi_{1-k/n}^{\mu} \left| \int_0^1 x^{\mu}\, dx^{-\gamma_1} \right| \\
&= \left(\frac{n}{k}\right)^{\frac{1}{2}-\frac{1}{2\eta}} \xi_{1-k/n}^{\mu} \frac{\gamma_1}{\mu_1 - \gamma_1} \sup_{0<x\le 1} \frac{|W_C(x\xi_{1-k/n}, 1)|}{(x\xi_{1-k/n})^{\mu}} \\
&= o(1),
\end{aligned} \tag{S2.32}$$

by recalling $\xi_{1-k/n} \to 0$ as $n \to \infty$.

For term I_1', it follows that

$$\begin{aligned}
I_1' = &\left(\frac{n}{k}\right)^{\frac{1}{2}-\frac{1}{2\eta}} \int_0^1 \sqrt{k} \left(\frac{n}{k}\right)^{\frac{1}{2}-\frac{1}{2\eta}} \Big\{ T_{n/k}\left(s_n(x^{-\gamma_1}), \hat{\theta}_Y(1-k/n)\right) \\
&\quad - C_{n/k}\left(s_n(x^{-\gamma_1}), \hat{\theta}_Y(1-k/n)\right) \Big\} \\
&\quad - W_C(s_n(x^{-\gamma_1}), \hat{\theta}_Y(1-k/n))\, dx^{-\gamma_1} \\
&+ \left(\frac{n}{k}\right)^{\frac{1}{2}-\frac{1}{2\eta}} \int_0^1 W_C(s_n(x^{-\gamma_1}), \hat{\theta}_Y(1-k/n)) - W_C(x\xi_{1-k/n}, \hat{\theta}_Y(1-k/n))\, dx^{-\gamma_1} \\
&+ \left(\frac{n}{k}\right)^{\frac{1}{2}-\frac{1}{2\eta}} \int_0^1 W_C(x\xi_{1-k/n}, \hat{\theta}_Y(1-k/n)) - W_C(x\xi_{1-k/n}, 1)\, dx^{-\gamma_1} \\
=: &\, I_{11}' + I_{12}' + I_{13}'.
\end{aligned}$$

We first consider I_{11}'. There exists a positive integer N_1 and $\delta_1 > 0$ such that, for all $n > N_1$, $s_n(x^{-\gamma_1}) < x$ and $\hat{\theta}_Y(1-k/n) \in (1-\delta_1, 1+\delta_1)$. Hence, for any $\beta \in \left(\gamma_1, \frac{1}{4}\left(3 - \frac{1}{\eta}\right)\right) \subset \left(\gamma_1, \frac{1}{2}\right)$, we have that,

$$
\begin{aligned}
&|I'_{11}| \\
\leq &\left(\frac{n}{k}\right)^{\frac{1}{2}-\frac{1}{2\eta}} \sup_{\substack{0<s\leq 1 \\ 1-\delta_1<t<1+\delta_1}} \left| \frac{\sqrt{k}\left(\frac{n}{k}\right)^{\frac{1}{2}-\frac{1}{2\eta}}\left(T_{n/k}(s,t) - C_{n/k}(s,t)\right) - W_C(s,t)}{s^\beta} \right| \\
&\times \left| \int_0^1 x^\beta \, dx^{-\gamma_1} \right| \\
\to &0,
\end{aligned}
$$

as $n \to \infty$, which follows from Proposition S1 and the integral $\int_0^1 x^\beta \, dx^{-\gamma_1} = \frac{\gamma_1}{\gamma_1 - \beta}$. On the other hand, applying Lemma S1 with $\mu \in (\gamma_1, 1/2)$, we have that,

$$
\sup_{0<x\leq 1} \frac{|W_C(x\xi_{1-k/n}, \hat{\theta}_Y(1-k/n))|}{(x\xi_{1-k/n})^\mu} < \infty.
$$

Moreover, $W_C(x,y)$ is continuous on $(0,T] \times (1/2, 2]$ (see Corollary 1.11 in Adler (1990)). Hence, by using Lemma S4, we have that, as $n \to \infty$,

$$
|I'_{12}| \leq \left(\frac{n}{k}\right)^{\frac{1}{2}-\frac{1}{2\eta}} \sup_{1-\delta_1<y<1+\delta_1} \left| \int_0^1 W_C(s_n(x^{-\gamma_1}), y) - W_C(x\xi_{1-k/n}, y) \, dx^{-\gamma_1} \right| \to 0.
$$

We continue to work on I'_{13}. By using (S1.13) for Y's version, we obtain that,

$$
\lim_{n\to\infty} \mathbb{P}\left(\left| \hat{\theta}_Y(1-k/n) - 1 \right| > k^{-1/4} \right) = 0.
$$

Hence, with probability tending to 1, and $\mu \in (\gamma_1, 1/2)$,

$$
\begin{aligned}
&\mathbb{P}(I'_{13} > \varepsilon)\\
\leq&\mathbb{P}\left(\left(\frac{n}{k}\right)^{\frac{1}{2}-\frac{1}{2\eta}} \sup_{|y-1|<k^{-1/4}} \left|\int_0^1 W_C(x\xi_{1-k/n}, y) - W_C(x\xi_{1-k/n}, 1)\, dx^{-\gamma_1}\right| > \varepsilon\right)\\
\leq&\mathbb{P}\left(2\left(\frac{n}{k}\right)^{\frac{1}{2}-\frac{1}{2\eta}} \sup_{0<s\leq 1, |y-1|<k^{-1/4}} \frac{|W_C(x\xi_{1-k/n}, y)|}{(x\xi_{1-k/n})^{\mu}} \xi^{\mu}_{1-k/n} \left|\int_0^1 x^{\mu}\, dx^{-\gamma_1}\right| > \varepsilon\right)\\
\leq&\mathbb{P}\left(\left(\frac{n}{k}\right)^{\frac{1}{2}-\frac{1}{2\eta}} \sup_{0<s\leq 1, |y-1|<k^{-1/4}} \frac{|W_C(x\xi_{1-k/n}, y)|}{(x\xi_{1-k/n})^{\mu}} \xi^{\mu}_{1-k/n} > \frac{\varepsilon(\mu-\gamma_1)}{2\gamma_1}\right)\\
\to&0,
\end{aligned}
$$

as $n \to \infty$, since both $\left(\frac{n}{k}\right)^{\frac{1}{2}-\frac{1}{2\eta}}$ and $\xi^{\mu}_{1-k/n}$ tend to 0. Now, taking arguments for $I'_{11}, I'_{12}, I'_{13}$ together implies that,

$$
I'_1 = o_{\mathbb{P}}(1). \tag{S2.33}
$$

Lastly, we deal with I'_3. It follows that,

$$
\begin{aligned}
&I'_3\\
=&\sqrt{k}\left(\frac{n}{k}\right)^{1-\frac{1}{\eta}} \int_0^1 \left(C_{n/k}\left(s_n(x^{-\gamma_1}), \hat{\theta}_Y(1-k/n)\right) - C\left(s_n(x^{-\gamma_1}), \hat{\theta}_Y(1-k/n)\right)\right) dx^{-\gamma_1}\\
&+\sqrt{k}\left(\frac{n}{k}\right)^{1-\frac{1}{\eta}} \int_0^1 \left(C\left(s_n(x^{-\gamma_1}), \hat{\theta}_Y(1-k/n)\right) - C\left(e_{1n}(x^{-\gamma_1}), \hat{\theta}_Y(1-k/n)\right)\right) dx^{-\gamma_1}\\
&+\sqrt{k}\left(\frac{n}{k}\right)^{1-\frac{1}{\eta}} \int_0^1 \left(C\left(e_{1n}(x^{-\gamma_1}), \hat{\theta}_Y(1-k/n)\right) - C\left(e_{1n}(x^{-\gamma_1}), 1\right)\right) dx^{-\gamma_1}\\
&+\sqrt{k}\left(\frac{n}{k}\right)^{1-\frac{1}{\eta}} \int_0^1 \left(C\left(e_{1n}(x^{-\gamma_1}), 1\right) - C_{n/k}\left(e_{1n}(x^{-\gamma_1}), 1\right)\right) dx^{-\gamma_1}\\
=:&I'_{31} + I'_{32} + I'_{33} + I'_{34}.
\end{aligned}
$$

For I'_{31}, using Assumption 2 (b) again, we have, there exists a absolute constant K such that,

$$\begin{aligned}
&|I'_{31}| \\
=&\left|\sqrt{k}\left(\frac{n}{k}\right)^{1-\frac{1}{\eta}}\int_0^1\left(C_{n/k}\left(s_n(x^{-\gamma_1}),\hat{\theta}_Y(1-k/n)\right)\right.\right. \\
&\left.\left.-C\left(s_n(x^{-\gamma_1}),\hat{\theta}_Y(1-k/n)\right)\right)\,dx^{-\gamma_1}\right| \\
\leq&\sqrt{k}\left(\frac{n}{k}\right)^{1-\frac{1}{\eta}}\left|\int_0^1\frac{C_{n/k}\left(s_n(x^{-\gamma_1}),\hat{\theta}_Y(1-k/n)\right)-C\left(s_n(x^{-\gamma_1}),\hat{\theta}_Y(1-k/n)\right)}{\left(s_n(x^{-\gamma_1})\right)^{\beta_1}}\right. \\
&\left.\times\left(s_n(x^{-\gamma_1})\right)^{\beta_1}\,dx^{-\gamma_1}\right| \\
\leq&K\sqrt{k}\left(\frac{n}{k}\right)^{1-\frac{1}{\eta}}\int_0^1\left(\frac{n}{k}\right)^{\alpha}\left(s_n(x^{-\gamma_1})\right)^{\beta_1}\,dx^{-\gamma_1} \\
\leq&K\sqrt{k}\left(\frac{n}{k}\right)^{1-\frac{1}{\eta}}\left(\frac{n}{k}\right)^{\alpha}\xi_{1-k/n}^{\beta_1}\left|\int_0^1 x^{\beta_1}\,dx^{-\gamma_1}\right| \\
\leq&K\sqrt{k}\left(\frac{n}{k}\right)^{1-\frac{1}{\eta}+\alpha}\left(\frac{k}{n}\right)^{\beta_1\left(2-\frac{1}{\eta}\right)} \\
=&Kn^{\frac{\iota}{2}+(\iota-1)\left(\frac{1}{\eta}-1-\alpha\right)}\left(\frac{k}{n}\right)^{\beta_1\left(2-\frac{1}{\eta}\right)} \\
\to&0,
\end{aligned}$$

as $n\to\infty$. It is because that $\iota<1-\frac{1}{1-2\alpha}<1-\frac{1}{2/\eta-1-2\alpha}$ and hence $\frac{\iota}{2}+(\iota-1)\left(\frac{1}{\eta}-1-\alpha\right)<0$. A similar result for I'_{34} gives that $I'_{34}=o_{\mathbb{P}}(1)$. Now, we work on I'_{33}. By using intermediate value theorem and the homogeneity of $C(\cdot,\cdot)$, we obtain that,

$$
\begin{aligned}
&|I'_{33}|\\
=&\left|\sqrt{k}\left(\frac{n}{k}\right)^{1-\frac{1}{\eta}}\int_0^1 (\hat{\theta}_Y(1-k/n))^{\frac{1}{\eta}}C\left(\frac{e_{1n}(x^{-\gamma_1})}{\hat{\theta}_Y(1-k/n)},1\right)-C\left(e_{1n}(x^{-\gamma_1}),1\right)\,dx^{-\gamma_1}\right|\\
\leq&\left|\sqrt{k}\left(\frac{n}{k}\right)^{1-\frac{1}{\eta}}(\hat{\theta}_Y(1-k/n))^{\frac{1}{\eta}}\int_0^1 C\left(\frac{e_{1n}(x^{-\gamma_1})}{\hat{\theta}_Y(1-k/n)},1\right)-C\left(e_{1n}(x^{-\gamma_1}),1\right)\,dx^{-\gamma_1}\right|\\
&+\left|\sqrt{k}\left(\frac{n}{k}\right)^{1-\frac{1}{\eta}}\left((\hat{\theta}_Y(1-k/n))^{\frac{1}{\eta}}-1\right)\int_0^1 C\left(e_{1n}(x^{-\gamma_1}),1\right)\,dx^{-\gamma_1}\right|\\
\leq&\left(\frac{n}{k}\right)^{1-\frac{1}{\eta}}\left|\hat{\theta}_Y(1-k/n)\right|^{\frac{1}{\eta}}\left|\int_0^1 C_1(\zeta,1)e_{1n}(x^{-\gamma_1})\sqrt{k}\left(\frac{1}{\hat{\theta}_Y(1-k/n)}-1\right)\,dx^{-\gamma_1}\right|\\
&+\left(\frac{n}{k}\right)^{1-\frac{1}{\eta}}\left|\sqrt{k}\left((\hat{\theta}_Y(1-k/n))^{\frac{1}{\eta}}-1\right)\right|\left|\int_0^1 C(x,1)\,dx^{-\gamma_1}\right|\\
\to&0,
\end{aligned}
$$

as $n\to\infty$, where ζ lies between $\frac{e_{1n}(x^{-\gamma_1})}{\hat{\theta}_Y(1-k/n)}$ and $e_{1n}(x^{-\gamma_1})$, tends to 0. The convergence in the last step follows from the asymptotic property of $\hat{\theta}_Y(1-k/n)$, Assumption 2 (c) and $(n/k)^{1-\frac{1}{\eta}}\to 0$. Finally, we work on I'_{32}. It follows that, using intermediate value theorem again,

$$
\begin{aligned}
I'_{32}&=\sqrt{k}\left(\frac{n}{k}\right)^{1-\frac{1}{\eta}}\int_0^1 C_1\left(\tilde{e},\hat{\theta}_Y(1-k/n)\right)\left(s_n(x^{-\gamma_1})-e_{1n}(x^{-\gamma_1})\right)\,dx^{-\gamma_1}\\
&=\sqrt{k}\left(\frac{n}{k}\right)^{1-\frac{1}{\eta}}\int_0^1 C_1\left(\tilde{e},\hat{\theta}_Y(1-k/n)\right)e_{1n}(x^{-\gamma_1})\left(e_{2n}(x^{-\gamma_1})-1\right)\,dx^{-\gamma_1}\\
&=\left(\frac{n}{k}\right)^{2-\frac{1}{\eta}}\int_0^1 C_1\left(\tilde{e},\hat{\theta}_Y(1-k/n)\right)e_{1n}(x^{-\gamma_1})\frac{k^{3/2}}{n}\left(e_{2n}(x^{-\gamma_1})-1\right)\,dx^{-\gamma_1},
\end{aligned}
$$

where $\tilde{e}$ lies between $s_n(x^{-\gamma_1})$ and $e_{1n}(x^{-\gamma_1})$, tending to 0. On the one hand, due to the asymptotic normality of $\widehat{\text{CoVaR}}_{X|Y}(1-k/n)$ established in (S2.29), we have,

$$\frac{k^{3/2}}{n}\left(e_{2n}(x^{-\gamma_1})-1\right)\xrightarrow{d}\left(\frac{1}{\eta}-2\right)C_2(0,1)W(1),$$

which is independent of x. On the other hand, there exists a constant K such that,

$$\int_0^1 C_1\left(\tilde{e},\hat{\theta}_Y(1-k/n)\right)e_{1n}(x^{-\gamma_1})\,dx^{-\gamma_1}\le KC_1(0,1)\int_0^1 x\,dx^{-\gamma_1}<\infty.$$

