Statistica Sinica Preprint No: SS-2024-0418	
Title	Tail Risk Equivalent Level Transition And Its Application for Estimating Extreme Lp-Quantiles
Manuscript ID	SS-2024-0418
URL	http://www.stat.sinica.edu.tw/statistica/
DOI	10.5705/ss.202024.0418
Complete List of Authors	Qingzhao Zhong and Yanxi Hou
Corresponding Authors	Yanxi Hou
E-mails	yxhou@fudan.edu.cn
Notice: Accepted author version.	

TAIL RISK EQUIVALENT LEVEL TRANSITION AND ITS APPLICATION FOR ESTIMATING EXTREME L_p-QUANTILES

Qingzhao Zhong and Yanxi Hou

School of Data Science, Fudan University, Shanghai, China.

Abstract: L_p-quantiles, as a generalization of Value-at-Risk and expectile, have gained increasing attention in risk management due to their feasibility and straightforwardness in statistical implementation. This paper introduces the concept of Tail Risk Equivalent Level Transition (TRELT) to capture changes in tail risk when transitioning between two L_p-quantiles. Motivated by PELVE in Li and Wang (2023) but tailored for tail risk, we investigate the theoretical properties of TRELT, including its existence, uniqueness, and asymptotic behavior. Additionally, we develop inference methods for TRELT and extreme L_p-quantiles using this risk transition, which serves as a novel extrapolation technique in extreme value theory. Simulation studies and real data analysis demonstrate the empirical performance of these methods.

Key words and phrases: Extreme value theory; Heavy-tailed data; L_p-quantile; TRELT.

Corresponding author: Yanxi Hou, yxhou@fudan.edu.cn.

1. Introduction

In recent years, a significant issue has emerged in the field of risk management: the transition between different risk measures. The central challenge is to develop a robust model that can accurately describe the shift from one risk measure to another. According to the Basel Committee on Banking Supervision (BCBS (2019)), banks are required to transition from Value-at-Risk (VaR) to Expected Shortfall (ES) as a more reliable metric for risk quantification. This recommendation is based on ES's coherence as a risk measure (Artzner et al. (1999)) and its enhanced capacity to capture tail risk. The risk transition naturally leads to an important question: how do different risk measures transition and subsequently influence regulatory frameworks, particularly in terms of their relative conservatism?

To characterize the equivalent transition between VaR and ES while ensuring consistency in risk quantification, Li and Wang (2023) introduced the concept of Probability Equivalent Level of VaR and ES (PELVE), which identifies the balancing point for this transition. For a random loss X and a (tail) risk level $\varepsilon \in (0,1)$, PELVE is essentially a constant multiplier $c \in [1, 1/\varepsilon]$ such that,

$$\mathrm{ES}_{1-c\varepsilon}(X) = \mathrm{VaR}_{1-\varepsilon}(X), \tag{1.1}$$

with $\mathrm{VaR}_\tau(X) = \inf_{x\in\mathbb{R}} \left\{ x \;\middle|\; \mathbb{P}(X \le x) \ge \tau \right\}$ and $\mathrm{ES}_\tau(X) = \frac{1}{1-\tau}\int_\tau^1 \mathrm{VaR}_q(X)\,dq$, where $\tau \in (0,1)$ is a confidence level. Li and Wang (2023) has also demonstrated that the existence and uniqueness of PELVE can be ensured under fairly general conditions. This finding has important practical implications. Specifically, PELVE effectively illustrates the increase in capital requirements needed to manage potential risks when ES is used to replace VaR, as exemplified by the inequality $\mathrm{ES}_{0.975} > \mathrm{VaR}_{0.99}$. Additionally, Asimit et al. (2019) developed an ES-based methodology for quantile capital allocation by utilizing the relationship given in (1.1) to adjust the confidence level. Along this line of research, Fiori and Rosazza Gianin (2023) extended the concept of PELVE by establishing equivalent transitions between various risk measures. Similarly, Ortega-Jiménez et al. (2024) introduced the concept of PELCoV by extending this transition to CoVaR and VaR.

The tail risk transition between a couple of risk measures post a novel challenge for statistical methodology on tail risk. In addition to PELVE, another important transition is that between the quantile and the expectile, both of which belong to the class of L_p-quantiles. Some recent studies have already adopted a certain transition mechanism between L_p-quantiles. For instance, Daouia, Girard and Stupfler (2019) examined the transition from quantile to L_p-quantile, specifically reflected in the extreme L_p-quantile extrapolation from quantile; Zou (2014) characterized distributions for which the expectile at level $\omega(\tau)$ corresponds the τ-quantile, for a suitable function $\omega(\cdot)$ while Xu et al. (2022) considered this equivalent transition within a regression framework; Bignozzi, Merlo and Petrella (2024) provided a more substantial conclusion that for a Student's t distribution with $\nu \in [1, \infty)$ degrees of freedom the $L_{\nu-j}$-quantile and L_{j+1}-quantile always coincide for any $j \in [0, \nu-1]$. However, the theoretical underpinnings of these transition have not been explored well, leaving the statistical methodology unguaranteed for tail risks. For example, the existence and uniqueness of these transitions are unknown, and thus the statistical methods proposed based on extreme value theory suffer some deficiencies, such as biased problems.

This paper explores the tail risk transition mechanism between L_p-quantile and L_q-quantile ($p > q \geq 1$), introducing the concept of Tail Risk Equivalent Level Transition (TRELT). L_p-quantiles are more informative than quantiles as they incorporate higher order partial moments and tail information, whereas quantiles only indicate whether observations are above or below the predictor. Moreover, L_p-quantiles offer a more flexible framework than quantiles and expectiles, capturing different distribution characteristics and handling complex data, especially heteroscedasticity. The motivation for this tail risk transition is threefold. Firstly, Daouia, Girard and Stupfler (2019) showed that L_p-quantiles with $p \in (1, 2)$ can cover a broad class of heavy-tailed distributions while maintaining robustness and effectiveness, making them advantageous for tail risk analysis. Secondly, existing studies on L_p-quantile-based risk transition are mostly limited to specific cases like $p = 2$ or 1, lacking a general theoretical framework. Thirdly, this transition enables efficient online predictions for extreme L_p-quantiles with multiple values of p by extrapolating from a specific L_q-quantile, reducing the need for frequent optimization. This will facilitate real-time risk management, improve prediction accuracy, and lower computational costs, with significant applications in insurance and risk management.

The L_p-quantile $\theta_p(\tau)$ is defined as

$$\theta_p(\tau) := \arg\min_{u\in\mathbb{R}} E(\rho_{p,\tau}(X - u)), \tag{1.2}$$

where X is a random variable with a distribution F, p is any given order greater than 1, τ is a risk level in $(0, 1)$, and $\rho_{p,\tau}(s)$ is a p-power loss function such that

$$\rho_{p,\tau}(s) = \left|\tau - 1_{\{s\le 0\}}\right| \cdot |s|^p = \tau s_+^p + (1 - \tau)s_-^p,$$

with $s_+ = \max\{s, 0\}$ and $s_- = \max\{-s, 0\}$. Then, TRELT addresses the tail risk transition between L_p- and L_q-quantiles by considering

$$\theta_p(1 - c\varepsilon) = \theta_q(1 - \varepsilon), \quad \varepsilon \in (0, 1 - \tau_0), \tag{1.3}$$

where c is a certain constant determined by p, q, ε, and τ_0 is a threshold for tail region. TRELT builds a bridge between two levels given tail equivalent risk measures, paving the way for intriguing new statistical inference problems for tail risk measures, especially the estimation problem of c in (1.3). The statistical inference methods of L_p-quantiles are well established in the literature, but they cannot address the probelms arised by TRELT given (1.3), without studying other properies of TRELT like existence and uniqueness of c. Later on, we will see that one useful property derived from (1.2) such that

$$\tau = \frac{E((X-\theta_p(\tau))_-^{p-1})}{E(|X-\theta_p(\tau)|^{p-1})}, \tag{1.4}$$

will be used for constructing the explicit estimators for TRELT. Given samples $X_1, ..., X_n$, at a fixed or an intermediate level, the estimator for $\theta_p(\tau)$ via solving the empirical form of (1.2) is

$$\hat{\theta}_p(\tau) = \arg\min_{u\in\mathbb{R}} \frac{1}{n}\sum_{i=1}^{n} |\tau - \mathbb{1}_{\{X_i \le u\}}| \cdot |X_i - u|^p. \tag{1.5}$$

When we consider an extreme level, due to the lack of sufficient samples on the tail, (1.5) may lead to an ineffective estimator and thus extrapolative technique must be employed. Daouia, Girard and Stupfler (2019) further put up a standard extrapolative estimator for $\theta_p(1-\varepsilon_n')$, given by

$$\tilde{\theta}_p^{\text{sta}}(1-\varepsilon_n') = \left(\frac{\varepsilon_n'}{\varepsilon_n}\right)^{-\hat{\gamma}} \hat{\theta}_p(1-\varepsilon_n), \tag{1.6}$$

where $\varepsilon_n, \varepsilon_n'$ are intermediate and extreme (tail) levels, and $\hat{\gamma}$ is a suitable estimator for extreme value index γ (see Assumption 1 below). For $p = 1$, it could be taken $X_{n-[n\varepsilon_n],n}$ as $\hat{\theta}_1(1-\varepsilon_n)$. Hence, Daouia, Girard and Stupfler (2019) provided the other extrapolative estimator,

$$\tilde{\theta}_p^{\text{qua}}(1-\varepsilon_n') = \left[\frac{\hat{\gamma}}{B(p, \hat{\gamma}^{-1} - p + 1)}\right]^{-\hat{\gamma}} \left(\frac{\varepsilon_n'}{\varepsilon_n}\right)^{-\hat{\gamma}} X_{n-[n\varepsilon_n],n}, \tag{1.7}$$

where $B(a,b) = \int_0^1 t^{a-1}(1-t)^{b-1}\,dt$ denotes the Beta function. Since our proposed method (4.38) will reduce to (1.7) when $q = 1$, we thus will conduct (1.6) as a benchmark in our simulation. For other studies on statistical inference of L_p-quantiles, refer to Daouia, Girard and Stupfler (2018); Girard, Stupfler and Usseglio-Carleve (2021, 2022) etc.

