Journal of Applied Mathematics and Physics, 2026, 14(6), 2250-2272
https://www.scirp.org/journal/jamp
ISSN Online: 2327-4379
ISSN Print: 2327-4352

Simultaneous Confidence Bands for Conditional Risk Measurement and Conditional Expected Loss Based on Generalized Estimators

Jiale Diao

School of Mathematics and Statistics, Southwest University, Chongqing, China
Email: 949241162@qq.com

How to cite this paper: Diao, J.L. (2026) Simultaneous Confidence Bands for Conditional Risk Measurement and Conditional Expected Loss Based on Generalized Estimators. *Journal of Applied Mathematics and Physics*, **14**, 2250-2272.
https://doi.org/10.4236/jamp.2026.146110

Received: May 22, 2026
Accepted: June 19, 2026
Published: June 22, 2026

Abstract

Accurate measurement and valid inference of extreme financial risks remain a fundamental challenge in risk management research. Existing literature on conditional risk measures predominantly focuses on pointwise estimation and confidence intervals for single tail thresholds, failing to support joint statistical inference over the entire extreme tail interval. Meanwhile, traditional asymptotic-based confidence bands suffer from severe finite-sample coverage bias, and the complex limiting Gaussian process under the intermediate asymptotic scenario impedes direct empirical application. Under the location-scale model framework, this paper proposes a multiplier bootstrap approach to construct simultaneous confidence bands (SCBs) for conditional risk measures of Pareto-type heavy-tailed return series. We adopt the generalized moment Hill estimator and select effective sample sizes via a data-driven criterion, with rigorous proof of the bootstrap procedure's asymptotic validity. Monte Carlo simulations confirm that the proposed method is highly robust to the tail index estimator's moment order α, with coverage probabilities converging to the 95% nominal level and confidence band lengths shrinking at the rate governed by the effective tail sample size, consistent with the theoretical $k^{-1/2}$ rate. The method delivers superior coverage accuracy for core relative risk measures, effectively mitigating finite-sample bias in traditional approaches, balancing theoretical rigor and practical tractability for financial risk management applications.

Keywords

Conditional Risk Measures, Simultaneous Confidence Bands, Extreme Value Theory, Multiplier Bootstrap, Heavy Tails

1. Introduction

In financial risk management, accurately measuring the tail risk of asset returns is a core task for investment decisions and risk control. Traditional Value-at-Risk (VaR), defined as the maximum possible loss of an asset at a given confidence level, has been widely adopted by financial institutions since its introduction by J.P. Morgan in 1994. However, VaR suffers from two major drawbacks: first, it is not subadditive and hence not a coherent risk measure; second, it fails to capture extreme losses beyond VaR, *i.e.*, it does not fully utilise information about tail risk. To overcome these deficiencies, Artzner [1] proposed an axiomatic system for coherent risk measures, among which Expected Shortfall (ES, also known as Conditional Value-at-Risk) has become a standard coherent risk measure. In this paper, we investigate both the conditional quantile (traditionally called Value-at-Risk, VaR) and the conditional tail expectation (ES). For consistency with Li *et al.* [2], we denote the conditional quantile by CVaR and the conditional tail expectation by CES; therefore, in our notation, CVaR corresponds to standard VaR and CES corresponds to standard ES/CVaR.

In actual financial time series, returns often exhibit typical features such as volatility clustering, heavy tails and asymmetry. To capture these features, researchers commonly model returns using a location-scale model:

$$R_t = m\left(A_{t-1}, \beta_0\right) + \sigma\left(A_{t-1}, \beta_0\right)\varepsilon_t,$$

where the conditional mean $m(\cdot)$ and conditional standard deviation $\sigma(\cdot)$ depend on the historical information set A_{t-1}, and the innovations ε_t are i.i.d. with zero mean and unit variance. Extensive empirical evidence indicates that the distribution of ε_t is often Pareto-type heavy-tailed, *i.e.*, its tail distribution function satisfies the regular variation condition:

$$\lim_{x\to\infty} \frac{1-F_\varepsilon(xs)}{1-F_\varepsilon(x)} = s^{-1/\gamma_R}, \quad \forall s > 0,$$

where $\gamma_R > 0$ is called the right-tail extreme value index. Similarly, the left-tail extreme value index $\gamma_L > 0$ is defined on the tail of $-\varepsilon_t$. The magnitude of the extreme value index directly determines the thickness of the tail and thereby affects the estimation accuracy of CVaR and CES.

Extreme value theory (EVT) based tail estimation methods provide effective tools for estimating CVaR and CES. McNeil and Frey [3] pioneered the combination of EVT with GARCH models, proposing a two-stage estimation framework: first, fit a conditional heteroskedasticity model to obtain standardised residuals; then, use EVT to estimate tail quantiles and tail expectations of the residuals, thereby obtaining estimates of conditional risk measures. This approach has become the standard paradigm for dynamic risk measurement. Within this framework, the estimation of the extreme value index is a core component. The classical Hill estimator [4] is widely used due to its simplicity and asymptotic normality:

$$\hat{\gamma}_R^{\text{Hill}} = \frac{1}{k_1}\sum_{i=1}^{k_1} \log\frac{\varepsilon_{(i)}}{\varepsilon_{(k_1+1)}},$$

where $\varepsilon_{(1)} \geq \cdots \geq \varepsilon_{(n)}$ are order statistics. However, the Hill estimator only uses first-order moment information of the log-excesses, is sensitive to outliers, and can have substantial finite-sample bias when the effective sample size k_1 is poorly chosen.

To overcome the limitations of the Hill estimator, researchers have developed several improved methods. Dekkers *et al.* [5] proposed the moment estimator, extending the estimation range to negative extreme value indices. Gomes and Martins [6] and Segers [7] introduced the generalised moment estimator:

$$\hat{\gamma}_R^{(\alpha)} = \left(\frac{M_n^{(\alpha)}(k_1)}{\Gamma(\alpha+1)} \right)^{1/\alpha}, \quad M_n^{(\alpha)}(k_1) = \frac{1}{k_1} \sum_{i=1}^{k_1} \left(\log \frac{\varepsilon_{(i)}}{\varepsilon_{(k_1+1)}} \right)^{\alpha},$$

where $\alpha > 0$ is a tuning parameter. This family of estimators allows a trade-off between first-order ($\alpha = 1$) and higher-order moments by adjusting α, thereby achieving a flexible balance between bias and variance. When $\alpha = 1$, the generalised moment estimator reduces to the Hill estimator. Thus, the generalised moment estimator provides a unified framework for extreme value index estimation.

Although pointwise interval estimation of CVaR and CES at a single tail level is relatively well developed (e.g., Chan *et al.* [8]; Hoga [9]), in risk management practice regulators and investors often need to monitor multiple tail levels simultaneously, for example, CVaR at 0.5%, 0.8% and 1%. Constructing confidence intervals separately for each tail level would cause the overall coverage probability to deviate severely from the nominal level due to the multiple testing problem. Simultaneous confidence bands (SCBs) effectively address this issue by guaranteeing that the coverage probability over the entire tail region asymptotically attains the nominal level. Li *et al.* [2] systematically constructed joint SCBs for CVaR and CES under the location-scale model for the first time, established asymptotic theories for both the extrapolation extreme scenario ($a_1 = 0$) and the data-rich intermediate scenario ($a_1 > 0$), and designed a multiplier bootstrap method to improve finite-sample performance. However, that study still used the classical Hill estimator and did not fully exploit the flexibility of the generalised moment estimator family.

Based on the joint SCB framework of Li *et al.* [2], this paper generalises the extreme value index estimation from the Hill estimator to the generalised moment estimator, aiming to construct a more general joint inference method for conditional risk measures. Specifically, this paper addresses the following issues:

First, the asymptotic properties of the generalised moment estimator under time series dependence have not yet been clarified. Existing theory is mostly based on the i.i.d. assumption, whereas actual financial data often exhibit conditional heteroskedasticity. This paper derives the linear expansion of the generalised moment estimator in the residual sequence of the location-scale model and proves its joint asymptotic normality with order statistics, thereby providing a theoretical foundation for constructing SCBs.

Second, the bootstrap algorithm in the original literature depends on the specific linear form of the Hill estimator. This paper will derive a new linear expansion for the generalised moment estimator and design an adapted multiplier bootstrap statistic, enabling effective approximation of the limiting distribution of the original statistic under the intermediate scenario, thus overcoming the slow convergence rate of theoretical SCBs and the difficulty of approximating the limiting distribution in the intermediate scenario.

Third, the introduction of the tuning parameter α provides a new dimension for the bias-variance trade-off of the estimator. Through systematic Monte Carlo simulations, this paper evaluates the coverage probability and relative length of SCBs under different values of α, analyses the influence of α under symmetric and left-skewed distributions, and provides recommendations for parameter choice in practical applications.

Relative to Li *et al.* [2], this paper makes the following novel contributions. The overall simultaneous confidence band framework, the location-scale model paradigm, and the multiplier bootstrap strategy are inherited from that work. The new elements are: i) replacing the Hill estimator by the generalized moment estimator of Gomes and Martins [6] and Segers [7], which introduces the moment order α and encompasses the Hill estimator as the special case $\alpha = 1$; ii) deriving the linear expansion of the generalized moment estimator and designing a modified bootstrap statistic adapted to this expansion, together with a formal statement of why the bootstrap validity theorem of Li *et al.* [2] continues to hold; iii) providing a complete asymptotic theory for the generalized moment estimator under the intermediate scenario ($a_1 > 0$), culminating in a joint inference procedure for conditional risk measures.

2. The Framework and Estimation of CVaR and CES

2.1. The Framework

Assumption 2.1. R_t *follows the location-scale model*:

$$R_t = m(A_{t-1}, \beta_0) + \sigma(A_{t-1}, \beta_0)\varepsilon_t, \tag{2.1}$$

where $m(A_{t-1}, \beta_0)$ and $\sigma(A_{t-1}, \beta_0)$ denote the conditional mean and the conditional standard deviation of the return R_t given the information set A_{t-1}, respectively. The specific functional forms of these two functions are known, but they depend on an unknown parameter vector β_0. The innovation ε_t is independent of the information set A_{t-1} and forms an independent and identically distributed (i.i.d.) sequence of continuous random variables with mean zero and variance one. The information set at time t is defined as:

$$A_{t-1} = \sigma(\varepsilon_{t-1}, \varepsilon_{t-2}, \cdots; \eta_{t-1}, \eta_{t-2}, \cdots)$$

that is, the σ-algebra generated by the historical innovation sequence $\{\varepsilon_{t-i}\}_{i=1}^{\infty}$ and the possibly existing additional random vector sequence $\{\eta_{t-i}\}_{i=1}^{\infty}$. Here, $\{\eta_t\}_{t=-\infty}^{\infty}$ is a sequence of random vectors independent of the innovation sequence $\{\varepsilon_t\}_{t=-\infty}^{\infty}$. Furthermore, the functions $m(A_{t-1}, \beta_0)$ and $\sigma(A_{t-1}, \beta_0)$ are measurable with respect to A_{t-1}.

