Taylor & Francis
Taylor & Francis Group

Journal of Business & Economic Statistics

ISSN: 0735-0015 (Print) 1537-2707 (Online) Journal homepage: www.tandfonline.com/journals/ubes20

A family of coherent risk measures: an infinite weighted average of Value-at-Risk

Rong Jiang, M.C. Jones, Jiangfeng Wang & Keming Yu

To cite this article: Rong Jiang, M.C. Jones, Jiangfeng Wang & Keming Yu (10 Aug 2026): A family of coherent risk measures: an infinite weighted average of Value-at-Risk, Journal of Business & Economic Statistics, DOI: 10.1080/07350015.2026.2714065

To link to this article: https://doi.org/10.1080/07350015.2026.2714065

A family of coherent risk measures: an infinite weighted average of Value-at-Risk

Rong Jiang

Shanghai University of International Business and Economics, China

and

M.C. Jones

The Open University, UK

and

Jiangfeng Wang*

Zhejiang Gongshang University, China

and

Keming Yu*

Brunel University, UK

*: Corresponding author keming.yu@brunel.ac.uk

Abstract

Value-at-Risk (VaR) remains one of the most widely used risk measures in finance due to its simplicity, interpretability, and regulatory prominence, particularly within the Basel framework. However, VaR is not a coherent risk measure, as it generally fails to satisfy the subadditivity property. This paper demonstrates that, although finite weighted combinations of VaR-type functionals do not necessarily yield coherent risk measures, appropriately constructed infinite weighted averages of quantiles can generate coherent risk measures. Motivated by this observation, we introduce a two-parameter weight function that jointly depends on the quantile level and a position parameter, and propose a novel coherent risk measurement framework termed generalized quantile regression (GQR). In particular, we establish explicit and verifiable conditions on the weight function under which the resulting risk functional satisfies coherence, reversibility, monotonicity, and comparability across quantile levels. The proposed framework is flexible, interpretable, and unifying: it encompasses a broad class of existing coherent risk measures as special cases while also generating several new and economically meaningful risk measures. These results provide a characterization that is not explicitly available in conventional distortion or spectral risk measure frameworks. We further develop corresponding nonparametric estimators, including in multi-dimensional settings, and investigate their asymptotic properties. Empirical studies demonstrate the effectiveness of the proposed GQR framework in both risk assessment and portfolio optimization.

Keywords: Coherent risk measure; Value-at-Risk; Tail regression model; Quantile regression.

1 Introduction

Financial systems are undergoing profound structural and operational transformations, which have given rise to increasingly complex risk profiles. Against this backdrop, safeguarding financial stability has become a central concern for policymakers, regulators, and market participants. Historical financial crises show that inadequate early warning mechanisms have been a critical weakness in existing risk management frameworks. At a deeper level, this weakness reflects the limitations of conventional risk measurement tools in capturing the complexity of modern financial environments. Consequently, numerous new methodologies for measuring risk have been proposed. Nevertheless, regardless of the specific approach adopted, Value-at-Risk (VaR) continues to serve as the foundational benchmark that motivates and underpins the development of these alternative measures.

Statistically, VaR is defined based on the τ-quantile of asset returns *Y*:

$$\mathrm{VaR}_\tau(Y) = -Q_Y(\tau), \quad (1.1)$$

where $Q_Y = F_Y^{-1}$ denotes the quantile function of *Y* and F_Y is its distribution function. VaR is grounded in statistical theory and formally represents the maximum potential loss that a financial position may incur over a specified time horizon at a given confidence level $100\tau\%$, typically set at 5% or 1%. Despite its widespread adoption, VaR exhibits theoretical limitations, most notably its failure to satisfy subadditivity. Subadditivity requires that the risk of a portfolio should not exceed the sum of the risks of its individual components.