Therefore, using dominated convergence theorem and $\left(\frac{k}{n}\right)^{\frac{1}{\eta}-2}\xi_{1-k/n}\to 1/C_1(0,1)$, it follows that,

$$I_3'\xrightarrow{d}\frac{\gamma_1}{1-\gamma_1}\left(2-\frac{1}{\eta}\right)C_2(0,1)W(1). \tag{S2.34}$$

Combining (S2.32), (S2.33), and (S2.34) yields that,

$$\begin{aligned}&\frac{k^{3/2}}{n}\left(\frac{\widehat{\text{CoES}}_{X|Y}(1-k/n)}{\widehat{\text{CoVaR}}_{X|Y}(1-k/n)}-\frac{\text{CoES}_{X|Y}(1-k/n)}{\text{CoVaR}_{X|Y}(1-k/n)}\right)\\&\xrightarrow{d}\frac{\gamma_1}{1-\gamma_1}\left(2-\frac{1}{\eta}\right)C_2(0,1)W(1).\end{aligned} \tag{S2.35}$$

Finally, the joint asymptotic normality for $\widehat{\text{CoVaR}}_{X|Y}(1-k/n)$ and $\widehat{\text{CoES}}_{X|Y}(1-k/n)$ gives that, using (S2.29), (S2.35) and $\frac{\text{CoES}_{X|Y}(1-k/n)}{\text{CoVaR}_{X|Y}(1-k/n)}=$

$\frac{1}{1-\gamma_1}+o(1)$, we have that,

$$
\begin{aligned}
&\frac{k^{3/2}}{n}\left(\frac{\widehat{\mathrm{CoVaR}}_{X|Y}(1-k/n)}{\mathrm{CoVaR}_{X|Y}(1-k/n)}-1,\frac{\widehat{\mathrm{CoES}}_{X|Y}(1-k/n)}{\mathrm{CoES}_{X|Y}(1-k/n)}-1\right)^\top\\
=&\frac{k^{3/2}}{n}\left(\frac{\widehat{\mathrm{CoVaR}}_{X|Y}(1-k/n)}{\mathrm{CoVaR}_{X|Y}(1-k/n)}-1\right)\left(1,\frac{\mathrm{CoVaR}_{X|Y}(1-k/n)}{\mathrm{CoES}_{X|Y}(1-k/n)}\frac{\widehat{\mathrm{CoES}}_{X|Y}(1-k/n)}{\widehat{\mathrm{CoVaR}}_{X|Y}(1-k/n)}\right)^\top\\
&+\left(0,\frac{\mathrm{CoVaR}_{X|Y}(1-k/n)}{\mathrm{CoES}_{X|Y}(1-k/n)}\frac{k^{3/2}}{n}\left(\frac{\widehat{\mathrm{CoES}}_{X|Y}(1-k/n)}{\widehat{\mathrm{CoVaR}}_{X|Y}(1-k/n)}-\frac{\mathrm{CoES}_{X|Y}(1-k/n)}{\mathrm{CoVaR}_{X|Y}(1-k/n)}\right)\right)^\top\\
=&\frac{k^{3/2}}{n}\left(\frac{\widehat{\mathrm{CoVaR}}_{X|Y}(1-k/n)}{\mathrm{CoVaR}_{X|Y}(1-k/n)}-1\right)\\
&\times\left(1,\frac{\mathrm{CoVaR}_{X|Y}(1-k/n)}{\mathrm{CoES}_{X|Y}(1-k/n)}\left(\frac{\mathrm{CoES}_{X|Y}(1-k/n)}{\mathrm{CoVaR}_{X|Y}(1-k/n)}+O_{\mathbb{P}}\left(\frac{n}{k^{3/2}}\right)\right)\right)^\top\\
&+\left(0,\frac{\mathrm{CoVaR}_{X|Y}(1-k/n)}{\mathrm{CoES}_{X|Y}(1-k/n)}\frac{k^{3/2}}{n}\left(\frac{\widehat{\mathrm{CoES}}_{X|Y}(1-k/n)}{\widehat{\mathrm{CoVaR}}_{X|Y}(1-k/n)}-\frac{\mathrm{CoES}_{X|Y}(1-k/n)}{\mathrm{CoVaR}_{X|Y}(1-k/n)}\right)\right)^\top\\
\xrightarrow{d}&(1,2)^\top\left(2-\frac{1}{\eta}\right)\gamma_1C_2(0,1)W(1).
\end{aligned}
$$

This proof is therefore complete. □

Proof of Theorem 2. We begin with the decomposition in (S2.15) for $i=2$, that is,

$$
\begin{aligned}
&\left(\frac{\widetilde{\mathrm{CoVaR}}^{(2)}_{X|Y}(\tau_n')}{\mathrm{CoVaR}_{X|Y}(\tau_n')}-1,\frac{\widetilde{\mathrm{CoES}}^{(2)}_{X|Y}(\tau_n')}{\mathrm{CoES}_{X|Y}(\tau_n')}-1\right)^\top\\
=&\left(1,\frac{1}{1-\hat\gamma_1}\frac{\mathrm{CoVaR}_{X|Y}(\tau_n')}{\mathrm{CoES}_{X|Y}(\tau_n')}\right)^\top\left(\frac{\widetilde{\mathrm{CoVaR}}^{(2)}_{X|Y}(\tau_n')}{\mathrm{CoVaR}_{X|Y}(\tau_n')}-1\right)^\top\\
&+\left(0,\frac{1}{1-\hat\gamma_1}\frac{\mathrm{CoVaR}_{X|Y}(\tau_n')}{\mathrm{CoES}_{X|Y}(\tau_n')}-1\right)^\top,
\end{aligned}
$$

which implies that we only need to work on $\frac{\widetilde{\mathrm{CoVaR}}^{(2)}_{X|Y}(\tau_n')}{\mathrm{CoVaR}_{X|Y}(\tau_n')}$. Following the proof of Theorem 1, we have that,

$$\begin{aligned}
&\frac{\widetilde{\text{CoVaR}}^{(2)}_{X|Y}(\tau_n')}{\text{CoVaR}_{X|Y}(\tau_n')}\\
=&\left(\frac{k}{n(1-\tau_n')}\right)^{\hat{\gamma}_1\left(3-\frac{1}{\hat{\eta}}\right)}\frac{\widehat{\text{CoVaR}}_{X|Y}(1-k/n)}{\text{CoVaR}_{X|Y}(1-k/n)}\frac{\text{CoVaR}_{X|Y}(1-k/n)}{\text{CoVaR}_{X|Y}(\tau_n')}\\
=&\left(\frac{k}{n(1-\tau_n')}\right)^{\hat{\gamma}_1\left(3-\frac{1}{\hat{\eta}}\right)-\gamma_1\left(3-\frac{1}{\eta}\right)}\frac{\widehat{\text{CoVaR}}_{X|Y}(1-k/n)}{\text{CoVaR}_{X|Y}(1-k/n)}(1+O(A_1(1/(1-\tau_n')))),
\end{aligned}$$

and hence,

$$\begin{aligned}
&\log\frac{\widetilde{\text{CoVaR}}^{(2)}_{X|Y}(\tau_n')}{\text{CoVaR}_{X|Y}(\tau_n')}\\
=&\left[\hat{\gamma}_1\left(3-\frac{1}{\hat{\eta}}\right)-\gamma_1\left(3-\frac{1}{\eta}\right)\right]\log d_n+\log\frac{\widehat{\text{CoVaR}}_{X|Y}(1-k/n)}{\text{CoVaR}_{X|Y}(1-k/n)}+o\left(\frac{n}{k^{3/2}}\right).
\end{aligned}$$

On the one hand, it follows that,

$$\begin{aligned}
&\frac{k^{3/2}}{n}\left[\hat{\gamma}_1\left(3-\frac{1}{\hat{\eta}}\right)-\gamma_1\left(3-\frac{1}{\eta}\right)\right]\log d_n\\
=&\sqrt{k}\left[(\hat{\gamma}_1-\gamma_1)\left(3-\frac{1}{\hat{\eta}}\right)-\gamma_1\left(\frac{1}{\hat{\eta}}-\frac{1}{\eta}\right)\right]\frac{k}{n}\log d_n\\
\to&0,
\end{aligned}$$

by recalling $\frac{k}{n}\log d_n\to 0$ as $n\to\infty$. On the other hand, using (S2.29), we have,

$$\frac{k^{3/2}}{n}\log\frac{\widetilde{\text{CoVaR}}^{(2)}_{X|Y}(\tau_n')}{\text{CoVaR}_{X|Y}(\tau_n')}\xrightarrow{d}\left(2-\frac{1}{\eta}\right)\gamma_1 C_2(0,1)W(1). \tag{S2.36}$$

Then, combining (S2.16), (S2.17) with (S2.36) yields the final asymptotic normality (2.32).

Similarly, for $\widetilde{\text{CoES}}^{(3)}_{X|Y}(\tau_n')$, we have,

$$\log \frac{\widetilde{\text{CoES}}^{(3)}_{X|Y}(\tau_n')}{\text{CoES}_{X|Y}(\tau_n')}$$
$$= \left[\hat{\gamma}_1\left(3-\frac{1}{\hat{\eta}}\right)-\gamma_1\left(3-\frac{1}{\eta}\right)\right]\log d_n + \log\frac{\widehat{\text{CoES}}_{X|Y}(1-k/n)}{\text{CoES}_{X|Y}(1-k/n)} + o\left(\frac{n}{k^{3/2}}\right).$$

Then, the joint asymptotic normality (2.33) follows from Proposition 3. We here omit the repetitive steps for space-saving and this proof is hence finished.

□

S3 Algorithms for computing intermediate adjustment factor and CoVaR

This section provides two algorithms for computing $\hat{\xi}_{1-k/n}$ and $\widehat{\text{CoVaR}}_{X|Y}(1-k/n)$ involved in (2.20) and (2.26). Note that, the reason why we define the integer $m := \lceil \frac{k^2}{n} \rceil$ in Algorithm S1 is that, from the definition (2.20), $\widehat{C}_{n/k}(\xi,1) \geq (k/n)^{2-\frac{1}{\hat{\eta}}}$ is equivalent to

$$\sum_{i=1}^{n} I\left(1-\widehat{F}_X(X_i) \leq \frac{k\xi}{n}, 1-\widehat{F}_Y(Y_i) \leq \frac{k}{n}\right) \geq \frac{k^2}{n},$$

which removes the influence of $\left(\frac{n}{k}\right)^{\frac{1}{\hat{\eta}}}$ and simplifies the calculation.

Algorithm S1 Procedure to compute $\hat{\xi}_{1-k/n}$.

Require: Sample size n, intermediate k, and bivariate samples $\{(X_i, Y_i)\}_{i=1}^n$.

1: Generate empirical distributions $\widehat{F}_X(\cdot)$ and $\widehat{F}_Y(\cdot)$; Generate samples of ranks statistics $\{R_i^X, R_i^Y\}_{i=1}^n$;

2: Calculate the positive integer $m := \lceil \frac{k^2}{n} \rceil$, where $\lceil \cdot \rceil$ denotes the ceiling function;

3: Denote $\{Z_i^X, Z_i^Y\} := \{1 - \widehat{F}_X(X_i), 1 - \widehat{F}_Y(Y_i)\}$ for $i = 1, ..., n$;

4: Filter samples by the indictor $I(Z_i^Y \leq k/n)$ and obtain a sub-samples

$$\{(\widetilde{Z}_i^X, \widetilde{Z}_i^Y)\} := \{(Z_i^X I(Z_i^Y \leq k/n), Z_i^Y I(Z_i^Y \leq k/n)\},$$

with sample size $k + 1$.

5: Then, it follows that $\hat{\xi}_{1-k/n} := \frac{n}{k} \widetilde{Z}_{m,k+1}^X$, where $\widetilde{Z}_{m,k+1}^X$ denotes the m-th order statistic of the sub-samples.

Algorithm S2 Procedure to compute $\widehat{\text{CoVaR}}_{X|Y}(1 - k/n)$.

Require: Sample size n, intermediate k, and bivariate samples $\{(X_i, Y_i)\}_{i=1}^n$.

1: Calculate the positive integer $m := \lceil \frac{k^2}{n} \rceil$, where $\lceil \cdot \rceil$ denotes the ceiling function;

2: Calculate the empirical quantile estimator for Y, $\widehat{\text{VaR}}_Y(1 - k/n) := Y_{n-k,n}$;

3: Filter samples by the indictor $I(Y_i \geq \widehat{\text{VaR}}_Y(1 - k/n))$ and obtain a sub-samples

$$\{(X_i^*, Y_i^*)\} = \{(X_i I(Y_i \geq \widehat{\text{VaR}}_Y(1 - k/n)), Y_i I(Y_i \geq \widehat{\text{VaR}}_Y(1 - k/n)))\},$$

with sample size $k + 1$.

4: Output,

$$\widehat{\text{CoVaR}}_{X|Y}(1 - k/n) := X_{k+2-m,k+1}^*,$$

where $X_{k+2-m,k+1}^*$ denotes as the $(k + 2 - m)$-th order statistic of the sub-samples.

S4 Consistency analysis

This section provides a brief discussion on the consistency of $\hat{\xi}_{1-k/n}$, *that is*, its consistency remains valid for all $\eta \in (1/2, 1)$ without restricting it to a narrower interval $\eta \in \left(\frac{7+\sqrt{17}}{16}, 1\right)$.