We suppose that ε_t follows a distribution with Pareto-type heavy tails. Let $F_\varepsilon(\cdot)=\Pr(\varepsilon_t\le\cdot)$ and $F_{-\varepsilon}(\cdot)=\Pr(-\varepsilon_t\le\cdot)$ be the distribution functions of ε_t and $-\varepsilon_t$. Then

$$\lim_{x\to\infty}\frac{1-F_\varepsilon(xs)}{1-F_\varepsilon(x)}=s^{-\gamma_R^{-1}},\ \text{ for all } s>0, \tag{2.2}$$

and

$$\lim_{x\to\infty}\frac{1-F_{-\varepsilon}(xs)}{1-F_{-\varepsilon}(x)}=s^{-\gamma_L^{-1}},\ \text{ for all } s>0, \tag{2.3}$$

where $\gamma_R>0$ and $\gamma_L>0$ denote the extreme value indices corresponding to the right and left tails, respectively. Note that under Assumption 2.1, the innovation ε_t is assumed to possess unit variance. This condition means that $\gamma_R<1/2$ and $\gamma_L<1/2$.

By model (2.1), we have

$$\text{U-CVaR}(\tau)=m(A_{t-1},\beta_0)+\sigma(A_{t-1},\beta_0)Q_\varepsilon(1-\tau), \tag{2.4}$$

$$\text{U-CES}(\tau)=m(A_{t-1},\beta_0)+\sigma(A_{t-1},\beta_0)E\left[\varepsilon\mid\varepsilon>Q_\varepsilon(1-\tau)\right], \tag{2.5}$$

$$\text{D-CVaR}(\tau)=-m(A_{t-1},\beta_0)-\sigma(A_{t-1},\beta_0)Q_\varepsilon(\tau), \tag{2.6}$$

$$\text{D-CES}(\tau)=-m(A_{t-1},\beta_0)-\sigma(A_{t-1},\beta_0)E\left[\varepsilon\mid\varepsilon<Q_\varepsilon(\tau)\right]. \tag{2.7}$$

where $Q_\varepsilon(\cdot)$ is the quantile function of ε_t. Assume that for each $\text{type}\in\{\text{U},\text{D},\text{R}\}$, we have estimators $\widetilde{\text{type-CVaR}}$ and $\widetilde{\text{type-CES}}$ for the true type-CVaR and type-CES quantities, respectively. Given a tail region $[\tau_l,\tau_u]$ satisfying $0<\tau_l\le\tau_u$, we define the maximum absolute log-ratios as

$$\sup_{\tau\in[\tau_l,\tau_u]}\left|\log\left(\frac{\widetilde{\text{type-CVaR}}(\tau)}{\text{type-CVaR}(\tau)}\right)\right| \text{ and } \sup_{\tau\in[\tau_l,\tau_u]}\left|\log\left(\frac{\widetilde{\text{type-CES}}(\tau)}{\text{type-CES}(\tau)}\right)\right|. \tag{2.8}$$

2.2. Estimation of CVaR and CES

Suppose that there is a consistent estimator $\hat\beta$ of β_0. Let $\tilde A_{t-1}$ be the truncated information set, which is generated by feasible information up to time $t-1$. The truncation is necessary when the information set relies on infinite past observations. For example, when $A_{t-1}=\{R_{t-1},R_{t-2},\cdots,R_0,R_{-1},\cdots\}$, the truncated information set is $\tilde A_{t-1}=\{R_{t-1},R_{t-2},\cdots,R_1\}$. We obtain the standardized residuals $\hat\varepsilon_t=\dfrac{R_t-m(\tilde A_{t-1},\hat\beta)}{\sigma(\tilde A_{t-1},\hat\beta)}$, $t=1,\cdots,n$. For $d_n<n$ and $d_n\to\infty$ as $n\to\infty$, we discard the residuals for $t<d_n$ and work with $\hat\varepsilon_t$ for $t=d_n,\cdots,n$. This discarding eliminates the effect of information truncation. Similar discarding can be applied to other information. Denote by $F_\varepsilon^{\leftarrow}$ the left continuous inverse of F_ε. Then, for $x>0$, the $(1-1/x)$-quantile of F_ε is

$$Q_\varepsilon(1-1/x) \equiv U_\varepsilon(x) = F_\varepsilon^{\leftarrow}(1-1/x).$$

By Theorem 1.2.1 and Corollary 1.2.10 of de Haan & Ferreira [10], (2.2) is equivalent to

$$\lim_{x\to\infty}\frac{U_\varepsilon(xs)}{U_\varepsilon(x)} = s^{\gamma_R}, \text{ for all } s>0. \tag{2.9}$$

Similarly, let $F_{-\varepsilon}^{\leftarrow}$ be the left continuous inverse of $F_{-\varepsilon}$ and $U_{-\varepsilon}(x) \equiv F_{-\varepsilon}^{\leftarrow}(1-1/x)$, for $x>0$. Then (2.3) is equivalent to

$$\lim_{x\to\infty}\frac{U_{-\varepsilon}(xs)}{U_{-\varepsilon}(x)} = s^{\gamma_L}, \text{ for all } s>0. \tag{2.10}$$

We first consider estimating the CVaR and CES of the right tail, namely $Q_\varepsilon(1-\tau)$ and $E\left[\varepsilon \mid \varepsilon > Q_\varepsilon(1-\tau)\right]$. For positive integers k_1, it satisfies that $k_1\to\infty$ and $k_1/n\to 0$ as $n\to\infty$. By (2.9), for a small τ implies

$$Q_\varepsilon(1-\tau) = U_\varepsilon\left(\frac{1}{\tau}\right) \sim U_\varepsilon\left(\frac{n}{k_1}\right)\left(\frac{n\tau}{k_1}\right)^{-\gamma_R}, \text{ as } n\to\infty \tag{2.11}$$

To obtain an estimate of $Q_\varepsilon(1-\tau)$, we need to estimate γ_R and $U_\varepsilon\left(\frac{n}{k_1}\right)$. A suitable estimator for $U_\varepsilon(n/k_1)$ is given by $\hat{\varepsilon}_{n-k_1,n}$, commonly referred to as the intermediate order statistic de Haan & Ferreira [10]. For estimation of $\gamma_R>0$, we use

$$\hat{\gamma}_R = \gamma_n^{(\alpha)}(k_1) = \left\{\frac{M_n^{(\alpha)}(k_1)}{\Gamma(\alpha+1)}\right\}^{\frac{1}{\alpha}} \tag{2.12}$$

with $\alpha>0$ (Gomes and Martins 2001; Segers 2001), where $\Gamma(\cdot)$ is the gamma function and

$$M_n^{(\alpha)}(k_1) = \frac{1}{k_1}\sum_{i=0}^{k_1}\left(\log\frac{\hat{\varepsilon}_{n-i,n}}{\hat{\varepsilon}_{n-k_1,n}}\right)^{\alpha} \tag{2.13}$$

By substituting $\hat{\varepsilon}_{n-k_1,n}$ and $\hat{\gamma}_R$ into equation (2.11), the following estimate of $Q_\varepsilon(1-\tau)$ can be obtained:

$$\hat{Q}_\varepsilon(1-\tau) = \hat{\varepsilon}_{n-k_1,n}\left(\frac{n\tau}{k_1}\right)^{-\hat{\gamma}_R} \tag{2.14}$$

By Proposition 4.1 of Pan, Leng, and Hu (2013) [11], we have

$$\lim_{\tau\to 0}\frac{E\left[\varepsilon \mid \varepsilon > Q_\varepsilon(1-\tau)\right]}{Q_\varepsilon(1-\tau)} = \frac{1}{1-\gamma_R} \tag{2.15}$$

Since τ is close to 0, the estimation of $E\left[\varepsilon \mid \varepsilon > Q_\varepsilon(1-\tau)\right]$ is

$$\hat{E}\left[\varepsilon \mid \varepsilon > Q_\varepsilon(1-\tau)\right] = \frac{\hat{Q}_\varepsilon(1-\tau)}{1-\hat{\gamma}_R} \tag{2.16}$$

Estimation of the left tail proceeds in the same way as what is done for the right tail. Note that $U_{-\varepsilon}(1/\tau) = -Q_\varepsilon(\tau)$. Therefore, estimation of $Q_\varepsilon(\tau)$ is equivalent to estimation of $-U_{-\varepsilon}(1/\tau)$. Then, for positive integers k_2, it satisfies $k_2 \to \infty$ and $k_2/n \to 0$ as $n \to \infty$, following similar arguments leading to (2.14) and (2.16), we obtain the estimator

$$\hat{Q}_\varepsilon(\tau) = \hat{\varepsilon}_{d_n+k_2,n}\left(\frac{n\tau}{k_2}\right)^{-\hat{\gamma}_L}$$

for $Q_\varepsilon(\tau)$, and the estimator

$$\hat{E}\left[|\varepsilon| \mid \varepsilon < Q_\varepsilon(\tau)\right] = \frac{\hat{Q}_\varepsilon(\tau)}{1-\hat{\gamma}_L}$$

for $E\left[|\varepsilon| \mid \varepsilon < Q_\varepsilon(\tau)\right]$, where $\hat{\gamma}_L$ is estimator of γ_L based on $\{-\hat{\varepsilon}_i\}$, defined as

$$\hat{\gamma}_L = \gamma_n^{(\alpha)}(k_2) = \left\{\frac{M_n^{(\alpha)}(k_2)}{\Gamma(\alpha+1)}\right\}^{\frac{1}{\alpha}}$$

and

$$M_n^{(\alpha)}(k_2) = \frac{1}{k_2}\sum_{i=0}^{k_2}\left(\log\frac{-\hat{\varepsilon}_{i+d_n,n}}{-\hat{\varepsilon}_{k_2+d_n,n}}\right)^\alpha = \frac{1}{k_2}\sum_{i=0}^{k_2}\left(\log\frac{\hat{\varepsilon}_{i+d_n,n}}{\hat{\varepsilon}_{k_2+d_n,n}}\right)^\alpha$$

Plugging the above estimators into (2.4)-(2.7), we have estimators of the risk measures

$$\widetilde{\text{U-CVaR}}(\tau) = m\left(\tilde{A}_n, \hat{\beta}\right) + \sigma\left(\tilde{A}_n, \hat{\beta}\right)\hat{Q}_\varepsilon(1-\tau), \tag{2.17}$$

$$\widetilde{\text{U-CES}}(\tau) = m\left(\tilde{A}_n, \hat{\beta}\right) + \sigma\left(\tilde{A}_n, \hat{\beta}\right)\hat{E}\left[\varepsilon \mid \varepsilon > Q_\varepsilon(1-\tau)\right], \tag{2.18}$$

$$\widetilde{\text{D-CVaR}}(\tau) = -m\left(\tilde{A}_n, \hat{\beta}\right) - \sigma\left(\tilde{A}_n, \hat{\beta}\right)\hat{Q}_\varepsilon(\tau), \tag{2.19}$$

and

$$\widetilde{\text{D-CES}}(\tau) = -m\left(\tilde{A}_n, \hat{\beta}\right) - \sigma\left(\tilde{A}_n, \hat{\beta}\right)\hat{E}\left[\varepsilon \mid \varepsilon < Q_\varepsilon(\tau)\right]. \tag{2.20}$$

And apparently, we have estimators of the relative risk measures

$$\widetilde{\text{R-CVaR}}(\tau) = \frac{\widetilde{\text{U-CVaR}}(\tau)}{\widetilde{\text{D-CVaR}}(\tau)} \text{ and } \widetilde{\text{R-CES}}(\tau) = \frac{\widetilde{\text{U-CES}}(\tau)}{\widetilde{\text{D-CES}}(\tau)}. \tag{2.21}$$

3. Asymptotic Theory

3.1. Assumptions

Assumption 3.1. *The estimator* $\hat{\beta}$ *of* β_0 *satisfies* $\left\|\hat{\beta}-\beta_0\right\| = O_p\left(n^{-\upsilon_0/2}\right)$ *for some positive* υ_0.