This paper makes two significant contributions. First, it establishes the mathematical conditions and properties of (1.3). It is crucial to highlight that (1.3) may only be meaningful in the one-sided tail region rather than across the entire uncertainty spectrum. This is attributed to the fact that the inequality $\theta_p(\tau) > \theta_q(\tau)$ is valid solely when $\tau \in (\tau_0, 1)$, as elaborated in Proposition 1. Under this specific condition, the definition of (1.3) is functional. Despite the apparent limitations in the definition of TRELT, it is adequate for the study of extreme risks, thereby anchoring our research in the extreme value theory for heavy-tailed distributions. Second, TRELT introduces several novel approaches to estimate extreme

L_p-quantiles, which prove to be more innovative and effective in characterizing uncertainty compared to traditional methods. Our TRELT-based extrapolation offers several improvements over (1.6) and (1.7). Firstly, our TRELT-based estimators are seamlessly integrated into the TRELT estimators themselves, introducing a novel form of uncertainty that can potentially offer a more nuanced depiction of risk uncertainty as risk measures fluctuate. Secondly, within our framework, $\theta_p(1-\varepsilon_n')$ can be estimated effectively from any $\theta_q(1-\varepsilon_n)$ via TRELT under $1-q < p - \frac{1}{2\gamma} < 1$, not just from $\theta_p(1-\varepsilon_n)$ or $\theta_1(1-\varepsilon_n)$. This implies that we can glean valuable insights about $\theta_p(1-\varepsilon_n')$ from $\theta_q(1-\varepsilon_n)$ through the method of tail risk transition, a feat that neither (1.6) nor (1.7) can achieve. Thirdly, simulation evidence corroborates the superior empirical performance of our methods, particularly in terms of Mean Square Relative Error (MSRE).

The remainder is organized as follows. The construction of TRELT for $\theta_p(\tau)$ and $\theta_q(\tau)$ is organized in Section 2, where we also discuss the existence and uniqueness. Sections 3 and 4 provide the statistical methodologies for the coefficient of TRELT and the extreme L_p-quantiles via TRELT, respectively. We conduct a series of simulation studies and real data analysis to illustrate the performance of our proposed estimators in Sections 5 - 6.

2. Tail Risk Equivalent Level Transition Between L_p-quantiles

Recall that given two orders p, q satisfy $p > q \geq 1$, TRELT between L_p- and L_q-quanitles via a coefficient c is given by

$$\theta_p(1 - c\varepsilon) = \theta_q(1 - \varepsilon). \tag{2.8}$$

However, it might fail to achieve such a transition for all $\varepsilon \in (0, 1)$. The main reason is that it is not clear about the size of $\theta_p(\tau)$ and $\theta_q(\tau)$; consequently, it is infeasible to find the range of the coefficient c because we are uncertain whether (2.8) has a solution c, or even if it does, its uniqueness is unknown. Thus, it needs more information about the relative sizes between $\theta_p(\tau)$ and $\theta_q(\tau)$, and the study of the existence and uniqueness of c is necessary. Fortunately, Proposition 1 provides a helpful interpretation through a limit relation over them. We can claim that $\theta_p(\tau) > \theta_q(\tau)$ (or $\theta_p(\tau) < \theta_q(\tau)$) always holds on the tail region $(\tau_0, 1)$ (or $(\tau_0', 1)$) for a certain threshold τ_0. This suggests that the risk equivalent level transition could be established on the tail, which is sufficient and feasible for predicting extreme risks in risk management.

Recall X is a random variable with a distribution F, then we denote $\overline{F} = 1 - F$ and U as the survival function of F and the left-continuous inverse of $1/\overline{F}$, (*i.e.* tail quantile function) respectively. In this paper, we study the right tail of F for $\theta_p(\tau)$ with a risk level τ close to 1. Assumption 1 states a first-order regular variation condition for the right tail of F.

Assumption 1 (First-order regular variation)**.** *The function $\overline{F}$ satisfies a (first-order) regular variation condition with an extreme value index $\gamma > 0$, i.e., for all $x > 0$,*

$$\lim_{t\to\infty} \frac{\overline{F}(tx)}{\overline{F}(t)} = x^{-\frac{1}{\gamma}}. \tag{2.9}$$

Equivalently, this can be reformulated in terms of U by

$$\lim_{t\to\infty} \frac{U(tx)}{U(t)} = x^{\gamma}. \tag{2.10}$$

A moment condition for the left tail of F is also necessary.

Assumption 2. $E(X_-^{p-1}) < \infty$.

Proposition 1. *Suppose F satisfies both Assumptions 1 and 2 with an order $p \in (1, 1+1/\gamma)$. Then, for all $q \in [1, p)$, we have that,*

$$\lim_{\varepsilon\downarrow 0} \frac{\theta_p(1-\varepsilon)}{\theta_q(1-\varepsilon)} = \left[\frac{B(p, \gamma^{-1}-p+1)}{B(q, \gamma^{-1}-q+1)}\right]^{\gamma} := \mathcal{L}(\gamma, p, q), \tag{2.11}$$

where $B(a,b) = \int_0^1 t^{a-1}(1-t)^{b-1}\,dt$ is the Beta function. Furthermore, we have the following three statements:

(a) If $p - 1/\gamma = 1 - q$, then $\mathcal{L}(\gamma, p, q) = 1$;

(b) If $1 - q < p - 1/\gamma < 1$, then $\mathcal{L}(\gamma, p, q) > 1$, there exists $\tau_0 \in [0, 1)$, such that $\theta_p(1-\varepsilon) > \theta_q(1-\varepsilon)$ for $\varepsilon \in (0, 1-\tau_0)$;

(c) *If* $p - 1/\gamma < 1 - q$*, then* $\mathcal{L}(\gamma, p, q) < 1$*, there exists* $\tau_0' \in [0, 1)$*, such that* $\theta_p(1-\varepsilon) < \theta_q(1-\varepsilon)$ *for* $\varepsilon \in (0, 1 - \tau_0')$*.*

Note that (2.11) is a straightforward consequence of Corollary 1 in Daouia, Girard and Stupfler (2019). Note also that the condition $p - 1 < 1/\gamma$ is not only indispensable to make $B(p, \gamma^{-1} - p + 1)$ work, but also implies $E(X_+^{p-1}) < \infty$ (see Lemma S1). When combined the condition $E(X_-^{p-1}) < \infty$, it implies that $E(|X|^{p-1}) < \infty$, and the L_p-quantile is indeed well-defined. It is also worth noting that we only need the existence of threshold τ_0. Taking the second case (b) as an example, it is readily seen that τ_0 is essentially a threshold such that $\theta_p(\tau) > \theta_q(\tau)$ for all $\tau > \tau_0$. From monotonicity and continuity of $\theta_p(\tau)$ (see Proposition S3), one alternative definition can be given by

$$\tau_0 = \begin{cases} \sup_{\tau \in (0,1)} \left\{ \tau \mid \theta_p(\tau) \le \theta_q(\tau) \right\}, & \text{if } \left\{ \tau \mid \theta_p(\tau) \le \theta_q(\tau) \right\} \neq \emptyset, \\ 0, & \text{if } \left\{ \tau \mid \theta_p(\tau) \le \theta_q(\tau) \right\} = \emptyset. \end{cases} \tag{2.12}$$

Obviously, the value of τ_0 depends on p, q and F. The value of τ_0' can be defined similarly.

Below, we focus on the case of $1 - q < p - 1/\gamma < 1$ first and redefine the positive constant $c := c(\varepsilon)$ the coefficient of TRELT (CTRELT) between $\theta_p(\tau)$ and $\theta_q(\tau)$ with the tail probability $\varepsilon := 1 - \tau$ and reformulate (2.8) as follows,

$$\theta_p(1-c\varepsilon) = \theta_q(1-\varepsilon), \quad \varepsilon \in (0, 1-\tau_0), \tag{2.13}$$

where τ_0 is given in Proposition 1. Given the orders p, q and threshold τ_0, the coefficient c varies with the value of ε and its range can be determined as $\left[1, \frac{1-\tau_0}{\varepsilon}\right]$ by $\tau_0 \leq 1 - c\varepsilon < 1$ and $1 - c\varepsilon \leq 1 - \varepsilon$. A formal definition for CTRELT can be given by,

$$\Pi_{p,q,\tau_0}(\varepsilon; X) = \inf_{c \in \left[1, \frac{1-\tau_0}{\varepsilon}\right]} \left\{ c \;\middle|\; \theta_p(1-c\varepsilon) \leq \theta_q(1-\varepsilon) \right\}. \tag{2.14}$$

We define the CTRELT $\Pi_{p,q,\tau_0}(\varepsilon; X)$ in (2.14) by an infimum rather than a definite point to prevent some infrequent situations from occurring, such as several values or even no value of $c \in \left[1, \frac{1-\tau_0}{\varepsilon}\right]$ satisfying (2.13). Since the threshold τ_0 can be determined as soon as p, q and F are known, consequently, the relevance of CTRELT to τ_0 is essentially due to p, q and F; moreover, the focus of this article is not to investigate the relationship between $\Pi_{p,q,\tau_0}(\varepsilon; X)$ and τ_0 at all. Therefore, we omit the dependence of $\Pi_{p,q,\tau_0}(\varepsilon; X)$ on τ_0 and X, and write $\Pi_{p,q,\tau_0}(\varepsilon; X)$ as $\Pi_{p,q}(\varepsilon)$ in the rest.

Next, we present an analysis of the existence and uniqueness for CTRELT.

Assumption 3. *For all $p > q \geq 1$, there exists a threshold τ_0 such that*

(a) $\theta_p(\tau_0) \leq \theta_q(1-\varepsilon)$ for all $\varepsilon \in (0, 1-\tau_0)$;

(b) both $\theta_p(\tau)$ and $\theta_q(\tau)$ are not constants on $[\tau_0, 1]$.