According to the definitions of $\hat{\xi}_{1-k/n}$ and $\xi^*_{1-k/n}$, we have, as n becomes large,

$$\begin{aligned}
O_{\mathbb{P}}\left(\frac{1}{n^{2-\frac{1}{\eta}}}\right) &= \widehat{C}_{n/k}(\hat{\xi}_{1-k/n}, 1) - C(\xi^*_{1-k/n}, 1) \\
&= \widehat{C}_{n/k}(\hat{\xi}_{1-k/n}, 1) - T_{n/k}(\hat{\xi}_{1-k/n}, 1) \\
&\quad + T_{n/k}(\hat{\xi}_{1-k/n}, 1) - C_{n/k}(\hat{\xi}_{1-k/n}, 1) \\
&\quad + C_{n/k}(\hat{\xi}_{1-k/n}, 1) - C(\hat{\xi}_{1-k/n}, 1) \\
&\quad + C_1(\tilde{\xi}, 1)\left(\hat{\xi}_{1-k/n} - \xi_{1-k/n}(1+o(1))\right) \\
&= I_1 + I_2 + I_3 + I_4,
\end{aligned}$$

where $\tilde{\xi}$ is an intermediate point between $\hat{\xi}_{1-k/n}$ and $\xi^*_{1-k/n}$. Using Proposition S1 and Assumption 2 (b), we can find $I_2 = o_{\mathbb{P}}(1)$ and $I_3 = o_{\mathbb{P}}(1)$. For I_1, we first note that,

$$\begin{aligned}
&\widehat{C}_{n/k}(\hat{\xi}_{1-k/n}, 1) \\
&= \left(\frac{n}{k}\right)^{\frac{1}{\eta}} \frac{1}{n} \sum_{i=1}^{n} I\left(1 - \widehat{F}_X(X_i) \le \frac{k}{n}\hat{\xi}_{1-k/n}, 1 - \widehat{F}_Y(Y_i) \le \frac{k}{n}\right) \\
&= \left(\frac{n}{k}\right)^{\frac{1}{\eta}} \frac{1}{n} \sum_{i=1}^{n} I\left(X_i \ge \widehat{\mathrm{VaR}}_X(1 - k\hat{\xi}_{1-k/n}/n), Y_i \ge \widehat{\mathrm{VaR}}_Y(1 - k/n)\right)
\end{aligned}$$

$$
\begin{aligned}
&= \left(\frac{n}{k}\right)^{\frac{1}{\hat{\eta}}} \frac{1}{n} \sum_{i=1}^{n} I\left(\overline{F}_X(X_i) \leq \frac{\overline{F}_X(\widehat{\mathrm{VaR}}_X(1-k\hat{\xi}_{1-k/n}/n))}{\overline{F}_X(\mathrm{VaR}_X(1-k\hat{\xi}_{1-k/n}/n))}\frac{k}{n}\hat{\xi}_{1-k/n}, \right. \\
&\quad \left. \overline{F}_Y(Y_i) \leq \frac{\overline{F}_Y(\widehat{\mathrm{VaR}}_Y(1-k/n))}{\overline{F}_Y(\mathrm{VaR}_Y(1-k/n))}\frac{k}{n}\right) \\
&=: T_{n/k}\left(\underbrace{\frac{\overline{F}_X(\widehat{\mathrm{VaR}}_X(1-k\hat{\xi}_{1-k/n}/n))}{\overline{F}_X(\mathrm{VaR}_X(1-k\hat{\xi}_{1-k/n}/n))}}_{\hat{\theta}_X}\hat{\xi}_{1-k/n}, \underbrace{\frac{\overline{F}_Y(\widehat{\mathrm{VaR}}_Y(1-k/n))}{\overline{F}_Y(\mathrm{VaR}_Y(1-k/n))}}_{\hat{\theta}_Y}\right) + o_{\mathbb{P}}(1).
\end{aligned}
$$

Then, term I_1 follows

$$
\begin{aligned}
&\widehat{C}_{n/k}(\hat{\xi}_{1-k/n}, 1) - T_{n/k}(\hat{\xi}_{1-k/n}, 1) \\
=& T_{n/k}\left(\hat{\theta}_X\hat{\xi}_{1-k/n}, \hat{\theta}_Y\right) - C_{n/k}\left(\hat{\theta}_X\hat{\xi}_{1-k/n}, \hat{\theta}_Y\right) \\
&+ C_{n/k}\left(\hat{\theta}_X\hat{\xi}_{1-k/n}, \hat{\theta}_Y\right) - C\left(\hat{\theta}_X\hat{\xi}_{1-k/n}, \hat{\theta}_Y\right) \\
&+ C\left(\hat{\theta}_X\hat{\xi}_{1-k/n}, \hat{\theta}_Y\right) - C\left(\hat{\xi}_{1-k/n}, 1\right) \\
&+ C\left(\hat{\xi}_{1-k/n}, 1\right) - C_{n/k}\left(\hat{\xi}_{1-k/n}, 1\right) \\
&+ C_{n/k}\left(\hat{\xi}_{1-k/n}, 1\right) - T_{n/k}\left(\hat{\xi}_{1-k/n}, 1\right) + o_{\mathbb{P}}(1) \\
=:& I_{11} + I_{12} + I_{13} + I_{14} + I_{15} + o_{\mathbb{P}}(1).
\end{aligned}
$$

Using intermediate value theorem again, it follows that

$$
I_{13} = C_1(\tilde{\theta}_X\hat{\xi}_{1-k/n}, \hat{\theta}_Y)\hat{\xi}_{1-k/n}(\hat{\theta}_X - 1) + C_2(\hat{\xi}_{1-k/n}, \tilde{\theta}_Y)(\hat{\theta}_Y - 1) = o_{\mathbb{P}}(1),
$$

where $\tilde{\theta}_X$ is between $\hat{\theta}_X$ and 1, while $\tilde{\theta}_Y$ is between $\hat{\theta}_Y$ and 1. Using Proposition S1 and Assumption 2 (b) frequently, we also have that $I_{11} = o_{\mathbb{P}}(1)$, $I_{12} = o_{\mathbb{P}}(1)$, $I_{14} = o_{\mathbb{P}}(1)$ and $I_{15} = o_{\mathbb{P}}(1)$. Now, combining above arguments yields that,

$$C_1(\tilde{\xi},1)(\hat{\xi}_{1-k/n}-\xi_{1-k/n})=O_{\mathbb{P}}\left(\frac{1}{n^{2-\frac{1}{\eta}}}\right),$$

and hence,

$$C_1(\tilde{\xi},1)\left(\frac{\hat{\xi}_{1-k/n}}{\xi_{1-k/n}}-1\right)=O_{\mathbb{P}}\left(\frac{1}{k^{2-\frac{1}{\eta}}}\right),$$

by $\xi_{1-k/n}\sim(k/n)^{2-\frac{1}{\eta}}$, which implies the consistency.

In summary, there is no need to impose a particular rate such that it converges to a certain Gaussian; therefore, it suffices to consider the universal range of $\eta\in(1/2,1)$ to guarantee the consistency. By implementing a similar analysis, we can also obtain the consistency for $\widehat{\mathrm{CoVaR}}_{X|Y}(1-k/n)$ and $\widehat{\mathrm{CoES}}_{X|Y}(1-k/n)$, as well as the proposed extreme estimators. We omit these more cumbersome details.

S5 Some theoretical clarifications

In this section, we provide a few additional intermediate steps to clarify several asymptotic relationships presented in the main article, where some derivations are presented in a condensed form and may be difficult to follow.

The first one involves (2.9) - (2.10). By the definition of $\mathrm{CoVaR}_{X|Y}(\tau)$ given in (1.1), we find

$$\mathbb{P}(X\geq\mathrm{CoVaR}_{X|Y}(\tau),Y\geq\mathrm{VaR}_Y(\tau))=(1-\tau)^2,$$

which implies that, using the equity $\mathrm{CoVaR}_{X|Y}(\tau) = \mathrm{VaR}_X(1-(1-\tau)\xi_\tau)$,

$$\begin{aligned}
&(1-\tau)^{2-\frac{1}{\eta}} \\
&= (1-\tau)^{-\frac{1}{\eta}}\,\mathbb{P}(X \geq \mathrm{CoVaR}_{X|Y}(\tau), Y \geq \mathrm{VaR}_Y(\tau)) \\
&= (1-\tau)^{-\frac{1}{\eta}}\,\mathbb{P}(\overline{F}_X(X) \leq \overline{F}_X(\mathrm{CoVaR}_{X|Y}(\tau)), \overline{F}_Y(Y) \leq \overline{F}_Y(\mathrm{VaR}_Y(\tau))) \\
&= (1-\tau)^{-\frac{1}{\eta}}\,\mathbb{P}(\overline{F}_X(X) \leq (1-\tau)\xi_\tau, \overline{F}_Y(Y) \leq 1-\tau) \\
&= (1-\tau)^{-\frac{1}{\eta}}\,\mathbb{P}(\overline{F}_X(X) \leq \xi_\tau/(1-\tau)^{-1}, \overline{F}_Y(Y) \leq 1/(1-\tau)^{-1}).
\end{aligned}$$

By letting $\tau \uparrow 1$ (implying $(1-\tau)^{-1} \to \infty$), and using the important limit (1.4), we can find a ξ_τ^*, which serves as an approximation of ξ_τ such that,

$$(1-\tau)^{-\frac{1}{\eta}}\,\mathbb{P}(\overline{F}_X(X) \leq \xi_\tau/(1-\tau)^{-1}, \overline{F}_Y(Y) \leq 1/(1-\tau)^{-1}) \sim C(\xi_\tau^*, 1)$$

and

$$C(\xi_\tau^*, 1) = (1-\tau)^{2-\frac{1}{\eta}},$$

which is achievable as long as τ is sufficiently large such that $(1-\tau)^{2-\frac{1}{\eta}} < C(1,1)$. Note that ξ_τ^* is not the limit but a good approximate of ξ_τ, and ξ_τ^* depends on τ as well.

The second one involves extrapolative route between $\mathrm{CoVaR}_{X|Y}(\tau_n')$ and $\mathrm{VaR}_X(1-k/n)$: (2.13) - (2.15). Using regular variation on $U_1(\cdot)$ yields that,

$$\begin{aligned}
\frac{\mathrm{CoVaR}_{X|Y}(\tau_n')}{\mathrm{VaR}_X(1-k/n)} &= \frac{\mathrm{VaR}_X(1-(1-\tau_n')\xi_{\tau_n'})}{\mathrm{VaR}_X(1-k/n)} \\
&= \frac{U_1\left(\frac{k}{n(1-\tau_n')\xi_{\tau_n'}} \cdot \frac{n}{k}\right)}{U_1\left(\frac{n}{k}\right)}
\end{aligned}$$

$$\sim \left(\frac{k}{n(1-\tau_n')\xi_{\tau_n'}}\right)^{\gamma_1} \quad \text{(by regular variation)}$$
$$= \left(\frac{k}{n(1-\tau_n')}\right)^{\gamma_1} \xi_{\tau_n'}^{-\gamma_1},$$

by noting that $n/k \to \infty$ as $n \to \infty$. For the extrapolation of $\xi_{\tau_n'}$, we use Lemma 1 to obtain that,

$$\frac{\xi_{\tau_n'}}{\xi_{1-k/n}} = \frac{\frac{\xi_{\tau_n'}}{(1-\tau_n')^{2-\frac{1}{\eta}}} \cdot (1-\tau_n')^{2-\frac{1}{\eta}}}{\frac{\xi_{1-k/n}}{(k/n)^{2-\frac{1}{\eta}}} \cdot (k/n)^{2-\frac{1}{\eta}}} \sim \left(\frac{n(1-\tau_n')}{k}\right)^{2-\frac{1}{\eta}}, \tag{S5.37}$$

which implies (2.14), *that is*,

$$\xi_{\tau_n'} \sim \left(\frac{n(1-\tau_n')}{k}\right)^{2-\frac{1}{\eta}} \xi_{1-k/n}.$$

Substituting the approximation of $\xi_{\tau_n'}$ into the above equation, we obtain that

$$\frac{\text{CoVaR}_{X|Y}(\tau_n')}{\text{VaR}_X(1-k/n)} \sim \left(\frac{k}{n(1-\tau_n')}\right)^{\gamma_1} \xi_{\tau_n'}^{-\gamma_1} \sim \left(\frac{k}{n(1-\tau_n')}\right)^{\gamma_1\left(3-\frac{1}{\eta}\right)} \xi_{1-k/n}^{-\gamma_1},$$

suggesting the extrapolative expression (2.15) of $\text{CoVaR}_{X|Y}(\tau_n')$ by moving $\text{VaR}_X(1-k/n)$ from the left-hand side to the right-hand side.

The third one involves the extrapolative route between $\text{CoVaR}_{X|Y}(\tau_n')$ and $\text{CoVaR}_{X|Y}(1-k/n)$. Using (2.7), (S5.37) and regular variation again, we find that,

$$\frac{\text{CoVaR}_{X|Y}(\tau_n')}{\text{CoVaR}_{X|Y}(1-k/n)} = \frac{\text{VaR}_X(1-(1-\tau_n')\xi_{\tau_n'})}{\text{VaR}_X(1-k\xi_{1-k/n}/n)}$$
$$= \frac{U_1\left(\frac{k\xi_{1-k/n}}{n(1-\tau_n')\xi_{\tau_n'}} \cdot \frac{n}{k\xi_{1-k/n}}\right)}{U_1\left(\frac{n}{k\xi_{1-k/n}}\right)}$$

$$\sim \left(\frac{k}{n(1-\tau_n')} \cdot \frac{\xi_{1-k/n}}{\xi_{\tau_n'}} \right)^{\gamma_1} \quad \text{(by regular variation)}$$
$$\sim \left(\frac{k}{n(1-\tau_n')} \right)^{\gamma_1\left(3-\frac{1}{\eta}\right)} \quad \text{(by (S5.37)),}$$

which suggests the extrapolation (2.29) when moving the $\text{CoVaR}_{X|Y}(1 - k/n)$ from the left-hand side to the right-hand side and substituting all the estimators of unknown quantities.

S6 Bootstrap analysis

In this section, we provide a bootstrap procedure to approximate the asymptotic variances and to construct confidence intervals for practical implementation. Before presenting the detailed procedures, we first review some representative works on bootstrapping extreme value statistics. Draisma et al. (1999) and Danielsson et al. (2001) present a bootstrap method to estimate extreme value index by choosing the suitable number of order statistics. Peng and Qi (2008) first derives a bootstrap approximation for a tail dependence function and then applies it to construct a confidence band for the tail dependence function. Bücher and Dette (2013) bootstraps the tail copula via a random weighted method. Recently, de Haan and Zhou (2024) develops a bootstrap analysis of tail quantile process for the Peaks-over-Threshold method and the Block Maxima method, with applications in constructing confidence intervals for Probability Weighted Moment estimator.