Assumption 3.2. *Let* B_0 *be a neighborhood of* β_0.

(i) $\sup_{\beta\in B_0} m(A_n, \beta_0) = O_p(1)$.

(ii) $\inf_{\beta\in B_0} \sigma(A_{t-1}, \beta) > c$ *for some* $c > 0$ *for all* $t = d_n, \cdots, n+1$.

(iii) *Both* $m(A_{t-1},\beta)$ *and* $\sigma(A_{t-1},\beta)$ *are differentiable with respect to* β *in* B_0 *for all* $t=d_n,\cdots,n+1$.

(iv) $E\left[\sup_{\beta\in B_0}\left\|\frac{\partial m(A_{t-1},\beta)}{\partial\beta}\right\|^{\upsilon_1}\right]<\infty$ *and* $E\left[\sup_{\beta\in B_0}\left\|\frac{\partial \sigma(A_{t-1},\beta)}{\partial\beta}\right\|^{\upsilon_1}\right]<\infty$ *for some* $\upsilon_1>2/\upsilon_0$ *for all* $t=d_n,\cdots,n+1$, *where* υ_0 *is defined in Assumption* 3.1.

Assumption 3.3.

(i) *For each* $\beta\in B_0$, $m(\tilde{A}_{t-1},\beta)$ *and* $\sigma(\tilde{A}_{t-1},\beta)$ *are measurable with respect to* A_{t-1}.

(ii) $\sup_{\beta\in B_0}\left|m(\tilde{A}_n,\beta)-m(A_n,\beta)\right|=O_p\left(n^{-\upsilon_0/2}\right)$ *and* $\sup_{\beta\in B_0}\left|\sigma(\tilde{A}_n,\beta)-\sigma(A_n,\beta)\right|=O_p\left(n^{-\upsilon_0/2}\right)$, *where* υ_0 *is defined in Assumption* 3.1.

(iii) $\sum_{t=d_n}^{n}E\left[\sup_{\beta\in B_0}\left|m(\tilde{A}_{t-1},\beta)-m(A_{t-1},\beta)\right|\right]=o(1)$ *and* $\sum_{t=d_n}^{n}E\left[\sup_{\beta\in B_0}\left|\sigma(\tilde{A}_{t-1},\beta)-\sigma(A_{t-1},\beta)\right|\right]=o(1)$.

Assumption 3.4.

(i) *There exist* $\rho_R<0$ *and a function* $A_R(\cdot)$ *which is eventually positive or negative with* $\lim_{x\to\infty}A_R(x)=0$ *such that*

$$\lim_{x\to\infty}\frac{\frac{U_\varepsilon(xs)}{U_\varepsilon(x)}-s^{\gamma_R}}{A_R(x)}=s^{\gamma_R}\frac{s^{\rho_R}-1}{\rho_R},\ \textit{for all}\ s>0,$$

where $k_1^{1/2}A_R(n/k_1)\to 0$ *as* $n\to\infty$.

(ii) *There exist* $\rho_L<0$ *and a function* $A_L(\cdot)$ *which is eventually positive or negative with* $\lim_{x\to\infty}A_L(x)=0$ *such that*

$$\lim_{x\to\infty}\frac{\frac{U_{-\varepsilon}(xs)}{U_{-\varepsilon}(x)}-s^{\gamma_L}}{A_L(x)}=s^{\gamma_L}\frac{s^{\rho_L}-1}{\rho_L},\ \textit{for all}\ s>0,$$

where $k_2^{1/2}A_L(n/k_2)\to 0$ *as* $n\to\infty$.

Assumption 3.5.

(i) $\tau_l\le\tau_u$ *satisfies* $\tau_l\to 0$ *as* $n\to\infty$ *and* $\tau_u=C\tau_l$ *for some constant* $C\ge 1$.

(ii) *There exist two constants* $a_0>0$ *and* $a_1>0$ *such that* $\lim_{n\to\infty}k_1/k_2=a_0$ *and* $\lim_{n\to\infty}n\tau_l/k_1=a_1$.

(iii) $\lim_{n\to\infty}k_1^{-1/2}\log(n\tau_l/k_1)=0$ *and* $k_1=o\left(n^{\delta_0}\right)$ *for some* $0<\delta_0<\upsilon_0$, *where* υ_0 *is defined in Assumption* 3.1.

iv) $d_n=o\left(k_1^{1/2}\right)$.

3.2. Scenarios of Asymptotics

Our theoretical findings are grounded in the following three propositions from

Proposition 3.1 and 3.2 of Li *et al.* [2].

Proposition 3.1. *Under Assumptions* 2.1 *and* 3.1 - 3.5, *we have, uniformly in* $\tau \in [\tau_l, \tau_u]$,

$$\frac{\widetilde{\text{U-CVaR}}(\tau)}{\text{U-CVaR}(\tau)} - 1 = \frac{\hat{Q}_\varepsilon(1-\tau)}{Q_\varepsilon(1-\tau)} - 1 + o_p\left(k_1^{-1/2}\log(n\tau/k_1)\right),$$

$$\frac{\widetilde{\text{D-CVaR}}(\tau)}{\text{D-CVaR}(\tau)} - 1 = \frac{\hat{Q}_\varepsilon(\tau)}{Q_\varepsilon(\tau)} - 1 + o_p\left(k_1^{-1/2}\log(n\tau/k_1)\right),$$

$$\frac{\widetilde{\text{U-CES}}(\tau)}{\text{U-CES}(\tau)} - 1 = \frac{\hat{E}\left[\varepsilon > Q_\varepsilon(1-\tau)\right]}{E\left[\varepsilon > Q_\varepsilon(1-\tau)\right]} - 1 + o_p\left(k_1^{-1/2}\log(n\tau/k_1)\right),$$

and

$$\frac{\widetilde{\text{D-CES}}(\tau)}{\text{D-CES}(\tau)} - 1 = \frac{\hat{E}\left[\varepsilon < Q_\varepsilon(\tau)\right]}{E\left[\varepsilon < Q_\varepsilon(\tau)\right]} - 1 + o_p\left(k_1^{-1/2}\log(n\tau/k_1)\right).$$

Proposition 3.2. *Under Assumptions* 2.1 and 3.1 - 3.5, *we have, uniformly in* $\tau \in [\tau_l, \tau_u]$,

$$\frac{\hat{Q}_\varepsilon(1-\tau)}{Q_\varepsilon(1-\tau)} - 1 = \left[\frac{\hat{\varepsilon}_{n-k_1,n}}{U_\varepsilon(n/k_1)} - 1\right] + \log\left(\frac{k_1}{n\tau}\right)(\hat{\gamma}_R - \gamma_R) + o_p\left(k_1^{-1/2}\log(n\tau/k_1)\right),$$

$$\frac{\hat{Q}_\varepsilon(\tau)}{Q_\varepsilon(\tau)} - 1 = \left[\frac{-\hat{\varepsilon}_{k_2+d_n,n}}{U_{-\varepsilon}(n/k_2)} - 1\right] + \log\left(\frac{k_2}{n\tau}\right)(\hat{\gamma}_L - \gamma_L) + o_p\left(k_1^{-1/2}\log(n\tau/k_1)\right),$$

$$\frac{\hat{E}\left[\varepsilon > Q_\varepsilon(1-\tau)\right]}{E\left[\varepsilon > Q_\varepsilon(1-\tau)\right]} - 1 = \left[\frac{\hat{\varepsilon}_{n-k_1,n}}{U_\varepsilon(n/k_1)} - 1\right] + \left[\log\left(\frac{k_1}{n\tau}\right) + \frac{1}{1-\gamma_R}\right](\hat{\gamma}_R - \gamma_R) + o_p\left(k_1^{-1/2}\log(n\tau/k_1)\right),$$

and

$$\frac{\hat{E}\left[\varepsilon < Q_\varepsilon(\tau)\right]}{E\left[\varepsilon < Q_\varepsilon(\tau)\right]} - 1 = \left[\frac{-\hat{\varepsilon}_{k_2+d_n,n}}{U_{-\varepsilon}(n/k_2)} - 1\right] + \left[\log\left(\frac{k_2}{n\tau}\right) + \frac{1}{1-\gamma_L}\right](\hat{\gamma}_L - \gamma_L) + o_p\left(k_1^{-1/2}\log(n\tau/k_1)\right).$$

Proposition 3.3. *Under Assumptions* 2.1 *and* 3.1 - 3.5, *as* $n \to \infty$,

$$\begin{pmatrix} k_1^{1/2}\gamma_R^{-1}\left[\hat{\varepsilon}_{n-k_1,n}/U_\varepsilon(n/k_1) - 1\right] \\ k_1^{1/2}\gamma_R^{-1}\tilde{V}^{-1/2}(\alpha)(\hat{\gamma}_R - \gamma_R) \\ k_2^{1/2}\gamma_L^{-1}\left[-\hat{\varepsilon}_{k_2+d_n,n}/U_{-\varepsilon}(n/k_2) - 1\right] \\ k_2^{1/2}\gamma_L^{-1}\tilde{V}^{-1/2}(\alpha)(\hat{\gamma}_L - \gamma_L) \end{pmatrix}$$

is asymptotically four-dimensional standard normal where
$\tilde{V}(\alpha) = \frac{1}{\alpha^2}\left[\frac{\Gamma(2\alpha+1)}{\Gamma^2(\alpha+1)} - 1\right]$.