To address these shortcomings, Artzner et al. (1999) proposed the coherent risk measure framework, under which a desirable risk measure should satisfy four axiomatic properties: monotonicity, subadditivity, positive homogeneity, and translation invariance. Since VaR generally fails to satisfy subadditivity, this limitation has motivated extensive research in both academia and industry and has led to the development of alternative coherent risk measures. Expected shortfall (ES; Acerbi and Tasche (2002)) measures the average loss in the tail region beyond the VaR threshold and therefore accounts for losses that VaR ignores. Spectral Risk Measures (SRM), introduced by Acerbi (2002), summarize risk as a risk-aversion-weighted average of losses across the distribution, assigning greater weight to more severe outcomes and satisfying coherence when the weighting scheme is nonnegative, normalized, and increasingly conservative. The Exponential Spectral Risk Measure (ESRM) (Dowd and Cotter, 2007) is an important special case that uses exponentially increasing tail weights to represent stronger aversion to extreme losses; for example, Wang et al. (2025) develop an ESRM-based risk-averse scheduling model for the optimal dispatch of active distribution networks under load and distributed photovoltaic uncertainty. Extremiles (Daouia et al., 2019) capture extreme risk by replacing a single tail threshold with a tail-weighted average of the entire distribution, making them sensitive to both the probability and magnitude of extreme outcomes while preserving coherent, spectral, law-invariant, and comonotonically additive risk-measure properties. More recently, Debrauwer et al. (2025) proposed Meanimiles, a unified optimization-based class of risk functionals that combines loss functions with distributional weighting and encompasses quantiles, expectiles, extremiles, and certain distortion risk measures.

While generalized risk measures such as SRM and Meanimiles possess broad theoretical generality, quantile-based measures parameterized by τ including VaR, ES, and Extremiles are more commonly adopted in practice. This preference stems from the difficulty of selecting weight functions for SRM and Meanimiles, whereas τ-parameterized measures offer superior interpretability, with τ directly indexing a specific quantile level. To accommodate this, we incorporate the quantile parameter into the risk measurement framework and propose a novel class of generalized coherent risk measures, significantly enhancing both interpretability and practical utility. Moreover, we establish explicit conditions on the weight function that ensure coherence and monotonicity with respect to τ, thereby guaranteeing that the proposed risk measure simultaneously functions as a nonparametric regression model akin to quantile regression and characterizes data features across different quantile positions. A further contribution is the development of estimation procedures for the proposed conditional risk functionals. For univariate covariates, we use kernel conditional distribution estimation; for multivariate covariates, we adopt a single-index conditional distribution model as a dimension-reduction device.

The remainder of this paper is organized as follows. Section 2 introduces the proposed generalized quantile regression (GQR) framework through a weight function $J_\tau(s)$ and demonstrates that it encompasses a broad class of existing risk measures and non-mean regression models as special cases. Basic properties of GQR are also obtained therein. Section 3 presents examples of new risk measures that can be constructed under this framework. Section 4 develops estimation methodology. Section 5 provides simulation studies and real-data applications to assess the performance and practical relevance of the proposed methods. Section 6 concludes the paper with a brief discussion. Technical proofs are relegated to the Supplementary Material.

2 General quantile regression

2.1 Definition of GQR and comparison with classical risk measures

By analyzing risk measures such as ES, Extremile, SRM and ESRM, it becomes evident that integral-form weighted averages of quantiles can exhibit subadditivity, a property that finite weighted sums do not necessarily possess. Consequently, this study adopts the infinite quantile weighted sum in integral form to develop a novel class of coherent risk measurement tools.

We first define the proposed risk functional without covariates. For a univariate random variable Y with quantile function $Q_Y(s)$ and $0<\tau<1$, let

$$\xi_\tau(Y) = \int_0^1 Q_Y(s) J_\tau(s) ds,$$

where $J_\tau(s)$ is a density function on [0, 1]. Let $\omega_\tau = \mathrm{I}(1/2 < \tau < 1) - \mathrm{I}(0 < \tau \le 1/2)$, where $\mathrm{I}(\cdot)$ denotes the indicator function. The functional $\omega_\tau \xi_\tau(Y)$ is interpreted as a risk measure when $J_\tau(s)$ satisfies condition **C1** in Section 2.2.