Here, we adapt the moving block bootstrap (MBB) method proposed by Künsch (1989) to approximate the theoretical asymptotic variances. Unlike resampling individual observations, the MBB resamples blocks of consecutive observations at a time, making it a widely used alternative in practice, see Gomes and Neves (2015); Fung et al. (2026). Given original sample $\{X_i, Y_i\}_{i=1}^{n}$, we define $\mathcal{B}_i = \{(X_i, Y_i), ..., (X_{i+l-1}, Y_{i+l-1})\}$, where $l = l(n)$ denotes the block length such that $l \to \infty$ and $m = \lfloor n/l \rfloor \to \infty$. In practice, one may choose $l = O\left(n^{1/3}\right)$ as suggested by Hall et al. (1995). The detailed procedure is given as follows:

- ***Step 1.*** Resample m blocks randomly with replacement from $\mathcal{B}_1, ..., \mathcal{B}_{n-l+1}$, and rearrange elements in all m blocks in a sequence to get the bootstrap sample $\{X_i^{*b}, Y_i^{*b}\}_{i=1}^{ml}$;

- ***Step 2.*** Based on the bootstrap sample $\{X_i^{*b}, Y_i^{*b}\}_{i=1}^{ml}$, letting $\tilde{n} = ml$, compute $\hat{\gamma}_1^{*b}$, $\hat{\eta}^{*b}$, $\hat{\xi}_{1-k/\tilde{n}}^{*b}$, $\widehat{\mathrm{VaR}}_X^{*b}(1-k/\tilde{n})$, $\widehat{\mathrm{CoVaR}}_{X|Y}^{*b}(1-k/\tilde{n})$, and $\widehat{\mathrm{CoES}}_{X|Y}^{*b}(1-k/\tilde{n})$;

- ***Step 3.*** Compute the extrapolative estimators: $\widetilde{\mathrm{CoVaR}}_{X|Y}^{(j),*b}(\tau_n')$ for $j = 1, 2$, and $\widetilde{\mathrm{CoES}}_{X|Y}^{(j),*b}(\tau_n')$ for $j = 1, 2, 3$;

- ***Step 4.*** Repeat the above steps B times to get $\left\{\widetilde{\mathrm{CoVaR}}^{(j),*b}_{X|Y}(\tau_n')\right\}_{b=1}^{B}$ for $j=1,2$, and $\left\{\widetilde{\mathrm{CoES}}^{(j),*b}_{X|Y}(\tau_n')\right\}_{b=1}^{B}$ for $j=1,2,3$. Denote
$$\delta_{CR}^{j,b}=\frac{k^{3/2}}{n}\left(\frac{\widetilde{\mathrm{CoVaR}}^{(j),*b}_{X|Y}(\tau_n')}{\widetilde{\mathrm{CoVaR}}^{(j)}_{X|Y}(\tau_n')}-1\right),$$
and
$$\delta_{CES}^{j,b}=\frac{k^{3/2}}{n}\left(\frac{\widetilde{\mathrm{CoES}}^{(j),*b}_{X|Y}(\tau_n')}{\widetilde{\mathrm{CoES}}^{(j)}_{X|Y}(\tau_n')}-1\right).$$
Then the approximations of asymptotic variances are the empirical variances of $\delta_{CR}^{j,b}$ for $\widetilde{\mathrm{CoVaR}}^{(j),*b}_{X|Y}(\tau_n')$ and $\delta_{CES}^{j,b}$ for $\widetilde{\mathrm{CoES}}^{(j),*b}_{X|Y}(\tau_n')$ over $b=1,...,B$, respectively.

Given the bootstrap procedure described above, two-sided confidence intervals can be constructed in such way, for a confidence level $q\in(0,1)$ (say 95%),
$$\mathrm{CI}^j_{\mathrm{CR}}:=\left[\frac{\widetilde{\mathrm{CoVaR}}^{(j)}_{X|Y}(\tau_n')}{\Delta^{\mathrm{CR}}_{B,[B(1+q)/2]}},\frac{\widetilde{\mathrm{CoVaR}}^{(j)}_{X|Y}(\tau_n')}{\Delta^{\mathrm{CR}}_{B,[B(1-q)/2]}}\right],$$
$$\mathrm{CI}^j_{\mathrm{CES}}:=\left[\frac{\widetilde{\mathrm{CoES}}^{(j)}_{X|Y}(\tau_n')}{\Delta^{\mathrm{CES}}_{B,[B(1+q)/2]}},\frac{\widetilde{\mathrm{CoES}}^{(j)}_{X|Y}(\tau_n')}{\Delta^{\mathrm{CES}}_{B,[B(1-q)/2]}}\right],$$
where $\Delta^{\mathrm{CR}}_{B,1}\le\cdots\le\Delta^{\mathrm{CR}}_{B,B}$ and $\Delta^{\mathrm{CES}}_{B,1}\le\cdots\le\Delta^{\mathrm{CES}}_{B,B}$ denote the order statistics of $\left\{\frac{\widetilde{\mathrm{CoVaR}}^{(j),*b}_{X|Y}(\tau_n')}{\widetilde{\mathrm{CoVaR}}^{(j)}_{X|Y}(\tau_n')}\right\}_{b=1}^{B}$ and $\left\{\frac{\widetilde{\mathrm{CoES}}^{(j),*b}_{X|Y}(\tau_n')}{\widetilde{\mathrm{CoES}}^{(j)}_{X|Y}(\tau_n')}\right\}_{b=1}^{B}$, respectively. This method is easy to implement in practice and avoids the cumbersome computation of asymptotic variances. Similar approaches for constructing confidence intervals have also been adopted in many studies, such as Leng et al. (2024),

Francq and Zakoïan (2025), Fung et al. (2026). We construct the this types of confidence intervals $\mathrm{CI}^{j}_{\mathrm{CR}}$, $\mathrm{CI}^{j}_{\mathrm{CES}}$ for empirical application to implement practical inferences in Section S8.3.

The analyses for theoretical validity of this MBB procedure remain substantial space and interest for further exploration and faces two main difficulties. On the one hand, as noted in the literature cited above, most studies on bootstrapping extreme value statistics are based on the standard bootstrap, rather than MBB. The asymptotic theory of MBB has been investigated in Bühlmann (1994), where the weak convergence of the blockwise bootstrapped empirical process is established. However, it can not be directly applied to the bootstrap of extreme value estimators, since bootstrapping extreme value estimators necessarily relies on the MBB analysis based on the "tail" empirical process. On the other hand, the theoretical analyses inevitably involves the joint study of several extreme value estimators, including $\hat{\gamma}_1^{*b}$, $\hat{\eta}^{*b}$, $\hat{\xi}^{*b}_{1-k/n}$, $\widehat{\mathrm{VaR}}^{*b}_X(1-k/n)$, $\widehat{\mathrm{CoVaR}}^{*b}_{X|Y}(1-k/n)$, and $\widehat{\mathrm{CoES}}^{*b}_{X|Y}(1-k/n)$, which substantially increases the technical complexity and may not be fully carried out within a limited space. Therefore, we leave this highly challenging task as a future work and only provide a feasible algorithmic procedure here.

S7 Additional simulation results

This section includes some additional results to support simulation study. Specifically, Figures S1 - S3 report the boxplots of the ratios between the eight methods and their true values for **Models 1 - 3** under different sample size and risk levels.

Moreover, we conduct a robustness analysis to evaluate the performance of the proposed methods under varying degrees of tail heaviness and extremal dependence, as well as to show the robustness and superiority of the proposed methods over the benchmarks. In connection with the empirical application, as presented in Table S1, we consider a wider range of $\gamma_1 \in \{0.2, 0.3, 0.4\}$ and $\eta \in \{0.7, 0.8, 0.9\}$ to reflect varying degrees of tail heaviness and extremal dependence. In particular, when $\eta = 0.9$, it approaches 1, corresponding to a setting where the model is close to the tail dependence case; when $\eta = 0.7$, it approaches the lower bound of the admissible region, *i.e.*, $(7+\sqrt{17})/16 \approx 0.695$. We consider a sample size of $n = 2000$, and the choices of k, k_1, k_2 follow as Section 3. Similar empirical findings are obtained for other sample sizes, and thus are omitted here for brevity. We report the values of MSRE in Tables S2 - S4.

The main observations from Tables S2 - S4 are given as follows: all extrapolative estimators consistently outperform the three benchmark methods in terms of MSRE, even when η approaches its lower bound, demonstrating the superiority of the proposed methods; the MSREs of these extrapolative estimators exhibit a high consistency across different parameter settings (γ_1 and η), demonstrating the parametric robustness; as the tail becomes heavier (γ_1 becomes larger), the MSREs of these methods tends to increase, while no clear pattern is observed with respect to changes in η; when $\eta = 0.9$, the model approaches a tail-dependent setting, in which case CoVaR-TD method surprisingly performs better than CoVaR-WE much.

S8 Additional empirical study

S8.1 Additional figures for empirical Study

This section includes some additional figures of Section 4. Figure S4 plots the estimations of γ_1 and η against k_1 and k_2 for choosing suitable values of k_1 and k_2. Figures S5 - S6 plot the estimations $\widetilde{\text{CoVaR}}^{(1)}_{X|Y}(\tau_n')$, $\widetilde{\text{CoVaR}}^{(2)}_{X|Y}(\tau_n')$ and $\widetilde{\text{CoES}}^{(3)}_{X|Y}(\tau_n')$ against k with $\tau_n' = 0.99$ and 0.999 for choosing a suitable k.

S8.2 Tail Quotient Correlation Coefficient (TQCC) test

In this section, we implement the tail quotient correlation coefficient (TQCC) test proposed by Zhang et al. (2017) to examine the null hypothesis of

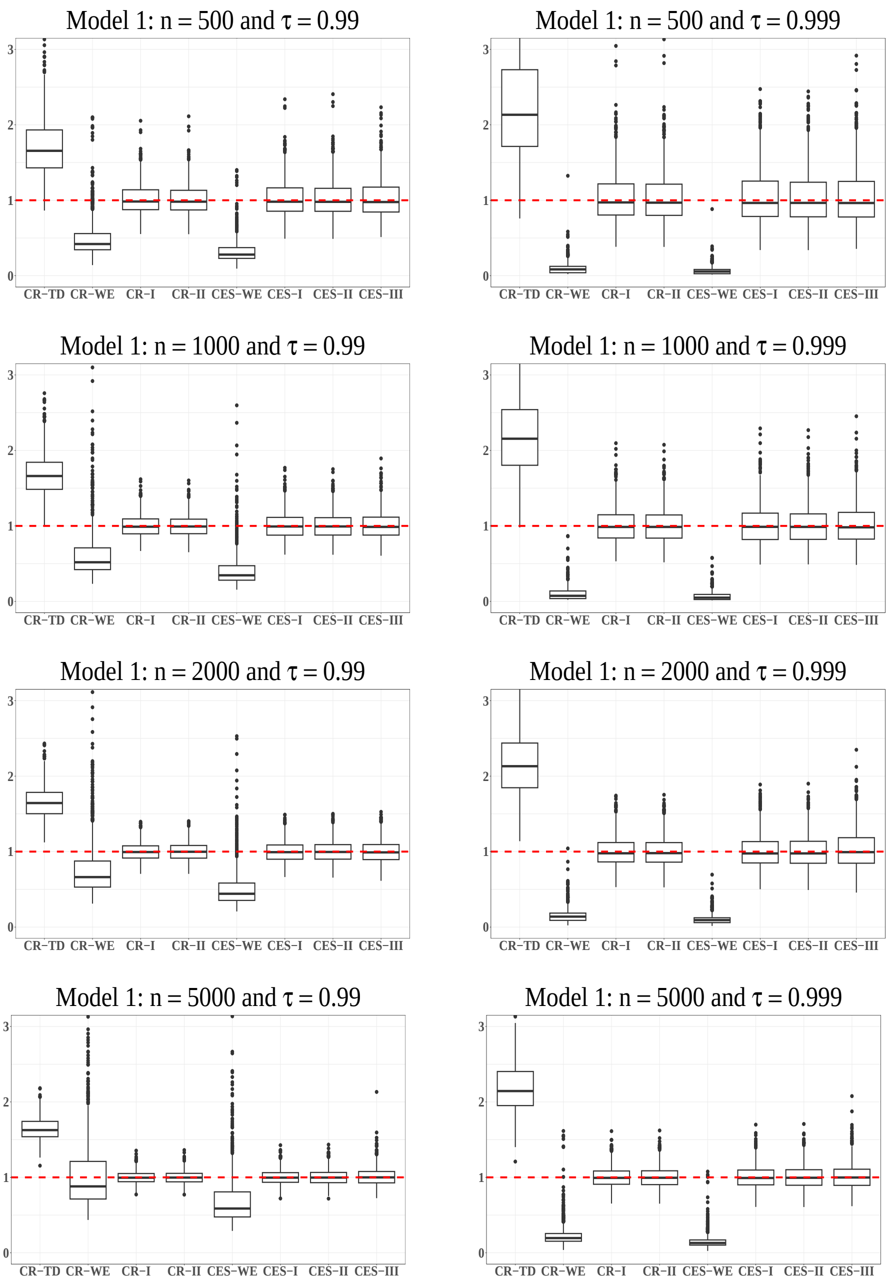

Figure S1: Boxplots of the ratios between CoVaR-TD, CoVaR-WE, CoVaR-I, CoVaR-II, CoES-WE, CoES-I, CoES-II, CoES-III and their true values for **Model 1** with $n \in \{500, 1000, 2000, 5000\}$ and $\tau_n' \in \{0.99, 0.999\}$, where "CR", "CES" denote "CoVaR", "CoES", respectively).