Proposition 2 (Existence and Uniqueness of CTRELT)**.** *Suppose F satisfies both Assumptions 1 and 2 with p, q satisfying $1 \leq q < p$, $1-q < p-\frac{1}{\gamma} < 1$. Then, for all $\varepsilon \in (0, 1-\tau_0)$, there exists $c \in \left[1, \frac{1-\tau_0}{\varepsilon}\right]$ such that* (2.13) *holds for $c = \Pi_{p,q}(\varepsilon)$ if and only if Assumption 3 (a) holds. Moreover, if Assumption 3 (b) also holds, then the $c \in \left[1, \frac{1-\tau_0}{\varepsilon}\right]$ in* (2.13) *is unique.*

Assumption 3 is a mild condition that many common distributions meet, including continuous heavy-tailed distributions with non-constant L_p-quantile. Proposition 2 indicates that there always exists a finite solution $\Pi_{p,q}(\varepsilon)$ to (2.13) such that $\Pi_{p,q}(\varepsilon) = c \in \left[1, \frac{1-\tau_0}{\varepsilon}\right]$ as long as $\theta_p(\tau_0) \leq \theta_q(1-\varepsilon)$. Assumption 3 (b) accounts for the strict monotonicity, which is essential for uniqueness. This condition can also be rephrased as the quantile function is not constant on $[\tau_0, 1]$, since a not-constant quantile implies a non-constant L_p-quantile as well. The existence and uniqueness provide the foundation for the statistical methodologies in Sections 3 - 4.

One critical feature of CTRELT is *location-scale invariance*, that is, $\Pi_{p,q}(\varepsilon; \lambda X + \mu) = \Pi_{p,q}(\varepsilon; X)$ for all $\lambda > 0$ and $\mu \in \mathbb{R}$. This property can be easily verified by *location-scale invariance* of L_p-quantile. A reasonable interpretation for this is that the value of CTRELT remains unchanged when a portfolio is scaled by a constant, or shifted by a constant loss or gain. CTRELT characterizes the shape of the distribution under risk transition without considering its location and scale. Hence, CTRELT may show some merits in measuring the variability of risk for an asset assessment, especially compared with some non-scale-free measures, such as variance.

Alternatively, the dual CTRELT is going to be defined when we move the multiplier in CTRELT from the θ_p side to the θ_q side, more precisely,

$$\theta_p(1-\varepsilon) = \theta_q\left(1-\frac{\varepsilon}{d}\right), \quad \varepsilon \in (0, 1-\tau_0). \tag{2.15}$$

Its formal definition can also be formulated similarly,

$$\pi_{p,q}(\varepsilon) = \inf_{d\in[1,\infty)} \left\{ d \,\middle|\, \theta_p(1-\varepsilon) \le \theta_q\left(1-\frac{\varepsilon}{d}\right) \right\}. \tag{2.16}$$

Compared to CTRELT, an advantage of using dual CTRELT is that we do not require the condition $\theta_p(\tau_0) \le \theta_q(1-\varepsilon)$ anymore and $\pi_{p,q}(\varepsilon)$ is always finite for all $\varepsilon \in (0, 1-\tau_0)$ assuredly.

Proposition 3 (Existence and Uniqueness of dual CTRELT)**.** *Suppose F satisfies both Assumptions 1 and 2 with p, q satisfying $1 \le q < p$, $1-q < p - \frac{1}{\gamma} < 1$. Then, for all $\varepsilon \in (0, 1-\tau_0)$, there exists $d \in [1,\infty)$ such that* (2.15) *holds for $d = \pi_{p,q}(\varepsilon)$. If Assumption 3 (b) holds, then the $d \in [1,\infty)$ in* (2.15) *is unique.*

The limit relationship for levels between expectile and quantile has been discussed in Proposition 1 of Xu et al. (2022), and a similar argument for

PELVE is also presented in Theorem 3 of Li and Wang (2023). The limit behavior for both $\Pi_{p,q}(\varepsilon)$ and $\pi_{p,q}(\varepsilon)$ can be described definitely as $\varepsilon \to 0$.

Proposition 4. *Suppose the conditions of Proposition 2 and Assumption 3 hold. Then, we have that*

$$\lim_{\varepsilon \downarrow 0} \Pi_{p,q}(\varepsilon) = \frac{B(p, \gamma^{-1} - p + 1)}{B(q, \gamma^{-1} - q + 1)} := \ell(\gamma, p, q), \tag{2.17}$$

$$\lim_{\varepsilon \downarrow 0} \pi_{p,q}(\varepsilon) = \frac{B(p, \gamma^{-1} - p + 1)}{B(q, \gamma^{-1} - q + 1)} := \ell(\gamma, p, q), \tag{2.18}$$

where $\ell(\gamma, p, q) = (\mathcal{L}(\gamma, p, q))^{1/\gamma}$.

Note that both $\Pi_{p,q}(\varepsilon)$ and $\pi_{p,q}(\varepsilon)$ converge to one same limit that depends on p, q and γ. This convergence suggests an approach to estimate $\Pi_{p,q}(\varepsilon)$ or $\pi_{p,q}(\varepsilon)$ by substituting an estimator of γ into this limit. Although this estimator is computationally tractable, it suffers some limitations and we discuss it in Section 3.1.

The following proposition provides a straightforward relationship between $\Pi_{p,q}(\varepsilon)$ and $\pi_{p,q}(\varepsilon)$. We remark that the condition $\theta_p(\tau_0) \leq \theta_q(1-\varepsilon)$ is sufficient to make $\pi_{p,q}(\Pi_{p,q}(\varepsilon)\varepsilon)$ sense in (2.20), while it is unnecessary for (2.19), since $\pi_{p,q}(\varepsilon)$ is always finite.

Proposition 5. *Suppose the conditions of Proposition 2 hold. For all* $\varepsilon \in (0, 1-\tau_0)$, *we have that*

$$\Pi_{p,q}(\varepsilon / \pi_{p,q}(\varepsilon)) = \pi_{p,q}(\varepsilon) \tag{2.19}$$

Moreover, if Assumption 3(a) also holds, then we have that

$$\pi_{p,q}(\Pi_{p,q}(\varepsilon)\varepsilon) = \Pi_{p,q}(\varepsilon). \tag{2.20}$$

The arguments of CTRELT/dual CTRELT for another case $p-\frac{1}{\gamma} < 1-q$ are included in Supplement S1 without proof.

3. Estimation of the CTRELT

Another contribution of this paper is to develop inference methods for extreme L_p-quantiles under the tail risk equivalent level transition. Suppose $X_1, ..., X_n$ are independent and identically distributed from a distribution F. Let $1-\varepsilon_n' \uparrow 1$ be extreme such that $n\varepsilon_n' \to a \in [0,\infty)$ and $1-\varepsilon_n \uparrow 1$ be intermediate such that $n\varepsilon_n \to \infty$. We aim to estimate the extreme L_p-quantile $\theta_p(1-\varepsilon_n')$ through an estimator of intermediate L_q-quantile $\theta_q(1-\varepsilon_n)$, which can be achieved by a novel TRELT-based extrapolation. Different from the classical extrapolations, for example, the one in (1.6), the proposed TRELT-based extrapolation methods have some additional elements to estimate due to the transition between the two risk measures. The key issue is to combine the extrapolative technique with (2.13) and then to plug in good estimators of CTRELT. We first sketch two approaches to extrapolate intermediate L_q-quantile to extreme L_p-quantile.

Firstly, by Proposition 1, the asymptotic relationship between $\theta_p(1-\varepsilon_n')$ and $\theta_p(1-c(\varepsilon_n)\varepsilon_n)$ follows that, as $\varepsilon_n', \varepsilon_n \downarrow 0$,

$$\frac{\theta_p(1-\varepsilon_n')}{\theta_p(1-c(\varepsilon_n)\varepsilon_n)} = \frac{\frac{\theta_p(1-\varepsilon_n')}{\theta_1(1-\varepsilon_n')}}{\frac{\theta_p(1-c(\varepsilon_n)\varepsilon_n)}{\theta_1(1-c(\varepsilon_n)\varepsilon_n)}} \times \frac{\theta_1(1-\varepsilon_n')}{\theta_1(1-c(\varepsilon_n)\varepsilon_n)} \sim \left(\frac{c(\varepsilon_n)\varepsilon_n}{\varepsilon_n'}\right)^\gamma. \quad (3.21)$$

Then, (2.13) motivates us to reformulate (3.21) as,

$$\theta_p(1-\varepsilon_n') \sim \left(\frac{c(\varepsilon_n)\varepsilon_n}{\varepsilon_n'}\right)^\gamma \theta_q(1-\varepsilon_n). \quad (3.22)$$

Secondly, the other path can be given by considering $\theta_p(1-\varepsilon_n')$ and $\theta_p(1-c(\varepsilon_n')\varepsilon_n')$, as $\varepsilon_n' \downarrow 0$,

$$\frac{\theta_p(1-\varepsilon_n')}{\theta_p(1-c(\varepsilon_n')\varepsilon_n')} = \frac{\frac{\theta_p(1-\varepsilon_n')}{\theta_1(1-\varepsilon_n')}}{\frac{\theta_p(1-c(\varepsilon_n')\varepsilon_n')}{\theta_1(1-c(\varepsilon_n')\varepsilon_n')}} \times \frac{\theta_1(1-\varepsilon_n')}{\theta_1(1-c(\varepsilon_n')\varepsilon_n')} \sim [c(\varepsilon_n')]^\gamma. \quad (3.23)$$

Then, by (2.13) and the extrapolation (1.6), (3.23) can be rewritten as,

$$\theta_p(1-\varepsilon_n') \sim [c(\varepsilon_n')]^\gamma \theta_q(1-\varepsilon_n') = \left(\frac{c(\varepsilon_n')\varepsilon_n}{\varepsilon_n'}\right)^\gamma \theta_q(1-\varepsilon_n). \quad (3.24)$$

From (3.22) and (3.24), it is readily seen that the estimation of $\theta_p(1-\varepsilon_n')$ is intricately linked to that of γ, $c(\varepsilon_n)$ or $c(\varepsilon_n')$, and $\theta_q(1-\varepsilon_n)$. We always estimate $\theta_q(1-\varepsilon_n)$ applying (1.5), whose asymptotic property has been studied well. For $\gamma > 0$, we take Hill estimator (Hill (1975)),

$$\hat{\gamma}_H = \frac{1}{k}\sum_{i=0}^{k-1} \log X_{n-i,n} - \log X_{n-k,n}, \quad (3.25)$$

as the estimator of γ throughout this paper. Here, $X_{1,n} \leq \cdots \leq X_{n,n}$ are the order statistics and $k := k(n)$ is an intermediate sequence satisfying

$k := k(n) \to \infty$ and $k/n \to 0$ as $n \to \infty$. We propose three estimators (3.28), (3.32) and (3.33) for CTRELT, which are motivated by the limit (2.17) and (1.4) respectively. Before that, it's necessary to put the second-order regular variation condition here.