We then consider its conditional version. Let $\boldsymbol{X} \in \mathbb{R}^p$ and let $Q_{Y|\boldsymbol{x}}(s)$ denote the conditional quantile function of Y given $\boldsymbol{X} = \boldsymbol{x}$. For $0 < \tau < 1$, define

$$\xi_\tau(Y \mid \boldsymbol{x}) = \int_0^1 Q_{Y|\boldsymbol{x}}(s) J_\tau(s) ds. \quad (2.1)$$

The parameter τ is aligned with the quantile level in VaR, making $\omega_\tau \xi_\tau(Y \mid \boldsymbol{x})$ comparable with VaR, ES, and extremiles. The regression problem studied in this paper is the estimation of the map $\boldsymbol{x} \mapsto \xi_\tau(Y \mid \boldsymbol{x})$; in this sense, generalized quantile regression (GQR) refers to the regression estimation of this conditional weighted-average quantile risk functional.

As detailed below, the formulation in (2.1) encompasses several widely adopted risk measures as special cases, thereby motivating our use of weighted averages of quantiles with interpretable τ-indexed weights.

(i) VaR in (1.1) arises by taking $\mathrm{J}_\tau(s) = \delta(s - \tau)$, where $\delta(\cdot)$ is the Dirac delta function ($\delta(u) = 0$ with $u \ne 0$ and $\int_{-\infty}^{+\infty} \delta(u) du = 1$). However, since VaR does not satisfy the properties of a coherent risk measure, taking $\mathrm{J}_\tau(s) = \delta(s - \tau)$ within the framework of VaR fails to satisfy condition **C1**.

(ii) ES_τ is equal to $\omega_\tau \xi_\tau(Y \mid \boldsymbol{x})$ with

$$\mathrm{J}_\tau(s) = \begin{cases} \mathrm{I}\,(s < \tau)/\tau, & \text{if } 0 < \tau \le 1/2, \\ \mathrm{I}(s \ge \tau)/(1-\tau), & \text{if } 1/2 < \tau < 1. \end{cases}$$

(iii) $\mathrm{Extremile}_\tau$ is equal to $\omega_\tau \xi_\tau(Y \mid \boldsymbol{x})$ with

$$\mathrm{J}_\tau(s) = \begin{cases} r_1(\tau)(1-s)^{r_1(\tau)-1}, & \text{if } 0 < \tau \le 1/2, \\ r_2(\tau) s^{r_2(\tau)-1}, & \text{if } 1/2 < \tau < 1, \end{cases}$$

with $r_1(\tau) = r_2(1-\tau) = \log(1/2)/\log(1-\tau)$.

(iv) ESRM is equal to $\omega_\tau \xi_\tau(Y \mid \boldsymbol{x})$ with

$$\mathrm{J}_\tau(s)=\begin{cases}(2\tau)^s\log(2\tau)/(2\tau-1), & \text{if } 0<\tau\le 1/2,\\ (2-2\tau)^{1-s}\log(2-2\tau)/(1-2\tau), & \text{if } 1/2<\tau<1.\end{cases}$$

For risk measurement, the proposed GQR is related to existing distortion risk measures (Wang, 2000) and spectral risk measures (Acerbi, 2002), which focus on weighting the quantile function solely using s as

$$\int_0^1 Q_Y(s)\psi(s)ds. \quad (2.2)$$

Chetverikov et al. (2025) extended this to a weighted-average quantile regression model of the form

$$\int_0^1 Q_{Y|\boldsymbol{X}}(s)\psi(s)ds=\boldsymbol{X}^\top\boldsymbol{\beta}, \quad (2.3)$$

which remains a linear regression model. Additionally, Debrauwer et al. (2025) proposed a generalized risk measure called meanimiles:

$$\text{Meanimile}_\tau=\min_{\theta\in\mathbb{R}}\mathbb{E}\left[d_\tau\{F_Y(Y)\}l_\delta(Y,\theta)\right], \quad (2.4)$$

where $d_\tau(\cdot)$ is a probability density function with support $(0,1)$ and $l_\delta(\cdot)$ is a loss function. The GQR in (2.1) can be viewed as a special case of (2.2); equivalently, it corresponds to (2.4) with a squared loss function. However, the general meanimile has practical application difficulties, as discussed in Debrauwer et al. (2025): it needs to select both the density function and the loss function simultaneously, so its operational complexity is relatively high.