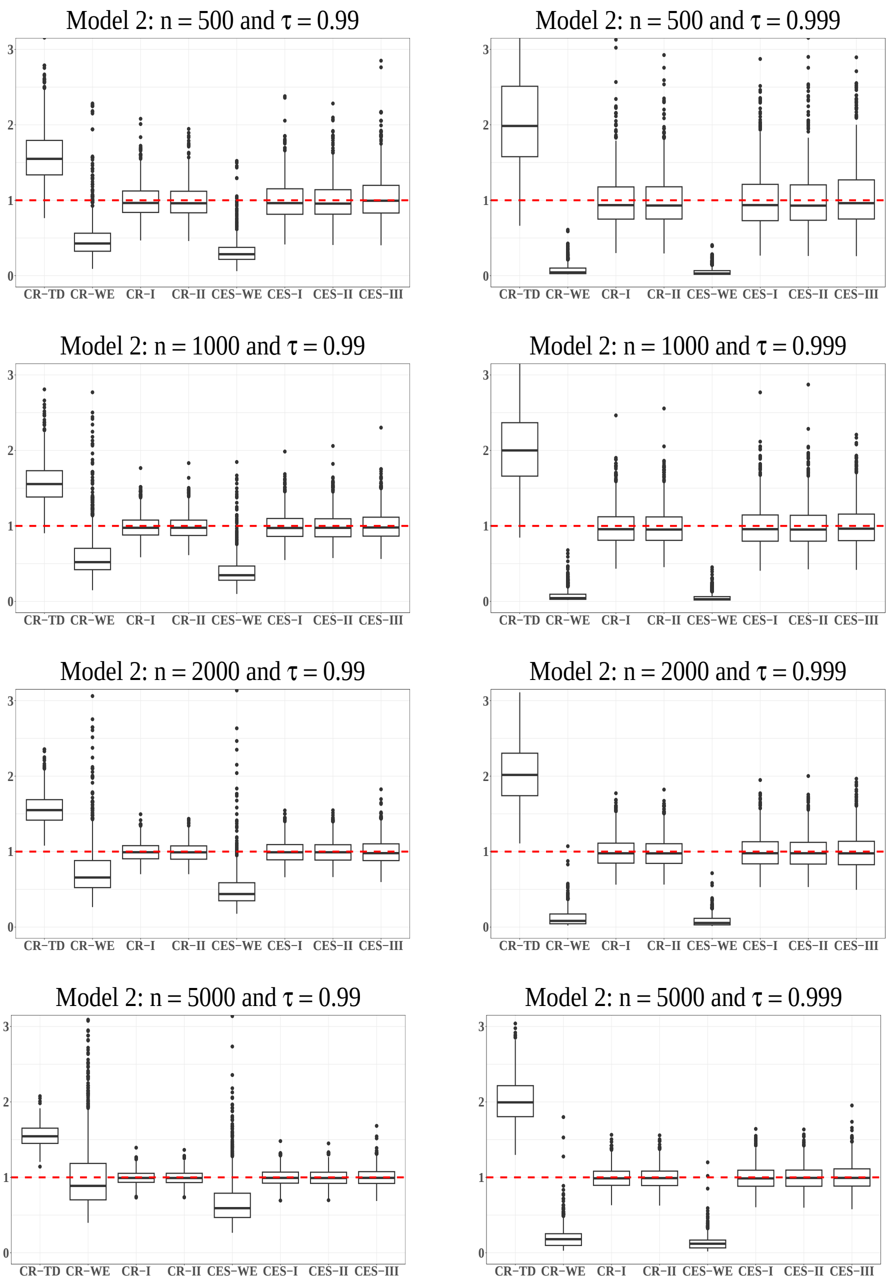

Figure S2: Boxplots of the ratios between CoVaR-TD, CoVaR-WE, CoVaR-I, CoVaR-II, CoES-WE, CoES-I, CoES-II, CoES-III and their true values for **Model 2** with $n \in \{500, 1000, 2000, 5000\}$ and $\tau_n' \in \{0.99, 0.999\}$, where "CR", "CES" denote "CoVaR", "CoES", respectively).

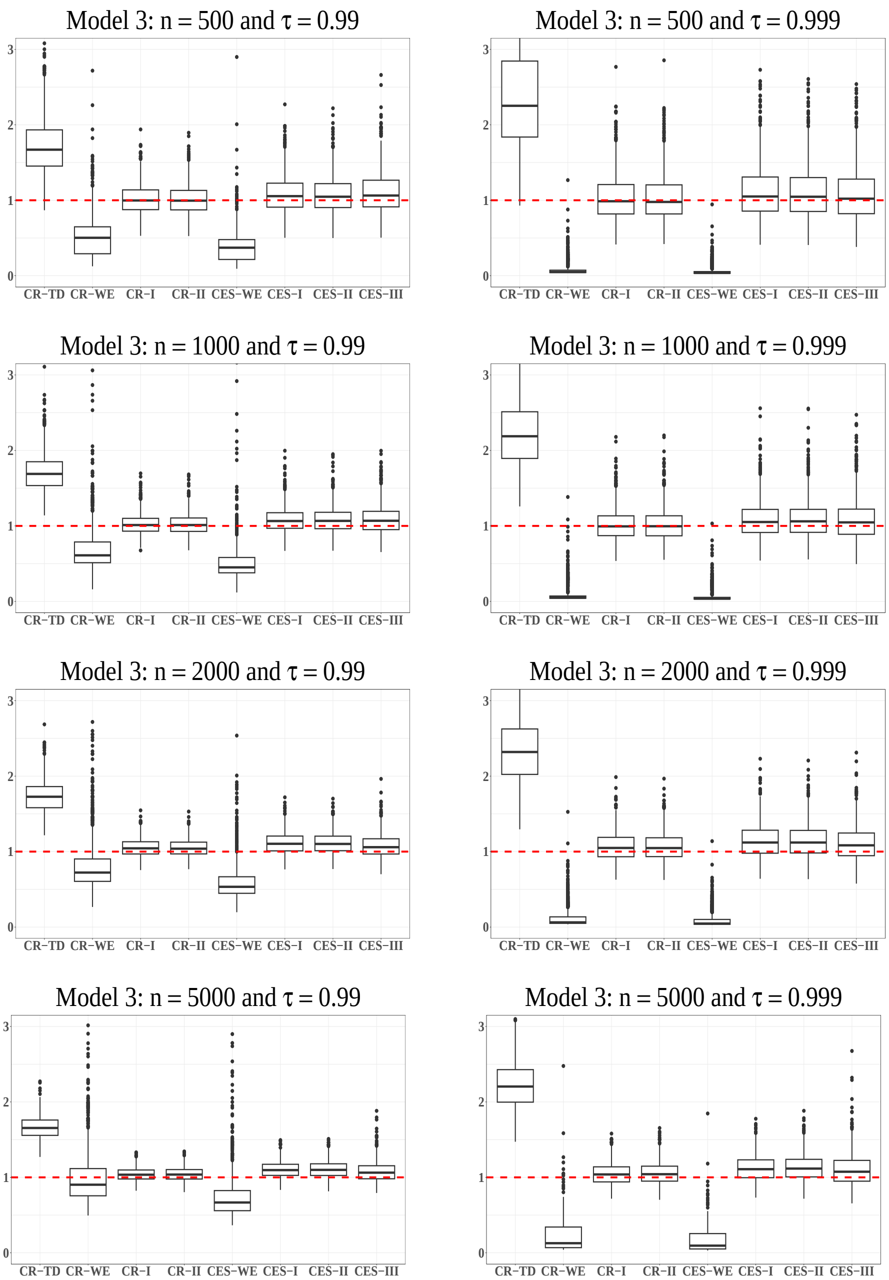

Figure S3: Boxplots of the ratios between CoVaR-TD, CoVaR-WE, CoVaR-I, CoVaR-II, CoES-WE, CoES-I, CoES-II, CoES-III and their true values for **Model 3** with $n \in \{500, 1000, 2000, 5000\}$ and $\tau_n' \in \{0.99, 0.999\}$, where "CR", "CES" denote "CoVaR", "CoES", respectively).

Table S1: Parameter specification for robustness analyses.

		Model 1	**Model 2**	**Model 3**
		$a = 5$		
	$\eta = 0.9$	$a_1 = 17/18$, $a_2 = 8/9$	$a_1 = a_2 = 8/9$	$b = 50/9$
$\gamma_1 = 0.2$	$\eta = 0.8$	$a_1 = 7/8$, $a_2 = 3/4$	$a_1 = a_2 = 3/4$	$b = 25/4$
	$\eta = 0.7$	$a_1 = 11/14$, $a_2 = 4/7$	$a_1 = a_2 = 4/7$	$b = 50/7$
		$a = 10/3$		
	$\eta = 0.9$	$a_1 = 17/18$, $a_2 = 8/9$	$a_1 = a_2 = 8/9$	$b = 100/27$
$\gamma_1 = 0.3$	$\eta = 0.8$	$a_1 = 7/8$, $a_2 = 3/4$	$a_1 = a_2 = 3/4$	$b = 25/6$
	$\eta = 0.7$	$a_1 = 11/14$, $a_2 = 4/7$	$a_1 = a_2 = 4/7$	$b = 100/21$
		$a = 5/2$		
	$\eta = 0.9$	$a_1 = 17/18$, $a_2 = 8/9$	$a_1 = a_2 = 8/9$	$b = 25/9$
$\gamma_1 = 0.4$	$\eta = 0.8$	$a_1 = 7/8$, $a_2 = 3/4$	$a_1 = a_2 = 3/4$	$b = 25/8$
	$\eta = 0.7$	$a_1 = 11/14$, $a_2 = 4/7$	$a_1 = a_2 = 4/7$	$b = 25/7$

Table S2: The MSREs of the estimators for $\text{CoVaR}_{X|Y}(\tau_n')$ and $\text{CoES}_{X|Y}(\tau_n')$ under **Models 1-3** with $\gamma_1 = 0.2$ and $n = 2000$.

η	**Models**	(k, k_1)	CoVaR-TD	CoVaR-WE	CoVaR-I	CoVaR-II	CoES-WE	CoES-I	CoES-II	CoES-III
					$\tau_n' = 0.99$					
	Model 1	(400,384)	0.02003	0.11009	0.00598	0.00599	0.15935	0.00812	0.00811	0.00851
0.9	**Model 2**	(337,384)	0.01937	0.09998	0.00613	0.00629	0.15500	0.00823	0.00839	0.00911
	Model 3	(400,353)	0.06417	0.07634	0.00603	0.00595	0.12722	0.00783	0.00769	0.00814
	Model 1	(337,384)	0.07540	0.09965	0.00522	0.00531	0.15361	0.00704	0.00713	0.00783
0.8	**Model 2**	(368,384)	0.07546	0.07611	0.00558	0.00565	0.14367	0.00748	0.00754	0.00831
	Model 3	(274,400)	0.11801	0.06290	0.00490	0.00489	0.10732	0.00790	0.00783	0.00809
	Model 1	(400,400)	0.22626	0.08149	0.00471	0.00477	0.14498	0.00630	0.00634	0.00694
0.7	**Model 2**	(384,353)	0.21087	0.09112	0.00564	0.00560	0.14984	0.00739	0.00732	0.00796
	Model 3	(179,400)	0.17570	0.05796	0.00550	0.00594	0.09743	0.01036	0.01094	0.01837
					$\tau_n' = 0.999$					
	Model 1	(400,384)	0.05307	0.46138	0.01544	0.01537	0.54891	0.01889	0.01879	0.01925
0.9	**Model 2**	(337,384)	0.05137	0.49045	0.01579	0.01592	0.57375	0.01913	0.01925	0.02017
	Model 3	(242,337)	0.13428	0.53122	0.01773	0.01795	0.59901	0.01978	0.01998	0.02091
	Model 1	(337,384)	0.20678	0.48739	0.01337	0.01346	0.57226	0.01630	0.01638	0.01722
0.8	**Model 2**	(368,384)	0.21228	0.55862	0.01446	0.01450	0.63308	0.01750	0.01753	0.01851
	Model 3	(274,400)	0.31428	0.55438	0.01222	0.01219	0.61176	0.01618	0.01607	0.01641
	Model 1	(400,400)	0.72357	0.49749	0.01226	0.01231	0.58090	0.01482	0.01486	0.01558
0.7	**Model 2**	(211,384)	0.63708	0.58364	0.01466	0.01475	0.65411	0.01731	0.01738	0.02039
	Model 3	(179,400)	0.63550	0.55666	0.01523	0.01588	0.60937	0.02548	0.02634	0.03686

Table S3: The MSREs of the estimators for $\text{CoVaR}_{X|Y}(\tau_n')$ and $\text{CoES}_{X|Y}(\tau_n')$ under **Models 1-3** with $\gamma_1 = 0.3$ and $n = 2000$.

η	**Models**	(k, k_1)	CoVaR-TD	CoVaR-WE	CoVaR-I	CoVaR-II	CoES-WE	CoES-I	CoES-II	CoES-III
$\tau_n' = 0.99$										
	Model 1	(400,384)	0.04975	0.28657	0.01362	0.01363	0.30969	0.01942	0.01936	0.02113
0.9	**Model 2**	(337,384)	0.04800	0.24458	0.01381	0.01413	0.29459	0.01941	0.01971	0.02252
	Model 3	(400,353)	0.16625	0.17773	0.01340	0.01322	0.24328	0.01857	0.01823	0.01926
	Model 1	(337,384)	0.19689	0.24849	0.01188	0.01207	0.29433	0.01678	0.01695	0.01935
0.8	**Model 2**	(368,384)	0.19712	0.15887	0.01253	0.01267	0.26365	0.01758	0.01769	0.02007
	Model 3	(274,400)	0.31565	0.14248	0.01129	0.01126	0.20684	0.02006	0.01985	0.02003
	Model 1	(400,400)	0.63812	0.18253	0.01068	0.01081	0.27038	0.01493	0.01503	0.01685
0.7	**Model 2**	(384,384)	0.59296	0.20843	0.01185	0.01196	0.27904	0.01625	0.01633	0.01836
	Model 3	(179,400)	0.48371	0.12888	0.01299	0.01408	0.18821	0.02680	0.02835	0.03833
$\tau_n' = 0.999$										
	Model 1	(384,384)	0.13978	0.65624	0.03543	0.03540	0.74717	0.04503	0.04493	0.04758
0.9	**Model 2**	(337,384)	0.13540	0.67880	0.03581	0.03603	0.76420	0.04495	0.04513	0.04857
	Model 3	(242,337)	0.37356	0.70105	0.03795	0.03840	0.77138	0.04431	0.04470	0.04842
	Model 1	(337,384)	0.59266	0.67926	0.03085	0.03102	0.76619	0.03894	0.03909	0.04176
0.8	**Model 2**	(368,384)	0.61023	0.73905	0.03264	0.03271	0.81025	0.04089	0.04092	0.04378
	Model 3	(274,400)	0.93042	0.72302	0.02804	0.02795	0.78025	0.04054	0.04022	0.04013
	Model 1	(400,400)	2.37639	0.68970	0.02806	0.02819	0.77354	0.03509	0.03519	0.03728
0.7	**Model 2**	(384,384)	2.28934	0.76330	0.03124	0.03131	0.82695	0.03859	0.03863	0.04068
	Model 3	(179,400)	2.03807	0.73758	0.03739	0.03913	0.78714	0.06827	0.07073	0.08340

Table S4: The MSREs of the estimators for $\text{CoVaR}_{X|Y}(\tau_n')$ and $\text{CoES}_{X|Y}(\tau_n')$ under **Models 1-3** with $\gamma_1 = 0.4$ and $n = 2000$.