Assumption 4 (Second-order regular variation)**.** *The function $\overline{F}$ satisfies a second-order regular variation condition with $\gamma > 0$, i.e., for all $x > 0$,*

$$\lim_{t\to\infty} \frac{1}{A\left(1/\overline{F}(t)\right)} \left[\frac{\overline{F}(tx)}{\overline{F}(t)} - x^{-1/\gamma}\right] = x^{-1/\gamma}\frac{x^{\rho/\gamma}-1}{\gamma\rho}, \tag{3.26}$$

where $\rho \leq 0$ and A is a positive or negative auxiliary function with $\lim_{t\to\infty} A(t) = 0$*. Equivalently, this can also be reformulated in terms of U,*

$$\lim_{t\to\infty} \frac{1}{A(t)} \left[\frac{U(tx)}{U(t)} - x^{\gamma}\right] = x^{\gamma}\frac{x^{\rho}-1}{\rho}. \tag{3.27}$$

Note that Assumption 4 implies Assumption 1 and further controls the rate of convergence. In addition, the function A is also regularly varying with index $\rho \leq 0$ (see Theorem 2.3.3 in De Haan and Ferreira (2006)).

3.1 Estimation of $\Pi_{p,q}(\varepsilon)$ via (2.17)

As previously discussed, the limit (2.17) provides a good approximation for $\Pi_{p,q}(\varepsilon)$ as $\varepsilon \to 0$, which inspires us to put up an estimator for $\Pi_{p,q}(\varepsilon)$ by plugging in $\hat{\gamma}_H$ directly, without considering what level ε is. It seems reasonable since ε always tends to 0 for both intermediate or extreme levels when the sample size is sufficiently large. This plug-in estimator is

$$\widehat{\Pi}_{p,q} = \ell(\hat{\gamma}_H, p, q) = \frac{B(p, \hat{\gamma}_H^{-1} - p + 1)}{B(q, \hat{\gamma}_H^{-1} - q + 1)}. \tag{3.28}$$

Indeed, this estimator is level-free and its uncertainty depends completely on $\hat{\gamma}_H$. Its asymptotic normality can be derived immediately by that of $\hat{\gamma}_H$ under second-order regular variation condition via Delta-method.

Theorem 1. *Suppose F satisfies Assumptions 2, 3 and 4 with p, q satisfying $1 \le q < p$, $1 - q < p - \frac{1}{\gamma} < 1$. Then, for all $\varepsilon_n, \varepsilon_n' \in (0, 1 - \tau_0)$, we have that as $n \to \infty$,*

$$\sqrt{k}\left(\widehat{\Pi}_{p,q} - \ell(\gamma, p, q)\right) \xrightarrow{d} \mathcal{N}\left(\frac{\partial}{\partial \gamma}\ell(\gamma, p, q)\frac{\lambda}{1-\rho}, \left(\frac{\partial}{\partial \gamma}\ell(\gamma, p, q)\right)^2 \gamma^2\right), \tag{3.29}$$

provided $\lim_{n\to\infty} \sqrt{k} A\left(\frac{n}{k}\right) = \lambda < \infty$. *The derivatives are given by*

$$\begin{aligned}
\frac{\partial}{\partial t}\ell(t, p, q) = \frac{\Gamma(p)}{\Gamma(q)} &\left[\frac{\Gamma(t^{-1} - q + 1)\left(\frac{\partial}{\partial t}\Gamma(t^{-1} - p + 1)\right)}{[\Gamma(t^{-1} - q + 1)]^2}\right. \\
&\left. - \frac{\Gamma(t^{-1} - p + 1)\left(\frac{\partial}{\partial t}\Gamma(t^{-1} - q + 1)\right)}{[\Gamma(t^{-1} - q + 1)]^2}\right], \\
\frac{\partial}{\partial t}\Gamma(t^{-1} - p + 1) = -\frac{1}{t^2}&\int_0^\infty s^{t^{-1} - p} e^{-s} \log s \, ds.
\end{aligned}$$

Although this estimator can work as a simple approximation to the CTRELT in practice, it suffers from several drawbacks under rigorous scrutiny. First, the CTRELT is a mapping from ε to $\Pi_{p,q}(\varepsilon)$ so that $\Pi_{p,q}(\varepsilon)$ may differ for different values ε. However, from the essence of this estimator, it makes no sense that $\widehat{\Pi}_{p,q}$ does not correspond to the value ε. Thus, the estimator (3.28) may only perform well when ε is sufficiently small and may show poor performance when it deviates far from 0 or the sample size is small. Second, the level $\tau = 1 - \varepsilon$ admits the transformation (1.4), but the limit in (2.17) fails to admit it. Thus, the estimator (3.28) does not truly estimate the tail risk transition given large risk levels. Third, the estimator (3.28) may be a bad statistical model when one considers the rates of convergence for an intermediate $\varepsilon_n \to 0$ and an extreme $\varepsilon'_n \to 0$ at the same time. It is because the rate of convergence of $\widehat{\Pi}_{p,q}$ is completely determined by the Hill estimator $\hat{\gamma}_H$. However, as $n \to \infty$, there is no difference between the convergence rates for $\varepsilon_n \to 0$ and $\varepsilon'_n \to 0$.

3.2 Estimation of $\Pi_{p,q}(\varepsilon_n)$ at intermediate level

We now propose other two empirical methods for estimating $\Pi_{p,q}(\varepsilon)$ via transformation (1.4). Using (1.4) and (2.13) yields that

$$1 - c\varepsilon = \frac{\mathbb{E}[(\theta_q(1-\varepsilon) - X)_+^{p-1}]}{\mathbb{E}[|X - \theta_q(1-\varepsilon)|^{p-1}]}, \text{ and } 1 - \varepsilon = \frac{\mathbb{E}[(\theta_q(1-\varepsilon) - X)_+^{q-1}]}{\mathbb{E}[|X - \theta_q(1-\varepsilon)|^{q-1}]}.$$

Then, we can derive $\Pi_{p,q}(\varepsilon)$ explicitly by

$$\Pi_{p,q}(\varepsilon) = \frac{\mathbb{E}[(X-\theta_q(1-\varepsilon))_+^{p-1}]}{\mathbb{E}[(X-\theta_q(1-\varepsilon))_+^{q-1}]} \times \frac{\mathbb{E}[|X-\theta_q(1-\varepsilon)|^{q-1}]}{\mathbb{E}[|X-\theta_q(1-\varepsilon)|^{p-1}]}, \tag{3.30}$$

and the corresponding estimator $\widetilde{\Pi}_{p,q}(\varepsilon)$ can be defined by using its empirical counterpart and plugging in $\hat{\theta}_q(1-\varepsilon)$ (see (1.5)) such that

$$\widetilde{\Pi}_{p,q}(\varepsilon) = \frac{\frac{1}{n}\sum_{i=1}^n (X_i-\hat{\theta}_q(1-\varepsilon))_+^{p-1}}{\frac{1}{n}\sum_{i=1}^n (X_i-\hat{\theta}_q(1-\varepsilon))_+^{q-1}} \times \frac{\frac{1}{n}\sum_{i=1}^n |X_i-\hat{\theta}_q(1-\varepsilon)|^{q-1}}{\frac{1}{n}\sum_{i=1}^n |X_i-\hat{\theta}_q(1-\varepsilon)|^{p-1}}. \tag{3.31}$$

The tail level ε in (3.31) can be either fixed or intermediate, which diverges to 0. When ε is intermediate, the estimator of $\widetilde{\Pi}_{p,q}(\varepsilon_n)$ gives,

$$\widetilde{\Pi}_{p,q}(\varepsilon_n) = \frac{\frac{1}{n}\sum_{i=1}^n (X_i-\hat{\theta}_q(1-\varepsilon_n))_+^{p-1}}{\frac{1}{n}\sum_{i=1}^n (X_i-\hat{\theta}_q(1-\varepsilon_n))_+^{q-1}} \times \frac{\frac{1}{n}\sum_{i=1}^n |X_i-\hat{\theta}_q(1-\varepsilon_n)|^{q-1}}{\frac{1}{n}\sum_{i=1}^n |X_i-\hat{\theta}_q(1-\varepsilon_n)|^{p-1}}. \tag{3.32}$$

Theorem 2. *Suppose F satisfies Assumptions 2, 3 and 4 with p, q satisfying $1 \le q < p$, $1-q < p - \frac{1}{2\gamma} < 1$. Then for all $\varepsilon_n \in (0, 1-\tau_0)$, we have that as $n \to \infty$,*

$$\sqrt{n\varepsilon_n}\left(\frac{\widetilde{\Pi}_{p,q}(\varepsilon_n)}{\Pi_{p,q}(\varepsilon_n)} - 1\right) \xrightarrow{d} \begin{cases} \mathcal{N}(\mathcal{E}_1(p,q,\gamma), \mathcal{V}_1(p,q,\gamma)), & \text{if } q > 1, \\ \mathcal{N}(\mathcal{E}_2(p,q,\gamma), \mathcal{V}_2(p,q,\gamma)), & \text{if } q = 1, \end{cases}$$

provided $\lim_{n\to\infty}\sqrt{n\varepsilon_n}A\left(\frac{1}{\varepsilon_n}\right) = \lambda < \infty$. Here, the asymptotic means $\mathcal{E}_1(p,q,\gamma)$, $\mathcal{E}_2(p,q,\gamma)$, and variances $\mathcal{V}_1(p,q,\gamma)$, $\mathcal{V}_2(p,q,\gamma)$ are provided in (S3.35) of Supplement S3.2.