Proof. Let $\varepsilon_1, \cdots, \varepsilon_n$ be i.i.d. random variables satisfying Assumptions 3.4(i) and (ii). Define

$$\hat{\gamma}_{\varepsilon,R} = \left\{ \frac{k_1^{-1} \sum_{i=0}^{k_1} \left[\log\left(\varepsilon_{n-i:n} / \varepsilon_{n-k_1:n} \right) \right]^{\alpha}}{\Gamma(\alpha+1)} \right\}^{\frac{1}{\alpha}},$$

$$\hat{\gamma}_{\varepsilon,L} = \left\{ \frac{k_2^{-1} \sum_{i=0}^{k_2} \left[\log\left(-\varepsilon_{i+1:n} / -\varepsilon_{k_2+1:n} \right) \right]^{\alpha}}{\Gamma(\alpha+1)} \right\}^{\frac{1}{\alpha}}.$$

By Lemma 1.5 in the supplementary material of Li, Peng and Song [2],

$$\begin{pmatrix} k_1^{1/2} \gamma_R^{-1} \left(\varepsilon_{n-k_1:n} / U_\varepsilon(n/k_1) - 1 \right) \\ k_2^{1/2} \gamma_L^{-1} \left(-\varepsilon_{k_2+1:n} / U_{-\varepsilon}(n/k_2) - 1 \right) \end{pmatrix} \xrightarrow{\mathcal{D}} N(0, I_2),$$

and the two components are asymptotically independent. Gomes and Martins [6] proved that for any $\alpha > 0$,

$$k_1^{1/2} \left(\hat{\gamma}_{\varepsilon,R} - \gamma_R \right) \xrightarrow{\mathcal{D}} N\left(0, \gamma_R^2 \tilde{V}(\alpha)\right), \; k_2^{1/2} \left(\hat{\gamma}_{\varepsilon,L} - \gamma_L \right) \xrightarrow{\mathcal{D}} N\left(0, \gamma_L^2 \tilde{V}(\alpha)\right).$$

Therefore, to obtain the joint convergence of the four-dimensional vector, it suffices to establish the asymptotic independence among the four components.

Using the probability integral transformation, set $U_i = F_\varepsilon(\varepsilon_i)$; then $U_i \sim U(0,1)$ are independent. Define the Pareto variables $\zeta_i = 1/(1-U_i)$, which are i.i.d. with distribution $F_\zeta(y) = 1 - 1/y$ $(y \ge 1)$, and we have $\varepsilon_i = U_\varepsilon(\zeta_i)$, where $U_\varepsilon(t) = Q_\varepsilon(1 - 1/t)$. The order statistics satisfy $\varepsilon_{i:n} = U_\varepsilon(\zeta_{i:n})$. For the left tail, let $V_i = F_{-\varepsilon}(-\varepsilon_i) = 1 - U_i$; then $-\varepsilon_i = U_{-\varepsilon}\left(1/(1-V_i)\right)$ and $-\varepsilon_{i:n} = U_{-\varepsilon}\left(\zeta_{i:n}^-\right)$ with $\zeta_{i:n}^- = 1/(1-V_{i:n})$.

By Rényi's representation [12], there exist i.i.d. $\mathrm{Exp}(1)$ random variables $E_1, \cdots, E_n$ such that

$$\zeta_{i:n} \overset{d}{=} \frac{n}{E_1 + \cdots + E_{n-i+1}}, \; i = 1, \cdots, n.$$

Consequently,

$$\zeta_{n-k_1:n} \overset{d}{=} \frac{n}{S_{k_1+1}}, \; S_{k_1+1} = E_1 + \cdots + E_{k_1+1},$$

and the ratios

$$\frac{\zeta_{n-i:n}}{\zeta_{n-k_1:n}} \overset{d}{=} \frac{S_{k_1+1}}{S_{k_1+1-i}}, \; i = 1, \cdots, k_1,$$

depend only on $E_1, \cdots, E_{k_1}$ and are independent of $\zeta_{n-k_1:n}$ (because $\zeta_{n-k_1:n}$ also depends on E_{k_1+1}). Similarly, for the left tail, $\zeta_{k_2+1:n}^-$ depends on $E_{n-k_2}, \cdots, E_n$. Since $k_1 + k_2 < n$ for sufficiently large n, the sets $\left\{E_1, \cdots, E_{k_1+1}\right\}$ and $\left\{E_{n-k_2}, \cdots, E_n\right\}$ are disjoint; hence the right-tail block $\left(\zeta_{n-k_1:n}, \left\{\zeta_{n-i:n} / \zeta_{n-k_1:n}\right\}_{i=1}^{k_1}\right)$ and the left-tail block $\left(\zeta_{k_2+1:n}^-, \left\{\zeta_{i:n}^- / \zeta_{k_2+1:n}^-\right\}_{i=1}^{k_2}\right)$ are independent.

Notice that $\hat{\gamma}_{\varepsilon,R}$ depends only on $\left\{\log\left(\zeta_{n-i:n}/\zeta_{n-k_1:n}\right)\right\}_{i=1}^{k_1}$, so $\hat{\gamma}_{\varepsilon,R}$ is independent of $\zeta_{n-k_1:n}$. Therefore,

$$k_1^{1/2}\gamma_R^{-1}\left(\varepsilon_{n-k_1:n}/U_\varepsilon\left(n/k_1\right)-1\right) \text{ and } k_1^{1/2}\gamma_R^{-1}\tilde{V}^{-1/2}\left(\alpha\right)\left(\hat{\gamma}_{\varepsilon,R}-\gamma_R\right)$$

are asymptotically independent. The same holds for the two standardized statistics on the left tail, and the left and right tails are independent of each other. Combining marginal asymptotic normality with the Cramér-Wold theorem yields

$$\begin{pmatrix} k_1^{1/2}\gamma_R^{-1}\left(\varepsilon_{n-k_1:n}/U_\varepsilon\left(n/k_1\right)-1\right) \\ k_1^{1/2}\gamma_R^{-1}\tilde{V}^{-1/2}\left(\alpha\right)\left(\hat{\gamma}_{\varepsilon,R}-\gamma_R\right) \\ k_2^{1/2}\gamma_L^{-1}\left(-\varepsilon_{k_2+1:n}/U_{-\varepsilon}\left(n/k_2\right)-1\right) \\ k_2^{1/2}\gamma_L^{-1}\tilde{V}^{-1/2}\left(\alpha\right)\left(\hat{\gamma}_{\varepsilon,L}-\gamma_L\right) \end{pmatrix} \xrightarrow{\mathcal{D}} N\left(0,I_4\right).$$

Now consider the standardized residuals $\hat{\varepsilon}_t$ defined in the proposition. By Lemma S1.9(ii) in the supplementary material of Li, Peng and Song [2],

$$\hat{\varepsilon}_{(k_1)} = \varepsilon_{(k_1)} + o_p\left(k_1^{-1/2}\right),\ \hat{\gamma}_R = \hat{\gamma}_{\varepsilon,R} + o_p\left(k_1^{-1/2}\right),$$

and analogous relations hold for the left tail. Hence, replacing the true residuals with the standardized residuals changes each component of the four-dimensional vector by $o_p(1)$, so the limiting distribution remains unchanged.

Moreover, the left-tail quantile in the proposition uses $\hat{\varepsilon}_{k_2+d_n,n}$, whereas the lemma uses $\varepsilon_{k_2+1,n}$. By Assumption 3.5(iv) and the fact that k_2 is of the same order as k_1, we have $d_n = o\left(\min\left(k_1^{1/2},k_2^{1/2}\right)\right)$, thus

$$k_2^{1/2}\left(\hat{\varepsilon}_{k_2+d_n,n}-\hat{\varepsilon}_{k_2+1,n}\right) = o_p(1),$$

so replacing k_2+1 with k_2+d_n does not affect the asymptotic distribution. This completes the proof of Proposition 3. □

3.3. Main Results

The following theorem establishes the uniform convergence of the maximum absolute log-ratios under the intermediate scenario $a_1 > 0$.

Theorem 3.1. *Suppose that Assumptions* 2.1 *and* 3.1 - 3.5 *hold. As* $n \to \infty$,

$$\sup_{u\in[1,C]}\left|k_1^{1/2}\left\{\log\left[\frac{\widetilde{\text{type-CVaR}}\left(u\tau_l\right)}{\text{type-CVaR}\left(u\tau_l\right)}\right]\right\}\right| \to \sup_{u\in[1,C]}\left|G_{1,\text{type}}(u)\right| \quad \textit{and}$$

$$\sup_{u\in[1,C]}\left|k_1^{1/2}\left\{\log\left[\frac{\widetilde{\text{type-CES}}\left(u\tau_l\right)}{\text{type-CES}\left(u\tau_l\right)}\right]\right\}\right| \to \sup_{u\in[1,C]}\left|G_{2,\text{type}}(u)\right|$$

in distribution, where $G_{1,\text{type}}(u)$ *and* $G_{2,\text{type}}(u)$ *are centered Gaussian processes with variance-covariance functions*

$\gamma_{1,\text{type}}\left(u_1,u_2\right) = \text{Cov}\left(G_{1,\text{type}}\left(u_1\right),G_{1,\text{type}}\left(u_2\right)\right)$ *and*

$\gamma_{2,\text{type}}\left(u_1,u_2\right) = \text{Cov}\left(G_{2,\text{type}}\left(u_1\right),G_{2,\text{type}}\left(u_2\right)\right)$ *satisfying that*

$$\gamma_{1,\text{type}}(u_1,u_2)=\begin{cases}\gamma_R^2\left[1+\tilde{V}(\alpha)\log(a_1u_1)\log(a_1u_2)\right], & \text{if type}=U,\\ a_0\gamma_L^2\left[1+\tilde{V}(\alpha)\log(a_0a_1u_1)\log(a_0a_1u_2)\right], & \text{if type}=D,\\ \gamma_R^2\left[1+\tilde{V}(\alpha)\log(a_1u_1)\log(a_1u_2)\right] & \\ \quad+a_0\gamma_L^2\left[1+\tilde{V}(\alpha)\log(a_0a_1u_1)\log(a_0a_1u_2)\right], & \text{if type}=R.\end{cases}$$

and

$$\gamma_{2,\text{type}}(u_1,u_2)$$
$$=\begin{cases}\gamma_R^2\left\{1+\tilde{V}(\alpha)\left[(1-\gamma_R)^{-1}-\log(a_1u_1)\right]\left[(1-\gamma_R)^{-1}-\log(a_1u_2)\right]\right\}, & \text{if type}=U,\\ a_0\gamma_L^2\left\{1+\tilde{V}(\alpha)\left[(1-\gamma_L)^{-1}-\log(a_0a_1u_1)\right]\left[(1-\gamma_L)^{-1}-\log(a_0a_1u_2)\right]\right\}, & \text{if type}=D,\\ \gamma_R^2\left\{1+\tilde{V}(\alpha)\left[(1-\gamma_R)^{-1}-\log(a_1u_1)\right]\left[(1-\gamma_R)^{-1}-\log(a_1u_2)\right]\right\} & \\ \quad+a_0\gamma_L^2\left\{1+\tilde{V}(\alpha)\left[(1-\gamma_L)^{-1}-\log(a_0a_1u_1)\right]\left[(1-\gamma_L)^{-1}-\log(a_0a_1u_2)\right]\right\}, & \text{if type}=R.\end{cases}$$

Proof. For $\text{type}\in\{\text{U}\}$,

$$\begin{aligned}F_U(u)&=k_1^{1/2}\left\{\log\left[\frac{\widehat{\text{U-CVaR}}(u\tau_l)}{\text{U-CVaR}(u\tau_l)}\right]\right\}\\&=k_1^{1/2}\left(\frac{\widehat{\text{U-CVaR}}(u\tau_l)}{\text{U-CVaR}(u\tau_l)}-1\right)+o_p(1)\\&=k_1^{1/2}\left(\frac{\tilde{Q}_\varepsilon(1-u\tau)}{Q_\varepsilon(1-u\tau)}-1\right)+o_p(1)\\&=k_1^{1/2}\left[\frac{\hat{\varepsilon}_{n-k_1,n}}{U_\varepsilon(n/k_1)}-1\right]+\left[\log\left(\frac{k_1}{n\tau}\right)-\log(u)\right]k_1^{1/2}(\hat{\gamma}_R-\gamma_R)+o_p(1)\end{aligned}$$