η	**Models**	(k, k_1)	CoVaR-TD	CoVaR-WE	CoVaR-I	CoVaR-II	CoES-WE	CoES-I	CoES-II	CoES-III
					$\tau_n' = 0.99$					
0.9	**Model 1**	(400,384)	0.09796	0.69324	0.02467	0.02463	0.50339	0.03760	0.03739	0.04484
	Model 2	(337,384)	0.09428	0.54999	0.02471	0.02523	0.46305	0.03700	0.03747	0.04778
	Model 3	(274,400)	0.31652	0.36832	0.02224	0.02278	0.38073	0.03140	0.03172	0.03900
0.8	**Model 1**	(337,384)	0.40837	0.56466	0.02146	0.02179	0.46494	0.03228	0.03257	0.04032
	Model 2	(368,384)	0.40898	0.27746	0.02233	0.02256	0.38823	0.03338	0.03353	0.04013
	Model 3	(274,400)	0.67120	0.27657	0.02063	0.02056	0.32253	0.04139	0.04090	0.04230
0.7	**Model 1**	(400,400)	1.43538	0.35790	0.01918	0.01942	0.40820	0.02853	0.02874	0.03392
	Model 2	(384,384)	1.32569	0.42427	0.02121	0.02141	0.42359	0.03095	0.03110	0.03738
	Model 3	(179,400)	1.06023	0.24179	0.02431	0.02646	0.29311	0.05593	0.05924	0.07489
					$\tau_n' = 0.999$					
0.9	**Model 1**	(384,384)	0.29481	0.78228	0.06569	0.06550	0.86103	0.08824	0.08784	0.09752
	Model 2	(337,384)	0.28439	0.79700	0.06501	0.06527	0.87110	0.08597	0.08613	0.09848
	Model 3	(242,337)	0.83088	0.80403	0.06497	0.06571	0.86784	0.08110	0.08174	0.09133
0.8	**Model 1**	(337,384)	1.36049	0.79964	0.05683	0.05711	0.87494	0.07543	0.07567	0.08407
	Model 2	(368,384)	1.40544	0.84249	0.05889	0.05898	0.90183	0.07763	0.07760	0.08518
	Model 3	(274,400)	2.21007	0.81863	0.05138	0.05116	0.87024	0.08312	0.08236	0.08195
0.7	**Model 1**	(400,400)	6.32171	0.80898	0.05124	0.05151	0.87995	0.06725	0.06748	0.07317
	Model 2	(384,384)	6.06671	0.86101	0.05686	0.05699	0.91157	0.07397	0.07404	0.07990
	Model 3	(179,400)	5.27668	0.83737	0.07299	0.07667	0.87905	0.14836	0.15406	0.16861

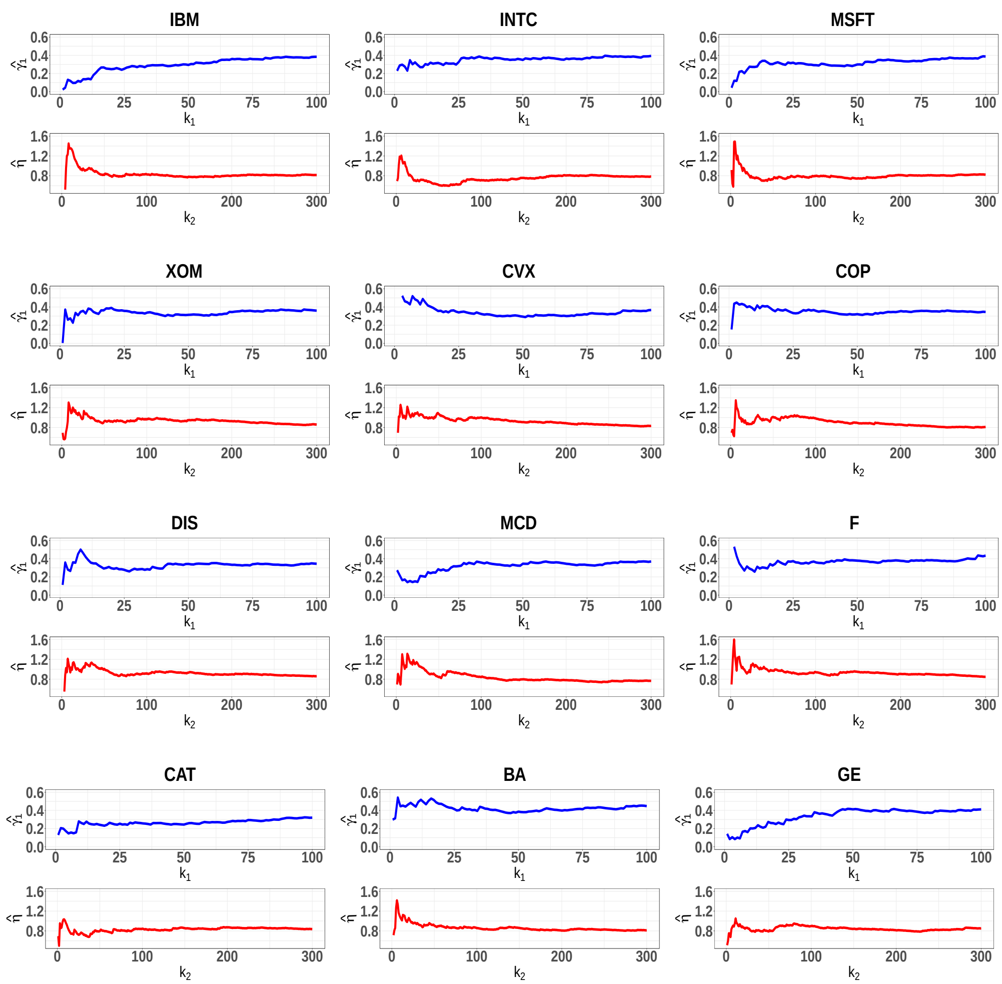

Figure S4: The estimations for γ_1 (blue lines) and η (red lines) against k_1 and k_2 for 12 individual stocks losses conditional on S&P500 Index loss.

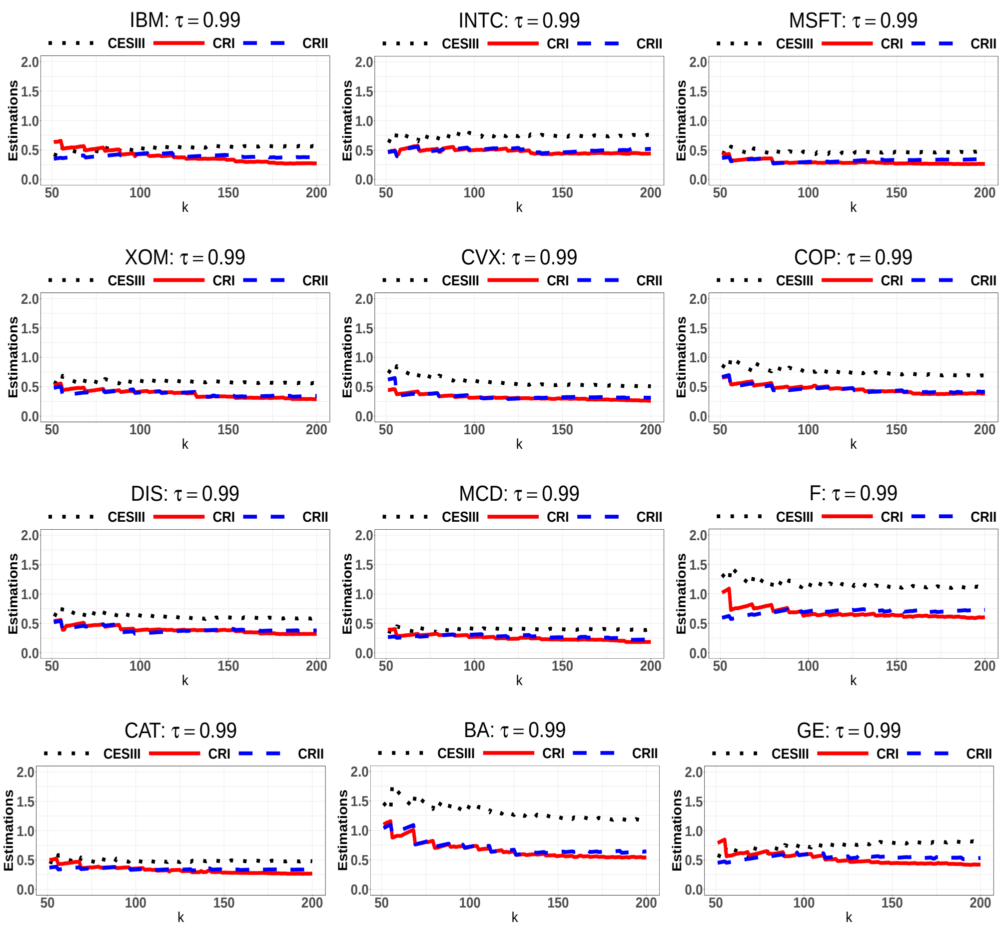

Figure S5: The estimations $\widetilde{\text{CoVaR}}^{(1)}_{X|Y}(\tau_n')$ (CRI in red solid lines), $\widetilde{\text{CoVaR}}^{(2)}_{X|Y}(\tau_n')$ (CRII in blue dashed lines) and $\widetilde{\text{CoES}}^{(3)}_{X|Y}(\tau_n')$ (CESIII in black dotted lines) against k for the 12 individual stocks conditional on S&P500 Index, with $\tau_n' = 0.99$.

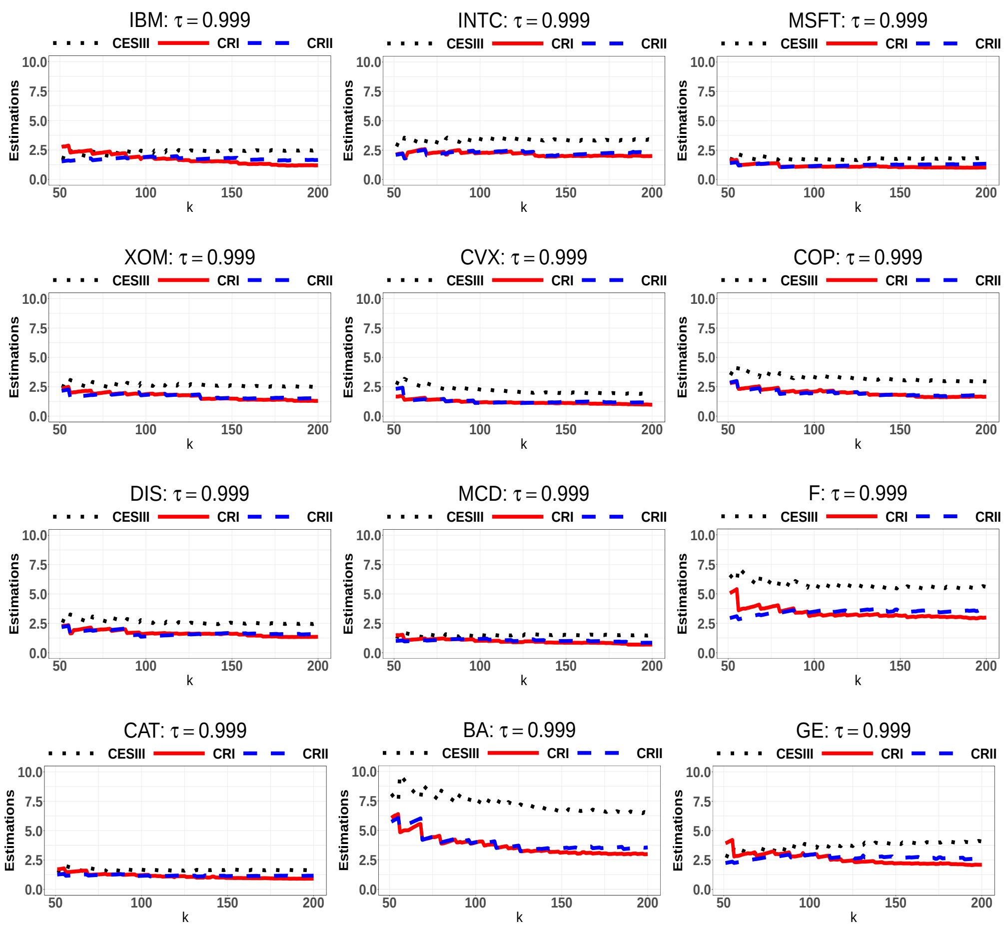

Figure S6: The estimations $\widetilde{\text{CoVaR}}^{(1)}_{X|Y}(\tau_n')$ (CRI in red solid lines), $\widetilde{\text{CoVaR}}^{(2)}_{X|Y}(\tau_n')$ (CRII in blue dashed lines) and $\widetilde{\text{CoES}}^{(3)}_{X|Y}(\tau_n')$ (CESIII in black dotted lines) against k for the 12 individual stocks conditional on S&P500 Index, with $\tau_n' = 0.999$.

asymptotic independence. The detailed steps of the test are as follows:

- Given samples $\{X_i, Y_i\}_{i=1}^n$, fit generalized extreme value distributions for X and Y, respectively,

$$H(x; \zeta, \mu, \sigma) = \exp\left\{-\left[1 + \frac{\zeta(x-\mu)}{\sigma}\right]_+^{-1/\zeta}\right\},$$

  where μ is a location parameter, $\sigma > 0$ is a scale parameter, and ζ is a shape parameter. Denote by $\hat{\zeta}_X, \hat{\mu}_X, \hat{\sigma}_X$ and $\hat{\zeta}_Y, \hat{\mu}_Y, \hat{\sigma}_Y$ the corresponding estimators of ζ, μ, σ for X and Y, respectively.

- Perform marginal transformations,

$$\hat{X}_i = -1/\log\left(H(X_i, \hat{\zeta}_X, \hat{\mu}_X, \hat{\sigma}_X)\right), \text{ and } \hat{Y}_i = -1/\log\left(H(Y_i, \hat{\zeta}_Y, \hat{\mu}_Y, \hat{\sigma}_Y)\right).$$

- Let u_n be a random threshold. Calculate the TQCC statistic,

$$\mathrm{TQCC}_n = \frac{\max_{1\le i\le n}\left\{\frac{\max\{\hat{X}_i, u_n\}}{\max\{\hat{Y}_i, u_n\}}\right\} + \max_{1\le i\le n}\left\{\frac{\max\{\hat{Y}_i, u_n\}}{\max\{\hat{X}_i, u_n\}}\right\} - 2}{\max_{1\le i\le n}\left\{\frac{\max\{\hat{X}_i, u_n\}}{\max\{\hat{Y}_i, u_n\}}\right\} \times \max_{1\le i\le n}\left\{\frac{\max\{\hat{Y}_i, u_n\}}{\max\{\hat{X}_i, u_n\}}\right\} - 1}.$$

  Following Zhang et al. (2017), we choose the smaller one of two empirical 95% percentiles of $\{\hat{X}_i\}$ and $\{\hat{Y}_i\}$ as u_n.