The forms of these asymptotic means and variances can be simplified as (S3.36) when $2 \le q < p$. It should be noted that $\widetilde{\Pi}_{p,q}(\varepsilon_n)$ is asymptotic unbiased when $\lambda = 0$ and (3.32) incorporates $\hat{\theta}_q(1-\varepsilon_n)$, making it impossible to study asymptotic normality without investigating $\hat{\theta}_q(1-\varepsilon_n)$. However, the existing asymptotic results of $\hat{\theta}_q(1-\varepsilon_n)$ (Theorem 1 in Daouia, Girard and Stupfler (2019)), which were established by the Lindeberg-type central limit theorem, may not be sufficient to support our findings here. We hence provide a more reinforced version of the asymptotic normality of $\hat{\theta}_q(1-\varepsilon_n)$ by employing the tail empirical process (see Proposition S4). An improvement lies in the relaxation of moment condition $\mathbb{E}\left[X_-^{(2+\delta)(q-1)}\right] < \infty$ with a $\delta > 0$, which was applied for Lyapunov condition.

3.3 Estimation of $\Pi_{p,q}(\varepsilon_n')$ at extremal level

Likewise, if we consider extreme level $1-\varepsilon_n'$, the estimator for $\Pi_{p,q}(\varepsilon_n')$ can be taken by substituting $\tilde{\theta}_q^{\text{sta}}(1-\varepsilon_n')$ into (3.31) such that

$$\widetilde{\Pi}_{p,q}(\varepsilon_n') = \frac{\frac{1}{n}\sum_{i=1}^n (X_i - \tilde{\theta}_q^{\text{sta}}(1-\varepsilon_n'))_+^{p-1}}{\frac{1}{n}\sum_{i=1}^n (X_i - \tilde{\theta}_q^{\text{sta}}(1-\varepsilon_n'))_+^{q-1}} \times \frac{\frac{1}{n}\sum_{i=1}^n |X_i - \tilde{\theta}_q^{\text{sta}}(1-\varepsilon_n')|^{q-1}}{\frac{1}{n}\sum_{i=1}^n |X_i - \tilde{\theta}_q^{\text{sta}}(1-\varepsilon_n')|^{p-1}}. \tag{3.33}$$

Recall $\tilde{\theta}_q^{\text{sta}}(1-\varepsilon_n')$ is the standard extrapolation for $\theta_q(1-\varepsilon_n')$ (see (1.6)),

$$\tilde{\theta}_q^{\text{sta}}(1-\varepsilon_n') = \left(\frac{\varepsilon_n'}{\varepsilon_n}\right)^{-\hat{\gamma}_H} \hat{\theta}_q(1-\varepsilon_n). \tag{3.34}$$

Regrettably, it fails to induce an asymptotic normality for $\widetilde{\Pi}_{p,q}(\varepsilon_n')$. It is because the convergence rate of (3.34) is much faster than $\sqrt{n\varepsilon_n'}$, which makes it impossible to find a suitable rate to ensure asymptotic normality. As shown in Daouia, Girard and Stupfler (2019), a critical technique to establish asymptotic normality for (3.34) involves reformulating the relationship between $\theta_p(\tau)$ and $\theta_1(\tau)$ through a second-order expansion, making the remainder term multiplied by the rate $\sqrt{n\varepsilon_n}/\log[\varepsilon_n/\varepsilon_n']$ be negligible. However, it needs some redundant conditions on the left tail of F. To improve this, we provide a more general version of the second-order expansion for $\theta_p(\tau)$ and $\theta_q(\tau)$ in Proposition S5, where we no longer care about whether the left tail is light or heavy, and only require Assumption 2. Our expansion (S2.8) applies to all p, q that satisfy $1 \le q < p < 1 + \frac{1}{\gamma}$. It is not difficult to verify the asymptotic normality of (3.34) still holds under (S2.8) and Assumption 5.

Assumption 5. *The $R(\cdot)$ in (S2.8) satisfies $\sqrt{n\varepsilon_n}R(p, 1, \gamma, 1-\varepsilon_n) = O(1)$.*

Table 1 provides a list of three distributions used in simulation satisfying second-order regular variation condition (Assumption 4), with corresponding values of $A(\cdot)$ and ρ. The validity of Assumption 5 is related to auxiliary functions $A(\cdot)$, discussed with details in Supplement S4.1.

Theorem 3. *Suppose F is strictly increasing and satisfies Assumptions 2, 3, 4 and 5 with p, q satisfying $1 \le q < p$, $1 - q < p - \frac{1}{2\gamma} < 1$. Then, for all*

Table 1: Summary of $A(t)$ and ρ of Pareto, Fréchet and Student-t distributions in second-order regular variation condition.

Distributions	Pareto	Fréchet	Student-t
$A(t)$	0	$\frac{\gamma}{2}t^{-1}$	$\frac{-2\gamma^2 D_\gamma}{C_\gamma^{2\gamma+1}t^{2\gamma}+D_\gamma}$
ρ	$-\infty$	-1	-2γ

Here, $C_\gamma = \frac{\Gamma\left(\frac{1/\gamma+1}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{1/\gamma}{2}\right)}\gamma^{1-\frac{1}{2\gamma}}$ and $D_\gamma = -\frac{(\gamma+1)\Gamma\left(\frac{1/\gamma+1}{2}\right)}{2\gamma(1+2\gamma)\sqrt{\pi}\Gamma\left(\frac{1/\gamma}{2}\right)}\gamma^{-\frac{1}{2\gamma}}$.

$\varepsilon_n, \varepsilon_n' \in (0, 1-\tau_0)$, *we have that as* $n \to \infty$,

$$\frac{\widetilde{\Pi}_{p,q}(\varepsilon_n')}{\Pi_{p,q}(\varepsilon_n')} \xrightarrow{\mathbb{P}} \begin{cases} \Delta, & \text{if } a > 0, \\ \frac{B(q,\gamma^{-1}-q+1)B(p,(2\gamma)^{-1}-p+1)}{B(p,\gamma^{-1}-p+1)B(q,(2\gamma)^{-1}-q+1)}, & \text{if } a = 0, \end{cases}$$

provided $n\varepsilon_n' \to a \in [0,\infty)$, $\lim_{n\to\infty} \frac{\sqrt{n\varepsilon_n}}{\log[\varepsilon_n/\varepsilon_n']} = \infty$ *and* $\lim_{n\to\infty} \sqrt{n\varepsilon_n} A\left(\frac{1}{\varepsilon_n}\right) = \lambda < \infty$. *Here,* Δ *is a random variable with density function*

$$f(y) = \frac{\sqrt{a}}{\sqrt{2\pi}} \exp\left\{-\frac{a(y-1)^2}{2(c_1-c_2y)^2}\right\} \frac{|c_1-c_2|}{(c_1-c_2y)^2}, \quad \left(y \neq \frac{c_1}{c_2}\right),$$

where $c_1 = \frac{[B(q,\gamma^{-1}-q+1)]^{1/2}B(p,(2\gamma)^{-1}-p+1)}{2\sqrt{\gamma}B(p,\gamma^{-1}-p+1)}$ *and* $c_2 = \frac{B(q,(2\gamma)^{-1}-q+1)}{2\sqrt{\gamma}[B(q,\gamma^{-1}-q+1)]^{1/2}}$.

4. Estimation of extreme L_p-quantiles via TRELT

Recall the unified form of extrapolation methods we developed,

$$\theta_p(1-\varepsilon_n') \sim \left(\frac{c\varepsilon_n}{\varepsilon_n'}\right)^\gamma \theta_q(1-\varepsilon_n), \quad \text{as } n \to \infty, \tag{4.35}$$

with $c := c(\varepsilon_n)$ or $c(\varepsilon'_n)$, where the existence and uniqueness of c was justified in Section 2. So far, all estimations have been well-established. Then, the two corresponding extrapolations for $\theta_p(1-\varepsilon'_n)$ can be given by

$$\tilde{\theta}_p^{\text{int}}(1-\varepsilon'_n) = \left(\frac{\widetilde{\Pi}_{p,q}(\varepsilon_n)\varepsilon_n}{\varepsilon'_n}\right)^{\hat{\gamma}_H} \hat{\theta}_q(1-\varepsilon_n), \tag{4.36}$$

$$\tilde{\theta}_p^{\text{ext}}(1-\varepsilon'_n) = \left(\frac{\widetilde{\Pi}_{p,q}(\varepsilon'_n)\varepsilon_n}{\varepsilon'_n}\right)^{\hat{\gamma}_H} \hat{\theta}_q(1-\varepsilon_n). \tag{4.37}$$

Given that their extrapolation formulations are the same and the only difference lies in $\widetilde{\Pi}_{p,q}(\varepsilon_n)$ or $\widetilde{\Pi}_{p,q}(\varepsilon'_n)$, the third one can thus be defined by using (3.28) such that,

$$\tilde{\theta}_p^{\text{lim}}(1-\varepsilon'_n) = \left(\frac{\widehat{\Pi}_{p,q}\varepsilon_n}{\varepsilon'_n}\right)^{\hat{\gamma}_H} \hat{\theta}_q(1-\varepsilon_n). \tag{4.38}$$

Beyond that of Daouia, Girard and Stupfler (2019), the above three estimators are all embedded in the estimators of TRELT, and their uncertainty is new and unknown. In practice, the TRELT-based extrapolation may better characterize the uncertainty of risks when risk measure varies from L_q-quantile to L_p-quantile. Hence, (4.36) - (4.38) allow the estimation of $\theta_p(1-\varepsilon'_n)$ using different $\theta_q(1-\varepsilon_n)$, rather than just p and 1. Finally, under mild conditions, the asymptotic properties of $\tilde{\theta}_p^{\text{int}}(1-\varepsilon'_n)$, $\tilde{\theta}_p^{\text{ext}}(1-\varepsilon'_n)$, and $\tilde{\theta}_p^{\text{lim}}(1-\varepsilon'_n)$ are well established in Theorems 4, 5, and 6 respectively.