For any $u_1,\cdots,u_m\in[1,C]$ and fixed $m\in\mathbb{N}$, by Lemma S1.11 in the supplementary material of Li *et al.* [2], Proposition 3 and the continuous mapping theorem, $\left(F_U(u_1),\cdots,F_U(u_m)\right)$ converges weakly to $\left(G_{1,U}(u_1),\cdots,G_{1,U}(u_m)\right)$. Let $X_1=k_1^{1/2}\left[\frac{\hat{\varepsilon}_{n-k_1,n}}{U_\varepsilon(n/k_1)}-1\right]$, $Y_1=\left[\log\left(\frac{k_1}{n\tau}\right)-\log(u_1)\right]k_1^{1/2}(\hat{\gamma}_R-\gamma_R)$, $Y_2=\left[\log\left(\frac{k_1}{n\tau}\right)-\log(u_2)\right]k_1^{1/2}(\hat{\gamma}_R-\gamma_R)$. Then

$$Cov\left(G_{1,U}(u_1),G_{1,U}(u_2)\right)=\gamma_R^2\left(1+\tilde{V}(\alpha)\log(a_1u_1)\log(a_1u_2)\right).$$

For $\text{type}\in\{\text{D}\}$,

$$\begin{aligned}F_D(u)&=k_1^{1/2}\left\{\log\left[\frac{\widehat{\text{D-CVaR}}(u\tau_l)}{\text{D-CVaR}(u\tau_l)}\right]\right\}\\&=k_1^{1/2}\left(\frac{\widehat{\text{D-CVaR}}(u\tau_l)}{\text{D-CVaR}(u\tau_l)}-1\right)+o_p(1)\\&=k_1^{1/2}\left(\frac{\tilde{Q}_\varepsilon(u\tau)}{Q_\varepsilon(u\tau)}-1\right)+o_p(1)\\&=\frac{k_1^{1/2}}{k_2^{1/2}}\left\{k_2^{1/2}\left[\frac{-\hat{\varepsilon}_{k_2+d_n,n}}{U_{-\varepsilon}(n/k_2)}-1\right]+\left[\log\left(\frac{k_2}{n\tau}\right)-\log(u)\right]k_2^{1/2}(\hat{\gamma}_L-\gamma_L)\right\}+o_p(1)\end{aligned}$$

$$Cov\left(G_{1,D}\left(u_1\right),G_{1,D}\left(u_2\right)\right)=a_0\gamma_L^2\left(1+\tilde{V}\left(\alpha\right)\log\left(a_0a_1u_1\right)\log\left(a_0a_1u_2\right)\right).$$

For $\text{type}\in\{\text{R}\}$,

$$\begin{aligned}F_R\left(u\right)&=k_1^{1/2}\left\{\log\left[\frac{\widetilde{\text{R-CVaR}}\left(u\tau_l\right)}{\text{R-CVaR}\left(u\tau_l\right)}\right]\right\}\\&=k_1^{1/2}\left\{\log\left[\frac{\widetilde{\text{U-CVaR}}\left(u\tau_l\right)}{\text{U-CVaR}\left(u\tau_l\right)}\right]-\log\left[\frac{\widetilde{\text{D-CVaR}}\left(u\tau_l\right)}{\text{D-CVaR}\left(u\tau_l\right)}\right]\right\}+o_p\left(1\right)\\&=k_1^{1/2}\left[\frac{\hat{\varepsilon}_{n-k_1,n}}{U_\varepsilon\left(n/k_1\right)}-1\right]+\left[\log\left(\frac{k_1}{n\tau}\right)-\log\left(u\right)\right]k_1^{1/2}\left(\hat{\gamma}_R-\gamma_R\right)\\&\quad-\frac{k_1^{1/2}}{k_2^{1/2}}\left\{k_2^{1/2}\left[\frac{-\hat{\varepsilon}_{k_2+d_n,n}}{U_{-\varepsilon}\left(n/k_2\right)}-1\right]+\left[\log\left(\frac{k_2}{n\tau}\right)-\log\left(u\right)\right]k_2^{1/2}\left(\hat{\gamma}_L-\gamma_L\right)\right\}+o_p\left(1\right)\end{aligned}$$

Because the right- and left-tail components are asymptotically independent,

$$\begin{aligned}Cov\left(G_{1,R}\left(u_1\right),G_{1,R}\left(u_2\right)\right)&=\gamma_R^2\left(1+\tilde{V}\left(\alpha\right)\log\left(a_1u_1\right)\log\left(a_1u_2\right)\right)\\&\quad+a_0\gamma_L^2\left(1+\tilde{V}\left(\alpha\right)\log\left(a_0a_1u_1\right)\log\left(a_0a_1u_2\right)\right).\end{aligned}$$

The proof method of CES is similar to that of CVaR, and the proof is omitted here. □

For the intermediate scenario $a_1>0$, the limiting distribution is a non-degenerate Gaussian process with a complicated covariance structure (Theorem 3.1). Direct computation of critical values is therefore difficult, and we instead employ the bootstrap method described in Section 4.

The results of this section extend the pointwise inference procedures of Chan *et al.* [8] and Hoga [9] to the joint inference setting, under the intermediate scenario that has been studied by Martins-Filho *et al.* [13] among others for pointwise estimation. Our simultaneous SCBs provide inference over the whole tail interval.

4. Bootstrap Implementation

Although Theorem 3.1 provides the limiting distributions of the maximum absolute log-ratio, two major difficulties arise when applying these theorems to construct simultaneous confidence bands (SCBs) in practice. First, because the effective sample sizes k_1,k_2 are small relative to the total sample size n and the tail level τ approaches zero, the convergence rate of the maximum absolute log-ratio is slow. As a result, theoretical SCBs directly derived from Theorem 3.1 may exhibit severely distorted coverage probabilities in finite samples. Second, the limiting distribution is a complicated Gaussian process whose covariance structure heavily depends on the underlying model and is difficult to approximate directly. To overcome these difficulties, this paper adopts the bootstrap method proposed by Li *et al.* [2] for finite-sample inference.

Proposition 4.1. *Under Assumptions* 2.1 *and* 3.1-3.5, *the generalized moment estimators admit the linear expansions*

$$\hat{\gamma}_R-\gamma_R=H_R\left(\alpha\right)+o_p\left(k_1^{-1/2}\right),\ \hat{\gamma}_L-\gamma_L=H_L\left(\alpha\right)+o_p\left(k_2^{-1/2}\right),$$

where

$$H_R(\alpha)=\frac{1}{\alpha\Gamma(\alpha+1)\gamma_R^{\alpha-1}}\left(M_{n,R}^{(\alpha)}(k_1)-\Gamma(\alpha+1)\gamma_R^{\alpha}\right)$$
$$=\frac{1}{\alpha\Gamma(\alpha+1)\gamma_R^{\alpha-1}}\left(\frac{1}{k_1}\sum_{t=d_n}^{n}\mathbb{I}\left(\varepsilon_t>\varepsilon_{n-k_1,n}\right)\left(\log\frac{\varepsilon_t}{\varepsilon_{n-k_1,n}}\right)^{\alpha}-\Gamma(\alpha+1)\gamma_R^{\alpha}\right), \tag{4.1}$$

and

$$H_L(\alpha)=\frac{1}{\alpha\Gamma(\alpha+1)\gamma_L^{\alpha-1}}\left(M_{n,L}^{(\alpha)}(k_2)-\Gamma(\alpha+1)\gamma_L^{\alpha}\right)$$
$$=\frac{1}{\alpha\Gamma(\alpha+1)\gamma_L^{\alpha-1}}\left(\frac{1}{k_2}\sum_{t=d_n}^{n}\mathbb{I}\left(\varepsilon_t<\varepsilon_{k_2+d_n,n}\right)\left(\log\frac{\varepsilon_t}{\varepsilon_{k_2+d_n,n}}\right)^{\alpha}-\Gamma(\alpha+1)\gamma_L^{\alpha}\right). \tag{4.2}$$

Proof. We only prove the case for the right tail; the proof for the left tail is completely analogous. According to Theorem 4.1 of Segers [7], under Assumptions 2.1 and 3.1-3.5, the sample moment $M_{n,R}^{(\alpha)}(k_1)$ satisfies the following asymptotic normality:

$$\sqrt{k_1}\left(M_{n,R}^{(\alpha)}(k_1)-\Gamma(\alpha+1)\gamma_R^{\alpha}\right)\xrightarrow{d}N\left(0,\sigma^2(\gamma_R,\alpha)\right),$$

where $\sigma^2(\gamma_R,\alpha)$ is the asymptotic variance; its explicit form can be found in Segers [7] or de Haan & Ferreira [10]. This result implies

$$M_{n,R}^{(\alpha)}(k_1)-\Gamma(\alpha+1)\gamma_R^{\alpha}=O_p\left(k_1^{-1/2}\right).$$

Consider $\hat{\gamma}_R$ as a differentiable function of $M_{n,R}^{(\alpha)}(k_1)$: define

$$g(t)=\left(\frac{t}{\Gamma(\alpha+1)}\right)^{1/\alpha},\quad t>0.$$

From (2.14), $\hat{\gamma}_R=g\left(M_{n,R}^{(\alpha)}(k_1)\right)$. Expanding g at the true value $t_0=\Gamma(\alpha+1)\gamma_R^{\alpha}$ using a first-order Taylor expansion gives

$$\hat{\gamma}_R-\gamma_R=g\left(M_{n,R}^{(\alpha)}(k_1)\right)-g(t_0)=g'(t_0)\left(M_{n,R}^{(\alpha)}(k_1)-t_0\right)+R_n,$$

where the remainder satisfies $R_n=o_p\left(\left|M_{n,R}^{(\alpha)}(k_1)-t_0\right|\right)$. Since $M_{n,R}^{(\alpha)}(k_1)-t_0=O_p\left(k_1^{-1/2}\right)$, we have $R_n=o_p\left(k_1^{-1/2}\right)$.