The main distinction between GQR and the above risk measures lies in the explicit conditions imposed on the weight function $\mathrm{J}_\tau(s)$. These conditions ensure coherence and monotonicity of the proposed risk measure with respect to τ, and allow the resulting conditional functional to be used as a nonparametric regression model analogous to quantile regression. In this sense, (2.3) can be viewed as a linear specification of a GQR-type conditional functional, whereas GQR itself is formulated nonparametrically. Moreover, the weight function depends on two arguments, s and τ: s determines how different quantile levels are weighted, while τ indexes the position of the fitted risk functional. As a result, the weighted average of quantiles remains a quantile-like function of τ. In applications, τ in GQR is typically chosen to match the quantile level used in QR, which facilitates comparative analysis, as illustrated in the empirical studies in Sections 5.2 and 5.3.

In summary, this paper introduces Generalized Quantile Regression (GQR), a novel class of generalized coherent risk measures. Its core innovation is a well-defined weighting function $\mathrm{J}_\tau(s)$ that ensures coherence and monotonicity in τ, thereby enabling GQR to simultaneously serve as a nonparametric regression model akin to quantile regression for characterizing data features across quantile positions. Furthermore, through flexible specification of $\mathrm{J}_\tau(s)$, GQR subsumes a wide variety of classical risk measures, providing a universal framework for risk assessment.

2.2 Basic properties of GQR

To establish the basic properties of $\xi_\tau(Y \mid \boldsymbol{x})$ in (2.1), the following technical conditions for weighting function $\mathrm{J}_\tau(s)$ are needed.

C1. $\mathrm{J}_\tau(s)$ is a density function in s satisfying: (i) being bounded for all $\tau \in (0,1)$ and $s \in [0,1]$; (ii) a reverse condition ($\mathrm{J}_\tau(s) = \mathrm{J}_{1-\tau}(1-s)$) and being monotonic with respect to s (non-increasing for $0 < \tau \le 1/2$ and non-decreasing for $1/2 \le \tau < 1$); (iii) the corresponding distribution function $\mathrm{G}_\tau(u) = \int_0^u \mathrm{J}_\tau(s)ds$ being a non-increasing function of τ for all $u \in (0,1)$.

Remark 2.1. The boundedness condition in (i) guarantees the existence of $\xi_\tau(Y \mid \boldsymbol{x})$. Condition (ii) is used to guarantee the coherence property. "Reverse" in condition (ii) is commonly present in risk measurement such as VaR, ES and Extremile; it ensures that both tails are treated similarly. The monotonicity in condition (ii) is related to the subadditivity coherence axiom. Condition (iii) ensures that $\xi_\tau(Y \mid \boldsymbol{x})$ is non-decreasing with respect to τ.

Theorem 2.1. Let Y given $\boldsymbol{X} = \boldsymbol{x}$ have a finite absolute first moment and $\mathrm{J}_\tau(s)$ satisfy condition **C1**. Then, for any $\tau \in (0,1)$, we have

(i) $\xi_\tau(Y \mid \boldsymbol{x})$ exists and is a non-decreasing function with respect to τ.

(ii) If the conditional distribution of Y given $\boldsymbol{X} = \boldsymbol{x}$ is symmetric, then $\xi_\tau(Y \mid \boldsymbol{x}) - \mathrm{E}(Y \mid \boldsymbol{X} = \boldsymbol{x}) = \mathrm{E}\,(Y \mid \boldsymbol{X} = \boldsymbol{x}) - \xi_{1-\tau}(Y \mid \boldsymbol{x})$, which means that the lower ($\tau < 1/2$) and upper ($\tau > 1/2$) GQR curves are symmetric about the regression mean.

(iii) If $Y = m(\boldsymbol{X}) + \sigma(\boldsymbol{X})\boldsymbol{\varepsilon}$, where $m(\cdot)$ and $\sigma(\cdot) > 0$ are unknown functions, and error random variable $\boldsymbol{\varepsilon}$ has a finite absolute first moment, then $\xi_\tau(Y \mid \boldsymbol{x}) = m(\boldsymbol{x}) + \sigma(\boldsymbol{x})\xi_\tau(\boldsymbol{\varepsilon})$, where $\xi_\tau(\boldsymbol{\varepsilon}) = \int_0^1 Q_{\varepsilon}(s)\mathrm{J}_\tau(s)ds$ and $Q_\varepsilon(s)$ is the quantile function of $\boldsymbol{\varepsilon}$. This result demonstrates parallel GQR curves under response homogeneity.