In this section, we always set $k = n\varepsilon_n$ or $\varepsilon_n = k/n$ where k is defined in (3.25). On the one hand, if k and ε_n are chosen independently, we can still establish similar asymptotic properties by modifying the rate of convergence and adding some suitable assumptions. On the other hand, in simulation or practice, we always begin by selecting an appropriate k according to the performance of $\hat{\gamma}_H$, and then define the intermediate level as $\varepsilon_n = k/n$.

Theorem 4. *Suppose F is strictly increasing, the conditions of Theorem 2 and Assumption 5 hold. Then, we have that as $n \to \infty$,*

$$\frac{\sqrt{n\varepsilon_n}}{\log[\varepsilon_n/\varepsilon_n']}\left(\frac{\tilde{\theta}_p^{\text{int}}(1-\varepsilon_n')}{\theta_p(1-\varepsilon_n')} - 1\right) \xrightarrow{d} \mathcal{N}\left(\frac{\lambda}{1-\rho}, \gamma^2\right),$$

provided $\lim_{n\to\infty} \frac{\sqrt{n\varepsilon_n}}{\log[\varepsilon_n/\varepsilon_n']} = \infty$ and $\lim_{n\to\infty} \sqrt{n\varepsilon_n} A\left(\frac{1}{\varepsilon_n}\right) = \lambda < \infty$.

Theorem 5. *Suppose the conditions of Theorem 3 hold. Then, we have that as $n \to \infty$,*

$$\frac{\tilde{\theta}_p^{\text{ext}}(1-\varepsilon_n')}{\theta_p(1-\varepsilon_n')} \xrightarrow{\mathbb{P}} \begin{cases} \Delta^\gamma, & \text{if } a > 0, \\ \left[\frac{B(q,\gamma^{-1}-q+1)B(p,(2\gamma)^{-1}-p+1)}{B(p,\gamma^{-1}-p+1)B(q,(2\gamma)^{-1}-q+1)}\right]^\gamma, & \text{if } a = 0, \end{cases}$$

provided $\lim_{n\to\infty} \frac{\sqrt{n\varepsilon_n}}{\log[\varepsilon_n/\varepsilon_n']} = \infty$ and $\lim_{n\to\infty} \sqrt{n\varepsilon_n} A\left(\frac{1}{\varepsilon_n}\right) = \lambda < \infty$. Here, Δ is defined in Theorem 3.

Theorem 6. *Suppose F is strictly increasing, the conditions of Theorem 1 and Assumption 5 hold. Then, we have that as $n \to \infty$,*

$$\frac{\sqrt{n\varepsilon_n}}{\log[\varepsilon_n/\varepsilon_n']}\left(\frac{\tilde{\theta}_p^{\text{lim}}(1-\varepsilon_n')}{\theta_p(1-\varepsilon_n')} - 1\right) \xrightarrow{d} \mathcal{N}\left(\frac{\lambda}{1-\rho}, \gamma^2\right),$$

provided $\lim_{n\to\infty} \frac{\sqrt{n\varepsilon_n}}{\log[\varepsilon_n/\varepsilon_n']} = \infty$ and $\lim_{n\to\infty} \sqrt{n\varepsilon_n} A\left(\frac{1}{\varepsilon_n}\right) = \lambda < \infty$.

Note that all $\tilde{\theta}_p^{\text{int}}(1-\varepsilon_n')$, $\tilde{\theta}_p^{\text{lim}}(1-\varepsilon_n')$, $\tilde{\theta}_p^{\text{sta}}(1-\varepsilon_n')$ and $\tilde{\theta}_p^{\text{qua}}(1-\varepsilon_n')$ share the same limit distribution, which is exactly the asymptotic distribution of Hill estimator $\hat{\gamma}_H$. This also implies that these four estimators are all asymptotically unbiased when $\lambda = 0$. Yet, $\tilde{\theta}_p^{\text{ext}}(1-\varepsilon_n')$ also fails to lead to asymptotic normality, primarily due to the inability of $\widetilde{\Pi}_{p,q}(\varepsilon_n')$ to achieve asymptotic normality.

5. Simulation Study

In this section, we implement some simulations to examine the empirical performance of the proposed methods. The following three distributions, all with an extreme value index γ, will be considered in the experiments:

- Pareto distribution with CDF $F(x) = 1 - x^{-1/\gamma},\ x > 1$;
- Fréchet distribution with CDF $F(x) = \exp\{-x^{-1/\gamma}\},\ x > 0$;
- Student-t distribution with PDF $f(x) = \frac{\Gamma\left(\frac{1/\gamma+1}{2}\right)}{\sqrt{\pi/\gamma}\,\Gamma\left(\frac{1/\gamma}{2}\right)}(1+\gamma x^2)^{-\frac{1/\gamma+1}{2}}$.

It is readily to verify that all these distributions satisfy Assumptions 1 - 5. To characterize the heavy tails, we set $\gamma = 1/3$ and 0.45. We take $(p,q) = \{(2.4, 1.8), (2.4, 2.0)\}$ for $\gamma = 1/3$ and $(p,q) = \{(2.0, 1.5), (2.0, 1.8)\}$ for $\gamma = 0.45$ respectively, which meets the constraint $1 - q < p - \frac{1}{2\gamma} < 1$ exactly. To find τ_0 empirically, we plot the curves of L_p-quantiles with given

p, q, against τ varying from 0 to 1 by step size 0.001. It shows that, for Pareto and Fréchet distributions, $\theta_p(\tau) > \theta_q(\tau)$ always holds on $(0, 1)$. As for Student-t distribution, due to symmetry, $\theta_p(\tau) > \theta_q(\tau)$ holds on $[0.5, 1)$ while $\theta_p(\tau) < \theta_q(\tau)$ holds on $(0, 0.5)$. Therefore, it is sufficient to take $\tau_0 = 0.5$ for the construction of CTRELT. Moreover, a tractable approach to choosing an intermediate $k(= n\varepsilon_n)$ is plotting the Hill estimator $\hat{\gamma}_H$ against k and then choosing a suitable k according to the first stable parts.

In our study, we set the sample size $n = 2000, 5000$, extreme level $\varepsilon'_n = 0.005$, intermediate level $\varepsilon_n = k/n$ and repeat the simulation $N = 1000$ times. We will implement the following methods to compare the finite-sample performance:

- BM: the standard extrapolative method $\tilde{\theta}_p^{\text{sta}}(1-\varepsilon'_n)$ (1.6) for $\theta_p(1-\varepsilon'_n)$, served as a benchmark;
- ExtraM-I: our proposed method $\tilde{\theta}_p^{\text{int}}(1-\varepsilon'_n)$ (4.36) for $\theta_p(1-\varepsilon'_n)$;
- ExtraM-II: our proposed method $\tilde{\theta}_p^{\text{ext}}(1-\varepsilon'_n)$ (4.37) for $\theta_p(1-\varepsilon'_n)$;
- ExtraM-III: our proposed method $\tilde{\theta}_p^{\text{lim}}(1-\varepsilon'_n)$ (4.38) for $\theta_p(1-\varepsilon'_n)$.

To evaluate the performance of all estimators, we calculate the *Mean*

Squared Relative Error (MSRE) based on N replications, which is given by,

$$\text{MSRE} = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{\hat{\theta}_n^{(i)}}{\theta} - 1\right)^2,$$

where $\hat{\theta}_n^{(i)}$ is the estimator we are interested in given the simulated data of the i-th replication, and θ is the true value. We use `uniroot` function in `R` to compute $\theta_p(1-\varepsilon)$ and compute $\Pi_{p,q}(\varepsilon)$ via (2.13). We only provide experimental results for BM, ExtraM-I, ExtraM-II, and ExtraM-III herein, while the results for TRELT are included in Supplement S4.2. Tables 2 - 3 collect the values of MSREs and Figures 1 - 2 present an intuitive comparison between our methods and the standard method. It is concluded immediately that all these methods show a more concentrated trend with lower MSREs as samples increase.

We provide some analyses based on these observations. First, the empirical performance of ExtraM-I is on par with that of the standard extrapolation BM, as evidenced in Figures 1 - 2, although some biases are observed within the context of a small sample size. However, ExtraM-I does present better performance than BM in terms of MSRE. Second, the most robust method appears to be ExtraM-II, which exhibits minimal bias and relatively smaller MSREs for both large and small sample sizes. Additionally, another well-performed method is ExtraM-III, which generally presents the lowest MSREs among these methods, albeit with some biases for Student-t samples. It may be because the extreme level ε_n' is sufficiently small, making the estimator $\widehat{\Pi}_{p,q}$ more reasonable, which in turn results in smaller MSREs of ExtraM-III. To sum up, compared to the standard method BM, our proposed methods ExtraM-I, ExtraM-II, and ExtraM-III, which utilize TRELT, indeed demonstrate certain merits in quantifying extreme risks via L_p-quantiles. They not only enjoy lower MSREs but also exhibit greater robustness than BM.

6. Real Data Analysis

In this section, we conduct a real data analysis to illustrate the empirical performance of our methods. The theories for statistical methodologies are derived for independent and identically distributed samples. To reduce the potential serial dependence, we utilize the weekly historical adjusted closing prices of the S&P500 Index. The data spans from May 22nd, 1967, to November 18th, 2024, consisting of 3001 trading records, obtained from Yahoo Finance (`https://finance.yahoo.com/`). We use weekly log-loss (negative log-return) data over the observation period with a moving window of 1800 trading days for weekly estimating γ, $\Pi_{p,q}$ and $\theta_p(1-\varepsilon_n')$. For instance, an estimation of γ on November 18th, 2024, will use 1800 pieces of data from May 21st, 1990, to November 11th, 2024. This rolling method

Table 2: The MSREs of BM, ExtraM-I, ExtraM-II, and ExtraM-III for Pareto, Fréchet and Student-t distributions with $\gamma = 1/3$. The bold numbers are the smallest values in each row.