Compute the derivative:

$$g'(t)=\frac{1}{\alpha}\left(\frac{t}{\Gamma(\alpha+1)}\right)^{\frac{1}{\alpha}-1}\cdot\frac{1}{\Gamma(\alpha+1)}=\frac{1}{\alpha\Gamma(\alpha+1)}\left(\frac{t}{\Gamma(\alpha+1)}\right)^{\frac{1}{\alpha}-1}.$$

Evaluating at $t=t_0$ yields

$$g'(t_0)=\frac{1}{\alpha\Gamma(\alpha+1)}\left(\frac{\Gamma(\alpha+1)\gamma_R^{\alpha}}{\Gamma(\alpha+1)}\right)^{\frac{1}{\alpha}-1}=\frac{1}{\alpha\Gamma(\alpha+1)}\left(\gamma_R^{\alpha}\right)^{\frac{1}{\alpha}-1}=\frac{1}{\alpha\Gamma(\alpha+1)}\gamma_R^{1-\alpha}.$$

Therefore,

$$\hat{\gamma}_R - \gamma_R = \frac{1}{\alpha\Gamma(\alpha+1)\gamma_R^{\alpha-1}}\left(M_{n,R}^{(\alpha)}(k_1) - \Gamma(\alpha+1)\gamma_R^{\alpha}\right) + o_p\left(k_1^{-1/2}\right).$$

This is exactly the definition of $H_R(\alpha)$. For the left tail, we replace ε_t by $-\varepsilon_t$ and symmetrically adjust the order statistics indicators to obtain the expression for $H_L(\alpha)$. This completes the proof of the proposition. □

Proposition 4.1 provides a linear expansion for the generalized moment estimator that is structurally identical to that of the Hill estimator (up to a deterministic factor). The multiplier bootstrap consistency theorem of Li *et al.* [2] (Theorem 4.1) relies only on such a linear representation and on the independence between the summands of the left- and right-tail expansions; therefore, it applies directly to the present setting. The only difference is the presence of the factor $\alpha\Gamma(\alpha+1)\gamma_R^{\alpha-1}$ (resp. γ_L) in the influence function; this factor is consistently estimated by replacing γ_R with $\hat{\gamma}_R$, and the bootstrap weights in (4.3)-(4.4) are constructed accordingly. Consequently, the bootstrap procedure described below yields asymptotically valid simultaneous confidence bands for the intermediate scenario.

Following the idea of the multiplier bootstrap in Li *et al.* [2], we use the linear expansion and the asymptotic independence of the left- and right-tail estimators (Proposition 3) to approximate the uncertainty characterised by $\hat{\gamma}_R - \gamma_R$ and $\hat{\gamma}_L - \gamma_L$. Specifically, assume the existence of two independent multiplier sequences $\{\xi_t\}_{t=d_n}^{n}$ and $\{\nu_t\}_{t=d_n}^{n}$ consisting of i.i.d. random variables with zero mean, unit variance, and bounded support, independent of the original data. By multiplying the sums in equations (4.1) and (4.2) by $\{\xi_t\}_{t=d_n}^{n}$ and $\{\nu_t\}_{t=d_n}^{n}$, respectively, we obtain the corresponding bootstrap approximations:

$$H_R(\alpha)^* = \frac{1}{\alpha\Gamma(\alpha+1)\hat{\gamma}_R^{\alpha-1}}\left(\frac{1}{k_1}\sum_{t=d_n}^{n}\left(\mathbb{I}\left(\hat{\varepsilon}_t > \hat{\varepsilon}_{n-k_1,n}\right)\left(\log\frac{\hat{\varepsilon}_t}{\hat{\varepsilon}_{n-k_1,n}}\right)^{\alpha} - \Gamma(\alpha+1)\hat{\gamma}_R^{\alpha}\right)\xi_t\right), \quad (4.3)$$

and

$$H_L(\alpha)^* = \frac{1}{\alpha\Gamma(\alpha+1)\hat{\gamma}_L^{\alpha-1}}\left(\frac{1}{k_2}\sum_{t=d_n}^{n}\left(\mathbb{I}\left(\hat{\varepsilon}_t < \hat{\varepsilon}_{k_2+d_n,n}\right)\left(\log\frac{\hat{\varepsilon}_t}{\hat{\varepsilon}_{k_2+d_n,n}}\right)^{\alpha} - \Gamma(\alpha+1)\hat{\gamma}_L^{\alpha}\right)\nu_t\right). \quad (4.4)$$

We now present the bootstrap inference procedure for the scenario $a_1 > 0$. Under Assumptions 2.1 and 3.1 - 3.5, for any type $\text{type} \in \{\text{U}, \text{D}, \text{R}\}$, the proof of Theorem 3.1 yields the following asymptotic expansions:

$$\begin{aligned} & k_1^{1/2}\log\left[\frac{\widetilde{\text{type-CVaR}}(\tau)}{\text{type-CVaR}(\tau)}\right] \\ & = \mathbb{I}\left(\text{type} \in \{\text{U},\text{R}\}\right)\left[\log\left(k_1/(n\tau)\right)\cdot k_1^{1/2}\left(\hat{\gamma}_R - \gamma_R\right) + \mathcal{X}_R\right] \\ & \quad - \mathbb{I}\left(\text{type} \in \{\text{D},\text{R}\}\right)\left\{\left(k_1/k_2\right)^{1/2}\left[\log\left(k_2/(n\tau)\right)\cdot k_2^{1/2}\left(\hat{\gamma}_L - \gamma_L\right) + \mathcal{X}_L\right]\right\} + o_p(1), \end{aligned}$$

and

$$
\begin{aligned}
& k_1^{1/2}\log\left[\frac{\widetilde{\text{type-CES}(\tau)}}{\text{type-CES}}(\tau)\right] \\
& = \mathbb{I}\left(\text{type}\in\{\text{U},\text{R}\}\right)\left\{\left[\log\left(k_1/(n\tau)\right)+\left(1-\gamma_R\right)^{-1}\right]k_1^{1/2}\left(\hat{\gamma}_R-\gamma_R\right)+\mathcal{X}_R\right\} \\
& \quad -\mathbb{I}\left(\text{type}\in\{\text{D},\text{R}\}\right)\left\{\left(k_1/k_2\right)^{1/2}\left[\log\left(k_2/(n\tau)\right)+\left(1-\gamma_L\right)^{-1}\right]k_2^{1/2}\left(\hat{\gamma}_L-\gamma_L\right)+\mathcal{X}_L\right\} \\
& \quad +o_p(1),
\end{aligned}
$$

where $\tau\in[\tau_l,\tau_u]$; $\mathcal{X}_R\sim N\left(0,\gamma_R^2\right)$, $\mathcal{X}_L\sim N\left(0,\gamma_L^2\right)$, and $\mathcal{X}_R$ and $\mathcal{X}_L$ are independent; $\mathcal{X}_R$ and $\mathcal{X}_L$ are independent of $\hat{\gamma}_R-\gamma_R$ and $\hat{\gamma}_L-\gamma_L$.

We therefore independently draw random samples $\mathcal{X}_R^*$ and $\mathcal{X}_L^*$ from normal distributions with zero mean and variances $\hat{\gamma}_R^2$ and $\hat{\gamma}_L^2$, respectively, and independent of the sequences $\{\xi_t\}_{t=d_n}^n$ and $\{\nu_t\}_{t=d_n}^n$.

$$
\begin{aligned}
& L_{\text{type-CVaR}}(\tau) \\
& = \mathbb{I}\left(\text{type}\in\{\text{U},\text{R}\}\right)\left[\log\left(k_1/(n\tau)\right)k_1^{1/2}H_R(\alpha)^*+\mathcal{X}_R^*\right] \\
& \quad -\mathbb{I}\left(\text{type}\in\{\text{D},\text{R}\}\right)\left\{\left(k_1/k_2\right)^{1/2}\left[\log\left(k_2/(n\tau)\right)k_2^{1/2}H_L(\alpha)^*+\mathcal{X}_L^*\right]\right\},
\end{aligned} \tag{4.5}
$$

where $H_R(\alpha)^*$ and $H_L(\alpha)^*$ are defined in (4.3) and (4.4).

$L_{\text{type-CVaR}}(\tau)$ is used to estimate $k_1^{1/2}\left\{\log\left[\widetilde{\text{type-CVaR}}(\tau)/\text{type-CVaR}(\tau)\right]\right\}$.

Similarly, for CES we have

$$
\begin{aligned}
& L_{\text{type-CES}}(\tau) \\
& = \mathbb{I}\left(\text{type}\in\{\text{U},\text{R}\}\right)\left\{\left[\log\left(k_1/(n\tau)\right)+\left(1-\hat{\gamma}_R\right)^{-1}\right]k_1^{1/2}H_R(\alpha)^*+\mathcal{X}_R^*\right\} \\
& \quad -\mathbb{I}\left(\text{type}\in\{\text{D},\text{R}\}\right)\left(k_1/k_2\right)^{1/2}\left\{\left[\log\left(k_2/(n\tau)\right)+\left(1-\hat{\gamma}_L\right)^{-1}\right]k_2^{1/2}H_L(\alpha)^*+\mathcal{X}_L^*\right\}.
\end{aligned} \tag{4.6}
$$

$L_{\text{type-CES}}(\tau)$ is used to estimate $k_1^{1/2}\left\{\log\left[\widetilde{\text{type-CES}}(\tau)/\text{type-CES}(\tau)\right]\right\}$.

Bootstrap Algorithm under $a_1>0$

1) **Model specification and residual extraction**

Choose an appropriate dynamic model and estimate it to obtain the residual series $\{\hat{\varepsilon}_t\}$. Compute the tail index estimates $\hat{\gamma}_R$ and $\hat{\gamma}_L$, as well as the maximum absolute log-ratios defined in (2.8) and (2.9).

2) **Generate random perturbations and construct statistics**

Generate independent i.i.d. random sequences $\{\xi_t\}$ and $\{\nu_t\}$ from a distribution with zero mean and unit variance, and compute $H_R(\alpha)^*$ and $H_L(\alpha)^*$. Then draw $\mathcal{X}_R^*$ and $\mathcal{X}_L^*$ from normal distributions with zero mean and variances $\hat{\gamma}_R^2$ and $\hat{\gamma}_L^2$, respectively. For $\tau\in[\tau_l,\tau_u]$, compute $L_{\text{type-CVaR}}(\tau)$ and $L_{\text{type-CES}}(\tau)$.

3) **Resampling to build empirical quantiles**

Take a sufficiently large integer B, repeat step 2 B times to obtain $\left\{L_{\text{type-CVaR}}(\tau)\right\}_{b=1}^B$ and $\left\{L_{\text{type-CES}}(\tau)\right\}_{b=1}^B$. Then, for a prescribed significance level $\alpha\in(0,1)$, compute $c_{\text{type},1-\alpha}^{*1}$ and $c_{\text{type},1-\alpha}^{*2}$, which are the $(1-\alpha)$ sample quantiles of $\left\{\sup_{\tau\in[\tau_l,\tau_u]}\left|L_{\text{type-CVaR}}(\tau)\right|\right\}_{b=1}^{B}$ and $\left\{\sup_{\tau\in[\tau_l,\tau_u]}\left|L_{\text{type-CES}}(\tau)\right|\right\}_{b=1}^{B}$, respectively.

4) **Construct asymptotic simultaneous confidence bands (SCBs)**

For $\text{type}\in\{\text{U},\text{D},\text{R}\}$, the $(1-\alpha)$ asymptotic SCBs for $\text{type-CVaR}(\tau)$ and $\text{type-CES}(\tau)$ are respectively:

- **Confidence interval for CVaR:**

$$\left[\widetilde{\text{type-CVaR}}(\tau)\exp\left(-c^{*1}_{\text{type},1-\alpha}k_1^{-1/2}\right),\widetilde{\text{type-CVaR}}(\tau)\exp\left(c^{*1}_{\text{type},1-\alpha}k_1^{-1/2}\right)\right]$$

- **Confidence interval for CES:**

$$\left[\widetilde{\text{type-CES}}(\tau)\exp\left(-c^{*2}_{\text{type},1-\alpha}k_1^{-1/2}\right),\widetilde{\text{type-CES}}(\tau)\exp\left(c^{*2}_{\text{type},1-\alpha}k_1^{-1/2}\right)\right]$$

The validity of this bootstrap method follows from Theorem 4.1 of Li *et al.* [2], which guarantees that the bootstrap statistic shares the same limiting distribution as the original statistic under the intermediate scenario.