- Calculate p-value by $\mathbb{P}\left(2n\{1 - \exp\{-1/u_n\}\}\mathrm{TQCC}_n > \chi_4^2\right)$.

We report the values of TQCC statistics and p-values in Table S5. Except for F, BA, and GE, the p-values for the remaining stocks are all close to 1, indicating significant tail independence between these stocks and

Table S5: Summary of TQCC statistics and p-values for 12 individual stocks.

	IBM	INTC	MSFT
TQCC	0.00042	0.95250×10^{-9}	4.07138×10^{-5}
p-values	0.99802	1.00000	0.99998
	XOM	CVX	COP
TQCC	0.00231	0.00799	0.00017
p-values	0.94977	0.63946	0.99959
	DIS	MCD	F
TQCC	6.36048×10^{-5}	0.000613	0.99152
p-values	0.99994	0.99580	0.00000
	CAT	BA	GE
TQCC	8.50098×10^{-6}	0.94398	0.17312
p-values	0.99999	0.00000	0.00000

the S&P 500 Index. Although there is no sufficient evidence to support tail independence between F, BA, GE and the S&P 500 Index, we still include these three stocks as comparative counterexamples in the empirical study to examine the empirical performance of the proposed methods.

S8.3 Bootstrapped confidence intervals

In this section, we implement the bootstrap-based confidence intervals proposed in Section S6 for empirical inference. Specifically, for the sample size of $n = 1565$, we set $B = 1000$, $l = 17$, and $m = 92$. We report the confidence intervals for the five extrapolative estimators in Tables S6 and S7 with $\tau_n' = 0.99$ and 0.999, respectively. As the level τ_n' becomes more extreme, the empirical performance of the corresponding confidence intervals deteriorates, with substantially larger interval lengths observed, particularly for the F, BA, and GE three stocks, which fail to pass the TQCC test for tail independence. This is intuitive, because more extreme levels make it harder to estimate accurately.

S8.4 Robustness analysis w.r.t k, k_1 and k_2

In this section, we conduct an analysis to demonstrate the robustness of our estimators with respect to the intermediate orders k, k_1 and k_2. Since $\widetilde{\text{CoES}}^{(j)}_{X|Y}(\tau_n')$ $(j = 1, 2)$ differ from $\widetilde{\text{CoVaR}}^{(j)}_{X|Y}(\tau_n')$ $(j = 1, 2)$ only by a $\hat{\gamma}_1$-dependent term, we focus on $\widetilde{\text{CoVaR}}^{(1)}_{X|Y}(\tau_n')$, $\widetilde{\text{CoVaR}}^{(2)}_{X|Y}(\tau_n')$ and $\widetilde{\text{CoES}}^{(3)}_{X|Y}(\tau_n')$ as representative examples in this analysis. Specifically, we adopt a control variable approach by fixing two of (k,k_1,k_2) while allowing the remaining one to vary over a range of plausible values, and then plot the corresponding estimates against k, k_1, or k_2, as shown in Figures S7 - S9. One can observe that all these estimators tend to stabilize as k, k_1, k_2 increase, except for F, BA, and GE, which fail to pass the TQCC test for

Table S6: Bootstrap-based 95% confidence intervals for the estimators of $\mathrm{CoVaR}_{X|Y}(\tau_n')$ and $\mathrm{CoES}_{X|Y}(\tau_n')$, with $\tau_n' = 0.99$.

	Est/CI	CoVaR-I	CoVaR-II	CoES-I	CoES-II	CoES-III
Information Technology Sector						
IBM	Est	0.37879	0.41245	0.58989	0.64230	0.56368
	CI	[0.18365,0.77923]	[0.23508,0.81078]	[0.25847,1.31281]	[0.33213,1.35906]	[0.32693,1.04111]
INTC	Est	0.50422	0.53746	0.80272	0.85564	0.73419
	CI	[0.30118,0.90864]	[0.33315,0.97427]	[0.42730,1.63669]	[0.46730,1.76219]	[0.43850,1.33377]
MSFT	Est	0.28707	0.30003	0.43175	0.45124	0.45597
	CI	[0.14396,0.41476]	[0.15930,0.44128]	[0.19493,0.67331]	[0.21465,0.71036]	[0.26025,0.75510]
Energy Sector						
XOM	Est	0.39466	0.39174	0.61186	0.60734	0.60774
	CI	[0.21100,0.83955]	[0.20736,0.79400]	[0.28859,1.47727]	[0.28345,1.40439]	[0.33241,1.26991]
CVX	Est	0.31093	0.29959	0.45566	0.43905	0.53785
	CI	[0.14007,0.53571]	[0.12668,0.45693]	[0.17412,0.88304]	[0.15653,0.75622]	[0.22448,1.14991]
COP	Est	0.46190	0.46588	0.71945	0.72564	0.77490
	CI	[0.22769,1.04294]	[0.24770,1.00684]	[0.31614,1.84041]	[0.33773,1.76236]	[0.37906,1.78182]
Consumer Discretionary Sector						
DIS	Est	0.38434	0.37703	0.58180	0.57075	0.58671
	CI	[0.19644,0.76629]	[0.21749,0.65951]	[0.26297,1.31125]	[0.29187,1.11842]	[0.29151,1.16009]
MCD	Est	0.26894	0.30968	0.40198	0.46287	0.40241
	CI	[0.11304,0.52089]	[0.18340,0.66104]	[0.14118,0.85742]	[0.22926,1.08295]	[0.19562,0.72103]
F	Est	0.63425	0.67251	1.02575	1.08764	1.18816
	CI	[0.28099,1.24068]	[0.36768,1.29802]	[0.40044,2.32754]	[0.52139,2.43072]	[0.52138,2.64594]
Industrials Sector						
CAT	Est	0.32756	0.34397	0.46409	0.48734	0.49316
	CI	[0.20997,0.54717]	[0.21228,0.53484]	[0.27928,0.82906]	[0.28063,0.81280]	[0.31291,0.79128]
BA	Est	0.59652	0.60089	1.01658	1.02404	1.25659
	CI	[0.20570,1.30717]	[0.19593,1.16672]	[0.26606,2.67920]	[0.25856,2.34174]	[0.42709,3.55283]
GE	Est	0.59528	0.59290	0.96896	0.96508	0.75045
	CI	[0.27940,1.42382]	[0.36797,1.31820]	[0.38678,2.68704]	[0.49903,2.46401]	[0.42399,1.40510]

Table S7: Bootstrap-based 95% confidence intervals for the estimators of $\text{CoVaR}_{X|Y}(\tau_n')$ and $\text{CoES}_{X|Y}(\tau_n')$, with $\tau_n' = 0.999$.

	Est/CI	CoVaR-I	CoVaR-II	CoES-I	CoES-II	CoES-III
			Information Technology Sector			
IBM	Est	1.65912	1.80652	2.58371	2.81326	2.46895
	CI	[0.55791,5.21356]	[0.71498,5.52397]	[0.79285,8.83392]	[0.99925,9.17402]	[1.00384,7.17319]
INTC	Est	2.27626	2.42634	3.62382	3.86274	3.31445
	CI	[0.97289,5.92674]	[1.09197,6.35072]	[1.34427,10.60115]	[1.50367,11.62269]	[1.41027,8.18503]
MSFT	Est	1.10047	1.15016	1.65509	1.72982	1.74794
	CI	[0.40126,2.18698]	[0.43832,2.28433]	[0.54715,3.51197]	[0.59969,3.77033]	[0.70268,3.78376]
			Energy Sector			
XOM	Est	1.76799	1.75493	2.74102	2.72078	2.72254
	CI	[0.64217,5.58788]	[0.64107,5.30229]	[0.86084,9.92633]	[0.88210,9.33755]	[1.02706,8.39891]
CVX	Est	1.15907	1.11682	1.69861	1.63669	2.00500
	CI	[0.32209,2.93316]	[0.29352,2.48791]	[0.40339,4.75858]	[0.36894,4.08254]	[0.53855,6.00204]
COP	Est	1.97902	1.99605	3.08247	3.10900	3.32006
	CI	[0.66394,7.07762]	[0.68967,6.78198]	[0.91818,12.46391]	[0.93758,12.11508]	[1.09493,12.04326]
			Consumer Discretionary Sector			
DIS	Est	1.61683	1.58610	2.44753	2.40102	2.46819
	CI	[0.55370,4.72491]	[0.61900,4.04216]	[0.73341,8.06896]	[0.82413,6.88392]	[0.81307,7.06446]
MCD	Est	1.01836	1.17262	1.52214	1.75271	1.52381
	CI	[0.24753,2.85535]	[0.42112,3.58969]	[0.31310,4.70395]	[0.53391,5.93175]	[0.43191,3.90320]
F	Est	3.14479	3.33453	5.08597	5.39283	5.89127
	CI	[0.91450,9.51896]	[1.20489,9.85344]	[1.29184,17.92137]	[1.69515,18.44516]	[1.65133,19.71296]
			Industrials Sector			
CAT	Est	1.11689	1.17285	1.58243	1.66170	1.68154
	CI	[0.57048,2.42338]	[0.57747,2.37598]	[0.74438,3.71242]	[0.77688,3.60787]	[0.86468,3.48884]
BA	Est	3.29230	3.31645	5.61073	5.65188	6.93541
	CI	[0.59144,12.46844]	[0.56976,10.82976]	[0.74917,25.36994]	[0.74477,22.34941]	[1.22089,33.45807]
GE	Est	2.95330	2.94148	4.80717	4.78793	3.72307
	CI	[0.88758,11.39808]	[1.16741 10.32110]	[1.20937,21.26723]	[1.58694,19.27782]	[1.38385,10.66487]

tail independence. Moreover, the plots of $\widetilde{\mathrm{CoVaR}}^{(j)}_{X|Y}(\tau'_n)$ $(j = 1, 2)$ almost overlap, indicating that the two extrapolative estimators exhibit very similar empirical performance. Moreover, the values of (k, k_1, k_2) we used in the empirical study remain reasonable, since they all make the estimations perform stable.

S8.5 Comparison with tail-dependence methods

In this section, we include an empirical comparison with $\mathrm{CoVaR}_{X|Y}(\tau'_n)$ and $\mathrm{CoES}_{X|Y}(\tau'_n)$ estimators developed under a framework of tail dependence. For estimating $\mathrm{CoVaR}_{X|Y}(\tau'_n)$, we also employ the semi-parametric method "CoVaR-TD" proposed by Nolde et al. (2022), which serves as a benchmark in Section 3. For estimating $\mathrm{CoES}_{X|Y}(\tau'_n)$, we define an estimator "CoES-TD" by scaling "CoVaR-TD" with the term $\frac{1}{1-\hat{\gamma}_1}$. To objectively evaluate the empirical performance of the competing methods, we conduct comparisons based on two types of scoring functions. The first is the classical scoring function for quantiles, which is also used in Nolde et al. (2022) for CoVaR evaluation,

$$S(r, x) = (1 - \tau - I\{x > r\})G(r) + I\{x > r\}G(x),$$

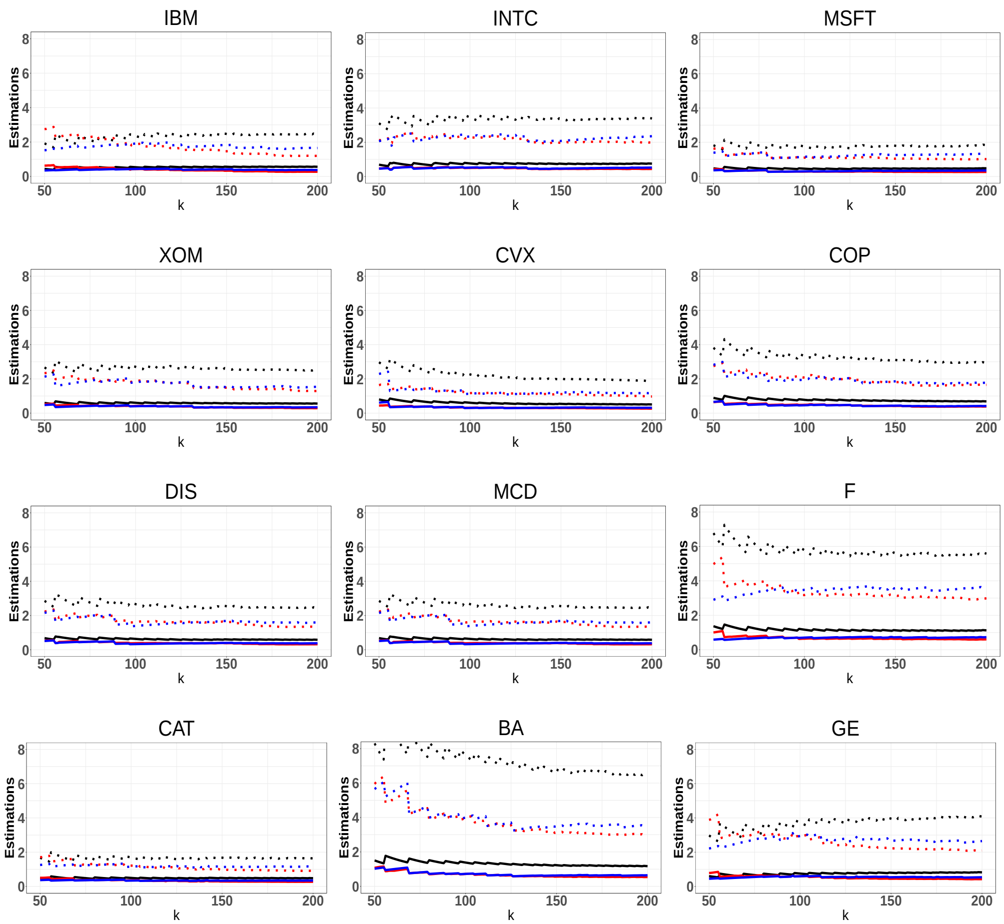

Figure S7: The values of $\widetilde{\mathrm{CoVaR}}^{(1)}_{X|Y}(\tau_n')$ (in red), $\widetilde{\mathrm{CoVaR}}^{(2)}_{X|Y}(\tau_n')$ (in blue) and $\widetilde{\mathrm{CoES}}^{(3)}_{X|Y}(\tau_n')$ (in black) against k, with $\tau_n' = 0.99$ (solid lines) and 0.999 (dotted lines) for 12 individual stocks conditional on S&P500 Index.