Methods			BM	ExtraM-I	ExtraM-II	ExtraM-III
	$\varepsilon_n(k)$	(p,q)		$n = 2000$		
Pareto	0.0290(58)	(2.4,1.8)	0.05596	0.06796	0.04594	**0.03977**
		(2.4,2.0)	0.05596	0.06335	**0.04295**	0.04459
Fréchet	0.0385(77)	(2.4,1.8)	0.11180	0.07991	0.05771	**0.03851**
		(2.4,2.0)	0.11180	0.08533	0.06004	**0.04956**
Student-t	0.0265(53)	(2.4,1.8)	0.06039	**0.05114**	0.05528	0.09832
		(2.4,2.0)	0.06039	0.05319	**0.05109**	0.08580
				$n = 5000$		
Pareto	0.0110(55)	(2.4,1.8)	0.02995	0.03341	0.03056	**0.02170**
		(2.4,2.0)	0.02995	0.03177	0.02869	**0.02307**
Fréchet	0.0160(80)	(2.4,1.8)	0.04911	0.03612	0.03270	**0.01974**
		(2.4,2.0)	0.04911	0.03757	0.03262	**0.02273**
Student-t	0.0120(60)	(2.4,1.8)	0.03313	**0.02638**	0.02997	0.03596
		(2.4,2.0)	0.03313	**0.02742**	0.02866	0.03301

Table 3: The MSREs of BM, ExtraM-I, ExtraM-II, and ExtraM-III for Pareto, Fréchet and Student-t distributions with $\gamma = 0.45$. The bold numbers are the smallest values in each row.

	Methods		BM	ExtraM-I	ExtraM-II	ExtraM-III
	$\varepsilon_n(k)$	(p,q)			$n = 2000$	
Pareto	0.0290(58)	(2.0,1.5)	0.07669	0.08439	0.07160	**0.06319**
		(2.0,1.8)	0.07669	0.07788	**0.06347**	0.06985
Fréchet	0.0385(77)	(2.0,1.5)	0.24294	0.10273	0.09877	**0.05934**
		(2.0,1.8)	0.24294	0.14290	0.11562	**0.09251**
Student-t	0.0285(57)	(2.0,1.5)	0.07961	**0.06417**	0.08161	0.08649
		(2.0,1.8)	0.07961	0.06928	**0.06721**	0.07611
					$n = 5000$	
Pareto	0.0108(54)	(2.0,1.5)	0.04262	0.04314	0.04350	**0.03327**
		(2.0,1.8)	0.04262	0.04148	0.03982	**0.03630**
Fréchet	0.0160(80)	(2.0,1.5)	0.08966	0.04772	0.05037	**0.03106**
		(2.0,1.8)	0.08966	0.05812	0.05373	**0.04015**
Student-t	0.0110(55)	(2.0,1.5)	0.05667	**0.03583**	0.04275	0.03921
		(2.0,1.8)	0.05667	0.04150	0.04194	**0.03789**

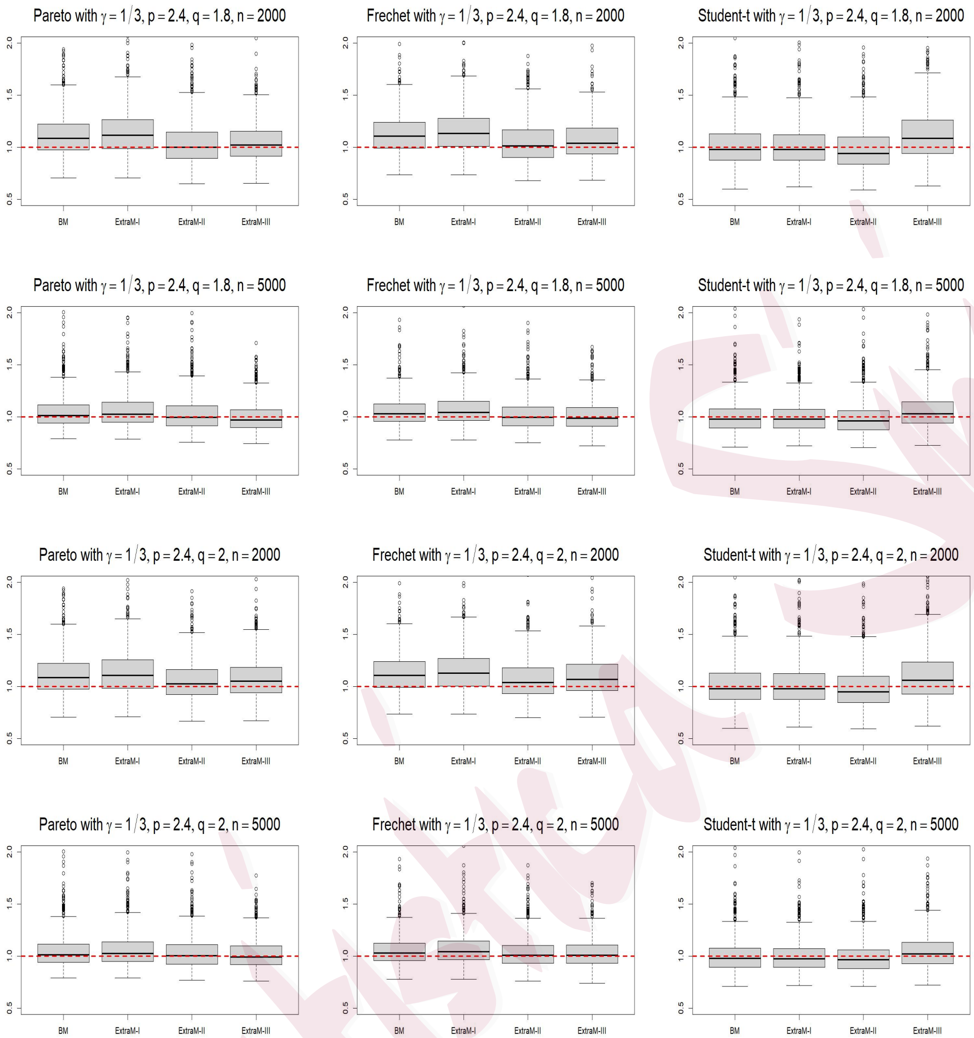

Figure 1: The boxplots of BM, ExtraM-I, ExtraM-II, and ExtraM-III for Pareto (left column), Fr$\acute{e}$chet (middle column) and Student-t (right column) distributions with $\gamma = 1/3$. The boxplots in the top two lines are drawn for $p = 2.4, q = 1.8$ with $n = 2000, 5000$ while the boxplots in the bottom two lines are drawn for $p = 2.4, q = 2$ with $n = 2000, 5000$ respectively.

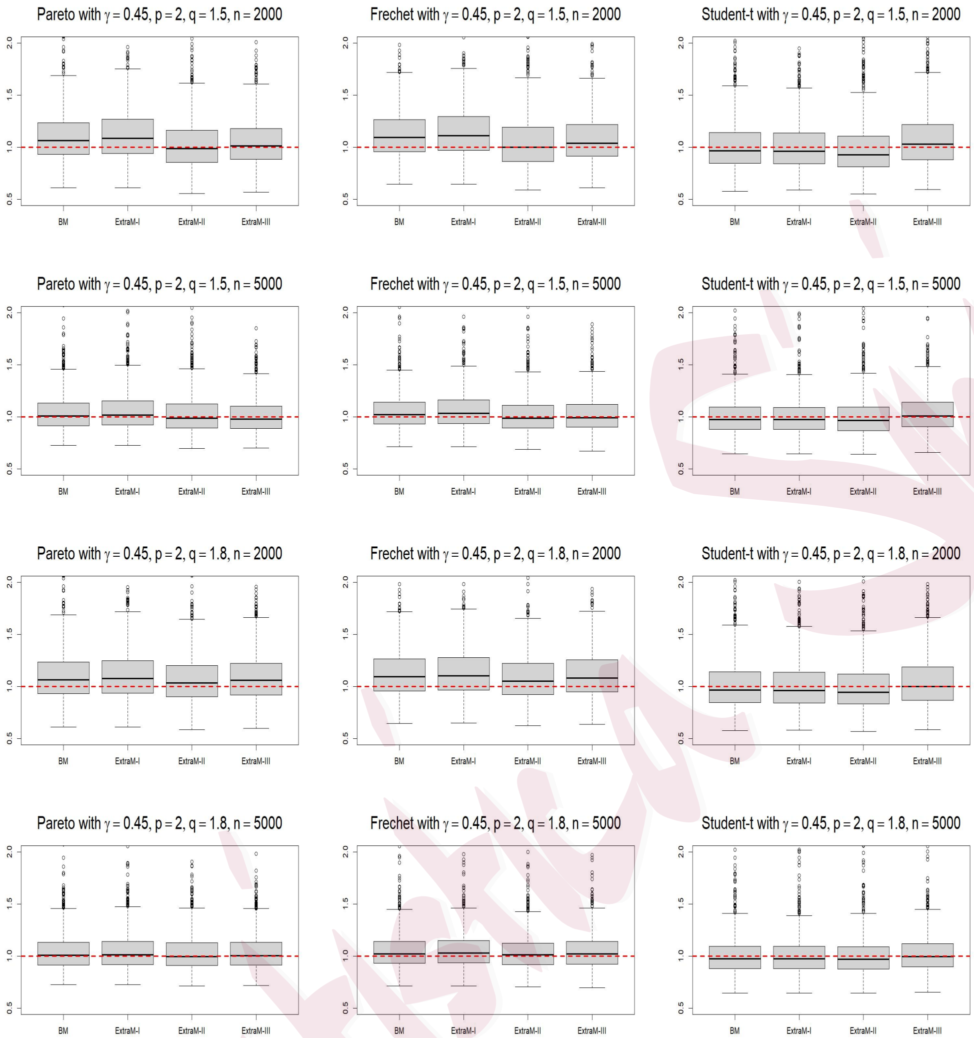

Figure 2: The boxplots of BM, ExtraM-I, ExtraM-II, and ExtraM-III for Pareto (left column), Fr$\acute{e}$chet (middle column) and Student-t (right column) distributions with $\gamma = 0.45$. The boxplots in the top two lines are drawn for $p = 2, q = 1.5$ with $n = 2000, 5000$ while the boxplots in the bottom two lines are drawn for $p = 2, q = 1.8$ with $n = 2000, 5000$ respectively.

typically covers a long period and can be used to see the dynamics of the tail heaviness γ.