5. Numerical Study

This section first presents a specific criterion for selecting the effective sample sizes k_1 and k_2 to implement our method, and then reports the results of Monte Carlo simulations.

5.1. Data-Driven Selection of k_1 and k_2

The implementation of the proposed method requires choosing the effective sample sizes k_1 and k_2. Studies on risk measure estimation based on extreme value theory (EVT) have confirmed that the choice of the effective sample size is crucial (e.g., Chan *et al.* [8]; Martins-Filho *et al.* [13]). Several empirical suggestions have been made. For instance, when constructing pointwise confidence intervals for downside CVaR, Chan *et al.* [8] suggested using $\left\lfloor 1.5(\log n)^2\right\rfloor$ (where $\lfloor\cdot\rfloor$ denotes the integer part) for a sample size $n=1000$. However, this suggestion was based on their simulation experiments with $n=1000$, and Spierdijk [14] found that it is no longer suitable for larger sample sizes. Compared with such ad-hoc suggestions, a data-driven rule is more desirable because it selects k_1 and k_2 according to the specific features of the data at hand.

In this section, we adopt the data-driven approach of Li *et al.* [2] for selecting k_1 and k_2. They follow the idea of Danielsson *et al.* [15] and use the maximum distance between the fitted Pareto-type tail and the empirical quantile to measure the approximation quality, choosing the effective sample size that minimises this measure as the optimal one. Since the estimation of conditional value at risk (CVaR) is intrinsically linked to quantile estimation, this quantile-based method is suitable for CVaR estimation. Our estimation procedure relies on a Pareto approximation to the tail of the distribution of ε, and k_1 and k_2 essentially determine where this approximation starts on the right and left tails, respectively. Therefore, it is reasonable to choose the effective sample sizes that optimise the

Pareto fit.

Following the ideas of Danielsson *et al.* [15] and Li *et al.* [2], for estimating the upper CVaR, k_1 is chosen as

$$k_1^* = \arg \min_{l=k_{\min},\cdots,k_{\max}} \left\{ \max_{j=1,\cdots,k_{\max}} \left[\left| \hat{\varepsilon}_{n-j,n} - \hat{\varepsilon}_{n-l+1,n} \left(j/l \right)^{-\hat{\gamma}_R(l)} \right| \right] \right\}, \quad (5.1)$$

where $1 \le k_{\min} < k_{\max} \le n$, and $\hat{\gamma}_R(l)$ is the estimator of γ_R with $k_1 = l$. Here $\hat{\varepsilon}_{n-j,n}$ is the empirical quantile, and $\hat{\varepsilon}_{n-l+1,n}(j/l)^{-\hat{\gamma}_R}$ is the Pareto-based approximate quantile estimator. Similarly, for estimating the lower CVaR, the effective sample size k_2 is chosen as k_2^* by substituting $-\hat{\varepsilon}_t$ into (5.1).

For CES estimation, a similar approach based on the tail mean is used to modify formula (5.1). Specifically, when estimating the upper CES, k_1 is chosen as

$$k_1^* = \arg \min_{l=k_{\min},\cdots,k_{\max}} \left\{ \max_{j=1,\cdots,k_{\max}} \left[\left| \frac{\sum_{t=d_n}^{n} \mathbb{I}\left(\hat{\varepsilon}_t > \hat{\varepsilon}_{n-j,n}\right)\hat{\varepsilon}_t}{\sum_{t=d_n}^{n} \mathbb{I}\left(\hat{\varepsilon}_t > \hat{\varepsilon}_{n-j,n}\right)} - \frac{\hat{\varepsilon}_{n-l+1,n}\left(j/l\right)^{-\hat{\gamma}_R(l)}}{1-\hat{\gamma}_R(l)} \right| \right] \right\}, \quad (5.2)$$

where $\frac{\sum_{t=d_n}^{n} \mathbb{I}\left(\hat{\varepsilon}_t > \hat{\varepsilon}_{n-j,n}\right)\hat{\varepsilon}_t}{\sum_{t=d_n}^{n} \mathbb{I}\left(\hat{\varepsilon}_t > \hat{\varepsilon}_{n-j,n}\right)}$ is the empirical tail mean at the $(1-j/n)$-quantile, and $\frac{\hat{\varepsilon}_{(l)}\left(j/l\right)^{-\hat{\gamma}_R(l)}}{1-\hat{\gamma}_R(l)}$ is the tail mean estimator based on the Pareto approximation. Similarly, for estimating the lower CES, the effective sample size k_2 is chosen as $\hat{k}_2^*$ by replacing $\{\hat{\varepsilon}_t\}$ with $\{-\hat{\varepsilon}_t\}$ in formula (5.2).

Finally, it should be noted that Assumptions 3.5(ii) and (iii) impose constraints on the selection of k_1 and k_2. We now provide specific suggestions that satisfy these conditions. First, k_1 and k_2 are both taken from the same interval $[k_{\min}, k_{\max}]$, which ensures the comparability required by Assumption 3.5(ii). Second, in the case $a_1 > 0$, k_1 and k_2 should be comparable to $n\tau$. To this end, we set $k_{\min} = an\tau_u$ and $k_{\max} = bn\tau_u$ with $0 < a < b$. For example, if the tail interval is $[0.5\%, 1\%]$, we may take $k_{\min} = 2\% \times n$ and $k_{\max} = 15\% \times n$, corresponding to $a = 2$ and $b = 15$. Moreover, this choice of $k_{\max}$ means that at most 15% of the sample is used, satisfying Assumption 3.5(iii) that $k_{\max}$ is small relative to n^{υ_0} with υ_0 typically equal or close to 1. Under the above selection rules, $k_1^{1/2}$ is typically larger than $\left|\log\left(n\tau_l/k_1\right)\right|$, so the limiting condition $\lim_{n\to\infty} k_1^{-1/2} \log\left(n\tau_l/k_1\right) = 0$ in Assumption 3.5(iii) holds automatically.

In practice, the simulation results in Section 5.2 demonstrate that the performance of the SCBs is highly robust to the choice of α. Hence, the user may simply fix $\alpha = 1$, which recovers the classical Hill estimator and avoids additional tuning. If a small improvement in efficiency is desired, any value in $[0.5, 2]$ yields similar coverage and length. The effective sample sizes k_1 and k_2 selected by the data-driven procedure above are then used for both CVaR and CES estimation; that is, the same k_1 (resp. k_2) is employed for the right-tail (resp. left-tail) risk measures. This ensures that the risk measures are based on the same tail subsample and facilitates coherent joint inference.

5.2. Monte Carlo Simulations

This section evaluates the finite-sample performance of the proposed method through a series of Monte Carlo simulations. To facilitate the comparison of model performance after changing the extreme value index, we follow the simulation design of Li *et al.* [2]. In the bootstrap implementation, the number of bootstrap replications is set to $B = 1000$, and the multipliers are drawn from a two-point distribution with $P(\xi = 1) = P(\xi = -1) = 1/2$. The confidence level is fixed at 95%. The number of discarded observations is defined as $d_n = \lfloor 5 \times (n\tau_l)^{1/3} \rfloor$, which satisfies $d_n = o(k_1^{1/2})$ to ensure Assumption 3.5 holds. Note that in the case $a_1 > 0$, k_1 is of the same order as $n\tau_l$. Therefore, setting $d_n = \lfloor 5 \times (n\tau_l)^{1/3} \rfloor$ satisfies the required order conditions for d_n. All simulation results reported are based on 1000 independent replications.

The data are generated from a GARCH(1,1) model:

$$R_t = \sigma_t \varepsilon_t, \ \ \sigma_t^2 = \omega + \omega_1 R_{t-1}^2 + \omega_2 \sigma_{t-1}^2, \tag{5.3}$$

with $\omega = 0.001$, $\omega_1 = 0.04$, $\omega_2 = 0.85$. The innovations $\varepsilon_t \sim \text{sstd}(\nu, \lambda)$; see Fernández and Steel [16], where the distribution is characterised by a shape parameter ν and a skewness parameter $\lambda = 1$.

We compare the finite-sample performance of simultaneous confidence bands (SCBs). The tail interval is $[0.5\%, 1\%]$ and the sample sizes are $n \in \{500, 1000, 1500\}$. The effective sample sizes k_1 and k_2 are selected using the data-driven method described in Section 5.1. We set $k_{\min} = 2\% \times n$ and $k_{\max} = 15\% \times n$; The confidence level is 95%, and the relative length is defined as the ratio of the upper bound to the lower bound of the SCBs.

In terms of parameter robustness, the choice of the moment order α has no significant impact on the performance of the proposed method. Across all sample sizes and risk measure specifications, the variation in coverage probabilities with respect to α does not exceed 0.05; notably, the variation in coverage probability for the core focus, the relative risk measure R-CVaR, is less than 0.015. Meanwhile, the variation in the relative length of confidence bands with respect to α is less than 0.03 for all specifications. This result verifies the robustness of the generalized moment-based tail index estimator, indicating that the proposed method does not require complex parameter tuning for α and exhibits strong practical applicability.

Regarding large-sample asymptotic properties, the experimental results are fully consistent with the theoretical conclusions of Theorem 3.1 in this paper. As the sample size increases from 500 to 1500, the coverage probabilities of all risk measures converge consistently towards the 95% nominal confidence level. Specifically, the average coverage probability of R-CVaR converges gradually from 0.957 to 0.948, and its deviation from the nominal level narrows from 0.007 to 0.002; the over-coverage issue of R-CES is also continuously alleviated as the sample size increases. Meanwhile, the relative length of the confidence bands decreases monotonically as the sample size grows. This behavior is consistent with the $k_1^{-1/2}$ rate implied by the asymptotic theory, since the effective tail sample size k_1 increases with n. The average relative length of R-CVaR decreases from 2.045 to 1.521, representing a 25.6% reduction, while that of R-CES decreases from 3.111 to 1.970, a 36.7% reduction, indicating a substantial improvement in the estimation precision of the method for large samples.