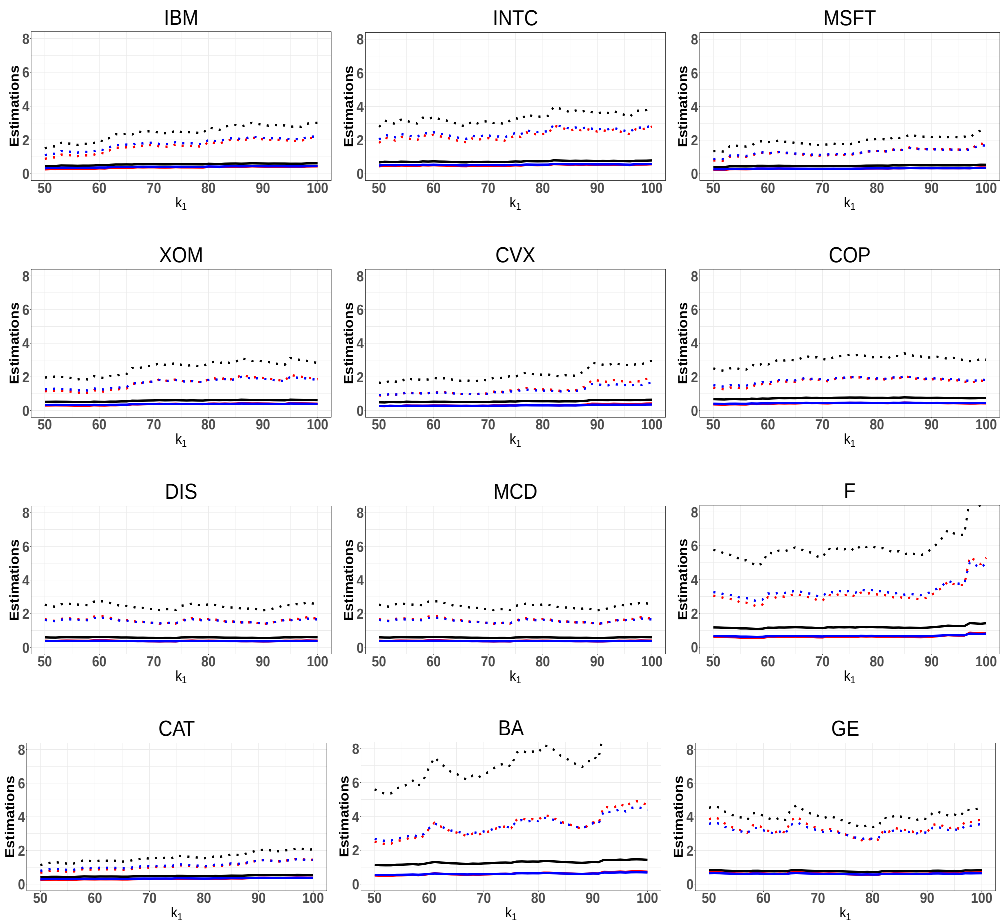

Figure S8: The values of $\widetilde{\text{CoVaR}}^{(1)}_{X|Y}(\tau_n')$ (in red), $\widetilde{\text{CoVaR}}^{(2)}_{X|Y}(\tau_n')$ (in blue) and $\widetilde{\text{CoES}}^{(3)}_{X|Y}(\tau_n')$ (in black) against k_1, with $\tau_n' = 0.99$ (solid lines) and 0.999 (dotted lines) for 12 individual stocks conditional on S&P500 Index.

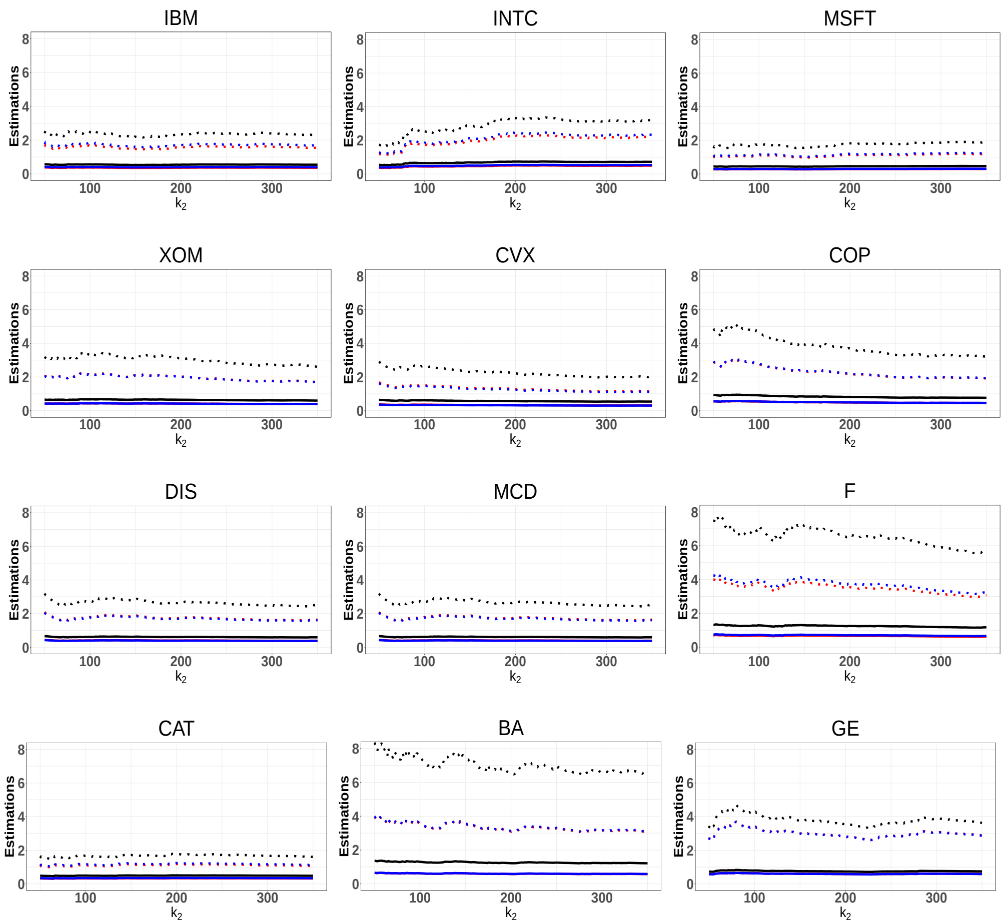

Figure S9: The values of $\widetilde{\mathrm{CoVaR}}^{(1)}_{X|Y}(\tau_n')$ (in red), $\widetilde{\mathrm{CoVaR}}^{(2)}_{X|Y}(\tau_n')$ (in blue) and $\widetilde{\mathrm{CoES}}^{(3)}_{X|Y}(\tau_n')$ (in black) against k_2, with $\tau_n' = 0.99$ (solid lines) and 0.999 (dotted lines) for 12 individual stocks conditional on S&P500 Index.

where $G(\cdot)$ is an increasing function. The second is the bivariate scoring function for pair $(\mathrm{VaR}(\tau), \mathrm{ES}(\tau))$ proposed by Fissler and Ziegel (2016),

$$S(r_1, r_2, x) = I\{x > r_1\}(-G_1(r_1) + G_1(x) - G_2(r_2)(r_1 - x))$$
$$+ (1-\tau)(G_1(r_1) - G_2(r_2)(r_2 - r_1) + \mathcal{G}_2(r_2)),$$

where $G_1(\cdot)$ is an increasing function, $\mathcal{G}_2'(\cdot) = G_2(\cdot)$ and $\mathcal{G}_2(\cdot)$ is increasing and concave. In practice, r, r_1, r_2 denote risk forecasts and x the observation, we let $G(x) = x$, $G_1(x) = 0$, and $\mathcal{G}_2(x) = x^{1/2}$. As conditional extensions of VaR and ES, it is appropriate to use these two scoring functions for the comparisons of CoVaR and CoES.

Moreover, we continue to use the rolling-window approach adopted in Section 4, *that is*, we estimate the risk forecasts for the first trading day of each month and calculate the averages of the scores based on these windows. We report the values of these scores in Table S8. It can be seen that the tail-dependence methods yield substantially higher scores than our methods, even in F, BA and GE, which fail to pass the TQCC test. The scores of the proposed methods are all very close across CoVaR-I, CoVaR-II and CoES-I, CoES-II, CoES-III. This suggests that our methods consistently outperform the tail-dependence-based approaches, while also confirming the robustness across different specifications of our proposed estimators.

Table S8: Scores ($\times 10^{-3}$) of extreme $\text{CoVaR}_{X|Y}(\tau_n')$ and $\text{CoES}_{X|Y}(\tau_n')$ estimates at $\tau_n' = 0.99$ and 0.999 (in brackets).

	CoVaR-TD	CoVaR-I	CoVaR-II	CoES-TD	CoES-I	CoES-II	CoES-III
			Information Technology Sector				
IBM	3.14430 (2.44932)	2.13841 (1.38689)	2.24706 (1.44530)	5.77078 (1.57167)	4.74273 (1.16416)	4.87032 (1.19393)	4.73898 (1.16051)
INTC	3.31405 (2.00512)	2.33788 (1.13644)	2.44999 (1.18776)	5.83506 (1.41395)	4.91623 (1.07047)	5.04236 (1.09713)	4.98456 (1.08447)
MSFT	3.05480 (1.97071)	2.32721 (1.27537)	2.20717 (1.21185)	5.65186 (1.41848)	4.93321 (1.13368)	4.80366 (1.10429)	4.71807 (1.08331)
			Energy Sector				
XOM	2.27287 (1.42551)	1.91556 (1.07369)	1.83834 (1.00877)	4.76232 (1.14663)	4.32784 (0.96710)	4.28593 (0.95415)	4.26649 (0.94923)
CVX	2.16126 (1.21623)	1.81594 (0.91041)	1.77863 (0.88050)	4.68933 (1.09192)	4.24242 (0.918843)	4.23061 (0.91250)	4.27673 (0.92352)
COP	2.82362 (1.74250)	2.33173 (1.27318)	2.50263 (1.37565)	5.32704 (1.29192)	4.75334 (1.05428)	4.91941 (1.09208)	4.85926 (1.07732)
			Consumer Discretionary Sector				
DIS	2.74334 (1.78150)	2.23915 (1.32114)	2.38331 (1.38061)	5.34429 (1.32645)	4.81050 (1.12623)	4.97298 (1.16037)	4.93475 (1.14963)
MCD	1.90143 (1.15421)	1.50603 (0.80706)	1.51416 (0.80149)	4.42045 (1.06122)	3.90838 (0.86129)	3.92813 (0.86339)	3.90935 (0.85980)
F	3.97940 (2.77568)	2.81969 (1.57595)	2.83754 (1.57981)	6.30674 (1.60313)	5.31905 (1.21281)	5.35175 (1.21890)	5.32512 (1.21237)
			Industrials Sector				
CAT	2.43837 (1.25539)	1.91568 (0.86793)	2.02734 (0.91248)	5.00441 (1.11937)	4.41389 (0.91898)	4.54822 (0.94582)	4.50121 (0.93558)
BA	3.04668 (1.96967)	2.45755 (1.37230)	2.60720 (1.43699)	5.56551 (1.36374)	5.00748 (1.14357)	5.17991 (1.18000)	5.19961 (1.18572)
GE	3.79462 (3.37301)	3.29076 (2.64625)	3.02566 (2.29058)	6.13168 (1.64764)	5.69624 (1.44836)	5.52021 (1.38995)	5.37056 (1.34201)

Bibliography

Adler, R.J. (1990). *An Introduction to Continuity, Extrema, and Related Topics for General Gaussian Processes*. Institute of Mathematical Statistics.

Bücher, A. and Dette, H. (2013). Multiplier bootstrap of tail copulas with applications. *Bernoulli* **19**, 1655–1687.

Bühlmann, P. (1994). Blockwise bootstrapped empirical process for stationary sequences. *The Annals of Statistics* **22**, 995–1012.

de Haan, L. and Ferreira, A. (2006). *Extreme Value Theory: An Introduction*. Springer, New York.

de Haan, L. and Zhou, C. (2024). Bootstrapping extreme value estimators. *Journal of the American Statistical Association* **119**, 382–393.

Danielsson, J., de Haan, L., Peng, L. and de Vries, C.G. (2001). Using a bootstrap method to choose the sample fraction in tail index estimation. *Journal of Multivariate analysis* **76**, 226–248.

Draisma, G., de Haan, L., Peng, L. and Pereira, T.T. (1999). A bootstrap-based method to achieve optimality in estimating the extreme-value index. *Extremes* **2**, 367–404.

Fissler, T., and Ziegel, J.F. (2016). Higher order elicitability and Osband's principle. *The Annals of Statistics* **44**, 1680–1707.

Francq, C. and Zakoïan, J.M. (2025). Inference on dynamic systemic risk measures. *Journal of Econometrics* **247**, 105936.

Fung, T.C., Li, Y., Peng, L. and Qian, L. (2026). Statistical inference for systemic risk-driven portfolio selection. *Journal of Econometrics* **253**, 106127.

Gomes, D.P. and Neves, M.M. (2015). Bootstrap and other resampling methodologies in statistics of extremes. *Communications in Statistics-Simulation and Computation* **44**, 2592–2607.

Hall, P., Horowitz, J.L. and Jing, B.Y. (1995). On blocking rules for the bootstrap with dependent data. *Biometrika* **82**, 561–574.

Künsch, H.R. (1989). The jackknife and the bootstrap for general stationary observations. *The Annals of Statistics* **17**, 1217–1241.

Leng, X., He, Y., Hou, Y. and Peng, L. (2024). Asymptotics of CoVaR Inference In Two-Quantile-Regression. *Available at SSRN 4816475.*

Nolde, N., Zhou, C. and Zhou, M. (2022). An extreme value approach to CoVaR estimation. *arXiv preprint arXiv:2201.00892.*

Orey, S. and Pruitt, W.E. (1973). Sample functions of the N-parameter Wiener process. *The Annals of Probability* **1**, 138–163.

Peng, L. and Qi, Y. (2008). Bootstrap approximation of tail dependence function. *Journal of Multivariate Analysis* **99**, 1807–1824.

Van der Vaart, A.W. and Wellner, J.A. (1996). *Weak Convergence and Empirical Processes with Applications to Statistics*. Springer, New York.

Van der Vaart, A.W. (2000). *Asymptotic statistics*. Cambridge University Press.

Zhang, Z., Zhang, C. and Cui, Q. (2017). Random threshold driven tail dependence measures with application to precipitation data analysis. *Statistica Sinica* **27**, 1807–1824.