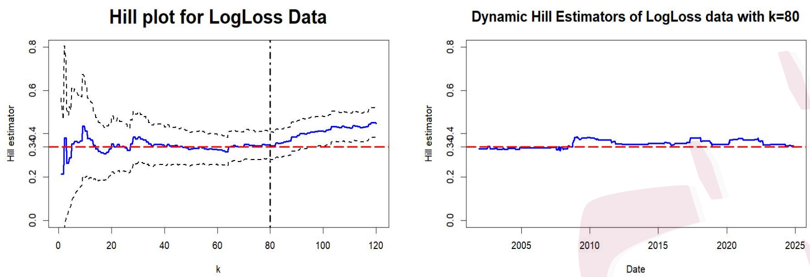

Figure 3: The left plot is drawn for choosing a suitable k, where the blue line is the Hill estimator against k, the upper and lower dashed lines are the 90% confidence bounds, and the vertical line shows the chosen $k = 80$. The right plot shows the dynamic Hill estimators with chosen k.

We first draw the Hill plots by using the recent 1800 pieces of data. As depicted in Figure 3, we choose $k = 80$, which stabilizes the Hill estimators around 0.34. We additionally draw the dynamic Hill plot with $k = 80$ using the rolling method. Obviously, the Hill estimators tend to stabilize around 0.34 again, suggesting that the choice of k is appropriate and the weekly log-loss data is empirically heavy-tailed. Based on these observations, we pick $(p, q) = \{(2, 1), (2.2, 1.5), (2.4, 2)\}$ according to $\gamma = 0.34$ and set $\varepsilon'_n = 0.005$. We present the curves of the estimations for both $\Pi_{p,q}(\varepsilon)$ and $\theta_p(1 - \varepsilon'_n)$ in Figure 4. Several observations can be made as follows:

- The estimator $\widehat{\Pi}_{p,q}$ and Hill estimator (3.25) share the same trend since the $\widehat{\Pi}_{p,q}$ is entirely dependent on the Hill estimator.

- Overall, the lines of $\widetilde{\Pi}_{p,q}(\varepsilon_n)$ and $\widetilde{\Pi}_{p,q}(\varepsilon_n')$ remain quite stable during the past three decades. Additionally, it is noteworthy that two significant changes, occurring around 2009 and 2020, are also well-reflected. This observation aligns with the widely held belief that both the financial crisis (2009) and the COVID-19 pandemic (2020) have had substantial impacts on financial markets.

- The empirical values of $\theta_p(1-\varepsilon_n')$ display the time-varying volatility observed over the past three decades. It is evident that the trends of the four methods are nearly identical. The two significant fluctuations (around 2009 and 2020) are also clearly depicted in Figure 4.

- The ranking of size for the four methods is as roughly follows: ExtraM-III $>$ ExtraM-I $\approx$ BM $>$ ExtraM-II. This outcome is highly consistent with the experimental results presented in Section 5. Due to lower risk preference, the regulators may opt for the ExtraM-II method as the most preferred choice, followed by either ExtraM-I or BM, for practical application of L_p-quantiles in quantifying extreme risks.

In summary, based on the substantial empirical studies, we can assert that TRELT serves as a tail-based measure of variability. Its values mirror the stability of the financial market and are capable of effectively identifying abnormal fluctuations. Moreover, these empirical studies also provide compelling evidence that our methods ExtraM-I, ExtraM-II, and ExtraM-III are more efficient for predicting extreme risks via L_p-quantiles.

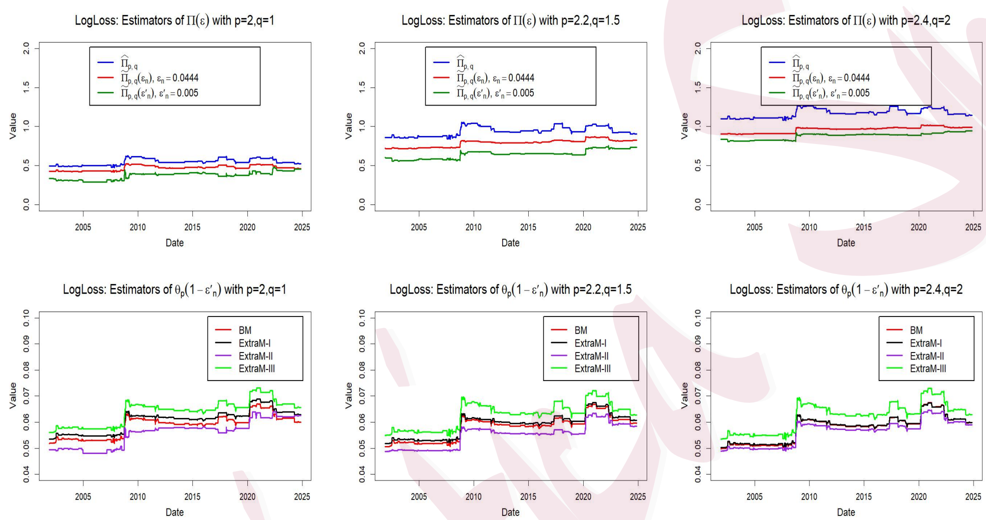

Figure 4: The dynamic estimations of $\Pi_{p,q}(\varepsilon)$ (top line) and $\theta_p(1-\varepsilon'_n)$ (bottom line) with a 1800 moving-window.

7. Conclusion

In this paper, we propose the concept of tail risk equivalent level transition (TRELT) between L_p-quantiles, which is motivated by the PELVE in Li and Wang (2023). The TRELT (and its dual) is developed under the extreme value theory to bridge different risk measures given tail equivalence of risks, which is novel in tail risk measurement. We study the theoretical properties, such as existence, uniqueness, and limiting properties, of the coefficient of TRELT, and further propose the estimation approach for it. In addition, to predict the extreme L_p-quantiles, we propose new extrapolative estimators based on the TRELT approach. Simulation studies show that our proposed estimators are effective for predicting extreme risks. As for further studies, it is of theoretical interest to study the TRELT between more general risk measures, as well as of practical interest to propose real applications of TRELT in tail risk measurement.

Supplementary Materials

The online Supplementary Material contains some theoretical statements, auxiliary results, all technical proofs and additional simulation results.

Acknowledgements

Yanxi Hou's work was supported by the National Natural Science Foundation of China Grants 72171055 and 71991471.

References

Artzner, P., Delbaen, F., Eber, J.M., and Heath, D. (1999). Coherent measures of risk. *Math. Finance* **9**, 203-228.

Asimit, V., Peng, L., Wang, R., and Yu, A. (2019). An efficient approach to quantile capital allocation and sensitivity analysis. *Math. Finance* **29**, 1131-1156.

BCBS. (2019). Minimum capital requirements for market risk. *Basel Committee on Banking Supervision Basel*.

Bellini, F., Klar, B., Müller, A. and Gianin, E.R. (2014). Generalized quantiles as risk measures. *Insur. Math. Econ.* **54**, 41–48.

Bignozzi, V., Merlo, L. and Petrella, L. (2024). Inter-order relations between equivalence for L_p-quantiles of the Student's t distribution. *Insur. Math. Econ.* **116**, 44–50.

Chen, Z. (1996). Conditional L_p-quantiles and their application to the testing of symmetry in non-parametric regression. *Stat. Probab. Lett.* **29**, 107–115.

Daouia, A., Girard, S. and Stupfler, G. (2018). Estimation of tail risk based on extreme expectiles. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **80**, 263–292.

Daouia, A., Girard, S., and Stupfler, G. (2019). Extreme M-quantiles as risk measures: From L^1 to L^p optimization. *Bernoulli* **25**, 264–309.

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

Fiori, A.M. and Rosazza Gianin, E. (2023). Generalized PELVE and applications to risk measures. *Eur. Actuar. J.* **13**, 307–339.

Girard, S., Stupfler, G. and Usseglio-Carleve, A. (2021). Extreme conditional expectile estimation in heavy-tailed heteroscedastic regression models. *Ann. Stat.* **49**, 3358–3382.

Girard, S., Stupfler, G. and Usseglio-Carleve, A. (2022). Nonparametric extreme conditional expectile estimation. *Scand. J. Stat.* **49**, 78–115.

Hill, B.M. (1975). A simple general approach to inference about the tail of a distribution. *Ann. Stat.* **3**, 1163–1174.

Jones, M.C. (1994). Expectiles and M-quantiles are quantiles. *Stat. Probab. Lett.* **20**, 149–153.

Li, H. and Wang, R. (2023). PELVE: Probability equivalent level of VaR and ES. *J. Econometrics* **234**, 353–370.

Ortega-Jiménez, P., Pellerey, F., Sordo, M.A. and Suárez-Llorens, A. (2024). Probability equivalent level for CoVaR and VaR. *Insur. Math. Econ.* **115**, 22-35.

Xu, W., Hou, Y. and Li, D. (2022). Prediction of extremal expectile based on regression models with heteroscedastic extremes. *J. Bus. Econ. Stat.* **40**, 522–536.

Ziegel, J.F. (2016). Coherence and elicitability. *Math. Finance* **26**, 901–918.

Zou, H. (2014). Generalizing Koenker's distribution. *J. Stat. Plan. Inference* **148**, 123–127.

School of Data Science, Fudan University, Shanghai, China.

E-mail: qzzhong22@m.fudan.edu.cn

School of Data Science, Fudan University, Shanghai, China.

E-mail: yxhou@fudan.edu.cn