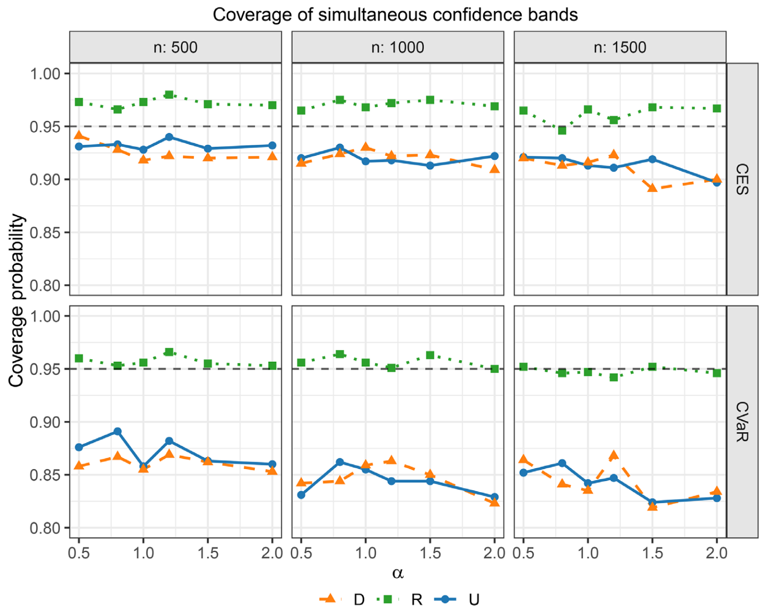

Figure 1. Relationship between CVaR, CES coverage and α with sample sizes of 500, 1000, and 1500.

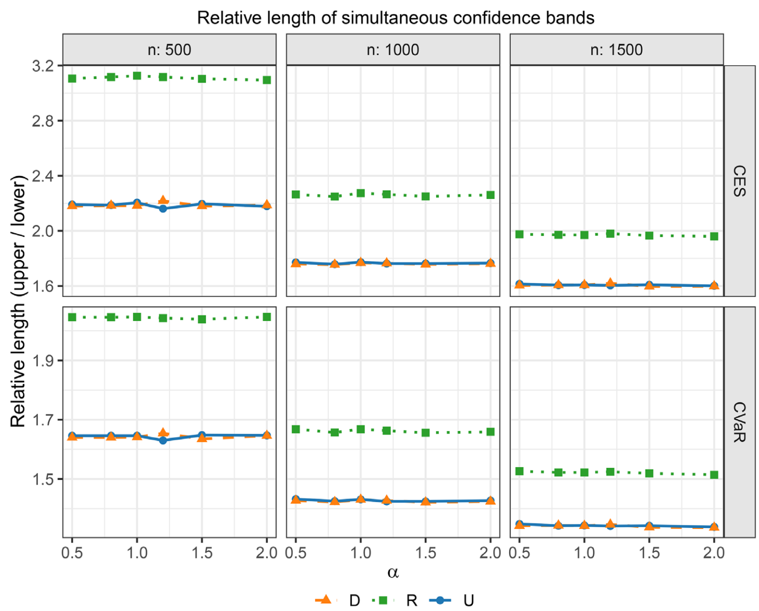

Figure 2. Relationship of relative lengths of CVaR and CES versus α for sample sizes of 500, 1000, and 1500.

Table 1. Coverage and relative length under different values of α ($n = 500$, $\varepsilon_t \sim \text{sstd}(5,1)$).

α	Coverage of CVaR			Coverage of CES			Relative length of CVaR			Relative length of CES		
(lr)2-4 (lr)5-7 (lr)8-10 (lr)11-13	U	D	R	U	D	R	U	D	R	U	D	R
0.5	0.876	0.858	0.960	0.931	0.941	0.973	1.646	1.640	2.046	2.191	2.180	3.106
0.8	0.891	0.867	0.953	0.933	0.928	0.966	1.646	1.640	2.046	2.187	2.184	3.117
1.0	0.858	0.855	0.956	0.928	0.918	0.973	1.646	1.641	2.047	2.205	2.183	3.126
1.2	0.882	0.869	0.966	0.940	0.922	0.980	1.630	1.654	2.043	2.161	2.218	3.117
1.5	0.863	0.862	0.955	0.929	0.920	0.971	1.648	1.635	2.039	2.196	2.180	3.104
2.0	0.860	0.853	0.953	0.932	0.921	0.970	1.647	1.646	2.047	2.179	2.187	3.095

Table 2. Coverage and relative length under different values of α ($n = 1000$, $\varepsilon_t \sim \text{sstd}(5,1)$).

α	Coverage of CVaR			Coverage of CES			Relative length of CVaR			Relative length of CES		
(lr)2-4 (lr)5-7 (lr)8-10 (lr)11-13	U	D	R	U	D	R	U	D	R	U	D	R
0.5	0.831	0.842	0.956	0.920	0.915	0.965	1.432	1.427	1.668	1.771	1.759	2.264
0.8	0.862	0.844	0.964	0.930	0.924	0.975	1.425	1.422	1.657	1.758	1.755	2.249
1.0	0.855	0.859	0.956	0.917	0.930	0.968	1.431	1.430	1.668	1.773	1.768	2.274
1.2	0.844	0.863	0.951	0.918	0.922	0.972	1.424	1.427	1.663	1.763	1.765	2.265
1.5	0.844	0.850	0.963	0.913	0.923	0.975	1.424	1.421	1.656	1.763	1.756	2.250
2.0	0.829	0.823	0.950	0.922	0.909	0.969	1.427	1.424	1.659	1.766	1.762	2.261

Table 3. Coverage and relative length under different values of α ($n = 1500$, $\varepsilon_t \sim \text{sstd}(5,1)$).

α	Coverage of CVaR			Coverage of CES			Relative length of CVaR			Relative length of CES		
(lr)2-4 (lr)5-7 (lr)8-10 (lr)11-13	U	D	R	U	D	R	U	D	R	U	D	R
0.5	0.852	0.864	0.952	0.921	0.920	0.965	1.348	1.342	1.526	1.615	1.604	1.975
0.8	0.861	0.841	0.946	0.920	0.913	0.946	1.342	1.343	1.522	1.606	1.607	1.971
1.0	0.842	0.835	0.947	0.913	0.916	0.966	1.343	1.342	1.522	1.607	1.607	1.970
1.2	0.847	0.868	0.942	0.911	0.923	0.956	1.341	1.346	1.524	1.604	1.619	1.980
1.5	0.824	0.819	0.952	0.919	0.891	0.968	1.342	1.337	1.519	1.608	1.597	1.966
2.0	0.828	0.834	0.946	0.897	0.900	0.967	1.338	1.336	1.514	1.601	1.599	1.960

In terms of performance across different risk types, the inference performance of the proposed method for relative risk measures is significantly superior to that for one-sided upside and downside risk measures. Across all sample sizes, the coverage probabilities of R-CVaR and R-CES remain stable in the interval $[0.94, 0.98]$, which is extremely close to the 95% nominal confidence level. In contrast, one-sided upside (U) and downside (D) risks exhibit mild systematic under-coverage. This result is consistent with the theoretical derivation in this paper: relative risk measures incorporate estimation information from both the right and left tails, effectively offsetting the bias in single-tail estimation, making them the core advantageous application scenario of the proposed method. A comparison of different risk measures shows that CES generally outperforms CVaR in terms of coverage accuracy, with coverage probabilities of CES up to 7% higher than those of CVaR in one-sided scenarios. However, since CES, as a tail conditional expectation, inherently has higher estimation uncertainty than quantile-based CVaR, the relative length of its confidence bands is also significantly larger than that of CVaR. The proposed method accurately captures the differences in the statistical properties of the two types of risk measures.

Overall, the trends presented in **Figure 1** and **Figure 2** are fully consistent with the numerical results in **Tables 1-3**. The bootstrap simultaneous confidence bands (SCBs) method proposed in this paper exhibits excellent finite-sample performance, strong parameter robustness, and large-sample convergence properties consistent with theoretical expectations, providing a reliable statistical tool for the joint inference of conditional risk measures.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

[1] Artzner, P. (1999) Application of Coherent Risk Measures to Capital Requirements in Insurance. *North American Actuarial Journal*, **3**, 11-25. https://doi.org/10.1080/10920277.1999.10595795

[2] Li, S., Peng, L. and Song, X. (2023) Simultaneous Confidence Bands for Conditional Value-at-Risk and Expected Shortfall. *Econometric Theory*, **39**, 1009-1043. https://doi.org/10.1017/s0266466622000275

[3] McNeil, A.J. and Frey, R. (2000) Estimation of Tail-Related Risk Measures for Heteroscedastic Financial Time Series: An Extreme Value Approach. *Journal of Empirical Finance*, **7**, 271-300. https://doi.org/10.1016/s0927-5398(00)00012-8

[4] Hill, B.M. (1975) A Simple General Approach to Inference about the Tail of a Distribution. *The Annals of Statistics*, **3**, 1163-1174. https://doi.org/10.1214/aos/1176343247

[5] Dekkers, A.L.M., Einmahl, J.H.J. and Haan, L.D. (1989) A Moment Estimator for the Index of an Extreme-Value Distribution. *The Annals of Statistics*, **17**, 1833-1855. https://doi.org/10.1214/aos/1176347397

[6] Gomes, M.I. and Martins, M.J. (2001) Generalizations of the Hill Estimator—Asymptotic versus Finite Sample Behaviour. *Journal of Statistical Planning and Inference*, **93**, 161-180. https://doi.org/10.1016/s0378-3758(00)00201-9

[7] Segers, J. (2001) Residual Estimators. *Journal of Statistical Planning and Inference*, **98**, 15-27. https://doi.org/10.1016/s0378-3758(00)00321-9

[8] Hang Chan, N., Deng, S., Peng, L. and Xia, Z. (2007) Interval Estimation of Value-At-Risk Based on GARCH Models with Heavy-Tailed Innovations. *Journal of Econometrics*, **137**, 556-576. https://doi.org/10.1016/j.jeconom.2005.08.008

[9] Hoga, Y. (2019) Confidence Intervals for Conditional Tail Risk Measures in ARMA-GARCH Models. *Journal of Business & Economic Statistics*, **37**, 613-624. https://doi.org/10.1080/07350015.2017.1401543

[10] de Haan, L. and Ferreira, A. (2006) Extreme Value Theory: An Introduction. Springer.

[11] Pan, X., Leng, X. and Hu, T. (2013) The Second-Order Version of Karamata's Theorem with Applications. *Statistics & Probability Letters*, **83**, 1397-1403. https://doi.org/10.1016/j.spl.2013.02.006

[12] Rényi, A. (1953) On the Theory of Order Statistics. *Acta Mathematica Academiae Scientiarum Hungaricae*, **4**, 191-231. https://doi.org/10.1007/bf02127580

[13] Martins-Filho, C., Yao, F. and Torero, M. (2018) Nonparametric Estimation of Conditional Value-at-Risk and Expected Shortfall Based on Extreme Value Theory. *Econometric Theory*, **34**, 23-67. https://doi.org/10.1017/s0266466616000517

[14] Spierdijk, L. (2016) Confidence Intervals for ARMA-GARCH Value-at-Risk: The Case of Heavy Tails and Skewness. *Computational Statistics & Data Analysis*, **100**, 545-559. https://doi.org/10.1016/j.csda.2014.08.011

[15] Danielsson, J., Ergun, L.M., de Haan, L. and de Vries, C.G. (2016) Tail Index Estimation: Quantile Driven Threshold Selection. *Computational Statistics and Data Analysis*, **100**, 545-559.

[16] Fernández, C. and Steel, M.F.J. (1998) On Bayesian Modeling of Fat Tails and Skewness. *Journal of the American Statistical Association*, **93**, 359-371. https://doi.org/10.1080/01621459.1998.10474117