(iv) $\omega_\tau \xi_\tau(Y \mid \boldsymbol{x})$ is a comonotonically additive coherent risk measure.

Theorem 2.1 demonstrates that the proposed GQR possesses several fundamental properties commonly associated with general risk measurement frameworks (Acerbi, 2002; Daouia et al., 2019, 2022; Debrauwer et al., 2025). Moreover, the condition that the random variable Y given $\boldsymbol{X} = \boldsymbol{x}$ has a finite absolute first moment is commonly employed in risk measures, as seen in Acerbi (2002), Daouia et al. (2019) and Debrauwer et al. (2025).

3 New risk measures from GQR

In this section, we propose some new coherent risk measures based on GQR (2.1), which can also be used as regression models for fitting tail features. Because of $\mathrm{J}_\tau(s)=\mathrm{J}_{1-\tau}(1-s)$ in condition **C1(ii)**, we only consider the case of $\tau\in(0,1/2]$, and the case of $\tau\in(1/2,1)$ can be similarly derived.

3.1 Generalized ES.

Note that $\mathrm{J}_\tau(s)=\mathrm{I}(0<s<\tau)/\tau$ in ES_τ, which is the density function of Uniform $(0,\tau)$ that provides a constant weight. However, in risk management, the greater the loss, the more attention is paid to it, so it should be given more weight. Therefore, we propose a generalized ES (GES), which gives greater weight to tails away from τ (see Figure 1a) as follows:

$$\mathrm{J}_\tau(s)=(1+a)\tau^{-1-a}(\tau-s)^a\times\mathrm{I}(0<s<\tau),\quad (3.1)$$

where $a\geq 0$ is a constant. When $a=0$, we have ES_τ. Moreover, when $a=1$, $\mathrm{J}_\tau(s)=2\tau^{-2}(\tau-s)\times\mathrm{I}(0<s<\tau)$, which is a decreasing linear density on $(0,\tau)$, while $\mathrm{J}_\tau(s)$ with $a=2$, that is, $3\tau^{-3}(\tau-s)^2\times\mathrm{I}(0<s<\tau)$, is a decreasing quadratic density on $(0,\tau)$. In fact, this $\mathrm{J}_\tau(s)$ in (3.1) is the density of τ times a $\mathrm{Beta}(1,a+1)$ random variable. Its distribution function is $\mathrm{G}_\tau(u)=1-(1-u/\tau)^{a+1}\times\mathrm{I}(0<u<\tau)$, which is a decreasing function of τ. In equation (3.1), τ is analogous to the quantile level in quantile regression. Accordingly, $\xi_\tau(Y\mid\boldsymbol{x})$ is a weighted average of values under the τ conditional quantile.

As *a* increases, $\mathrm{J}_\tau(s)$ has a larger value in the tail (see the left side of Figure 1a), which also results in a larger value for $\omega_\tau\xi_\tau(Y\mid\boldsymbol{x})$ (see Theorem 3.1(i) to follow and the right side of Figure 1a), where GES1 and GES2 are GES with $a=1$ and 2. Moreover, the values of $\omega_\tau\xi_\tau(Y\mid\boldsymbol{x})$ with GES ($a=0,1,2$) are all larger than QR. Indeed, this is true of any GQR where, for $0<\tau<1/2$, $\mathrm{J}_\tau(s)$ has support $(0,\tau)$: for such $\mathrm{J}_\tau(s)$,
$\xi_\tau(Y\mid\boldsymbol{x})=\int_0^\tau Q_{Y|\boldsymbol{x}}(s)\mathrm{J}_\tau(s)ds\leq Q_{Y|\boldsymbol{x}}(\tau)\int_0^\tau\mathrm{J}_\tau(s)ds\leq Q_{Y|\boldsymbol{x}}(\tau)$ (and similarly for $\tau>1/2$). Therefore, users can choose the appropriate *a* based on their risk preferences. A higher value means a more cautious approach.

3.2 Generalized Extremile.

Note that, for $0<\tau<1/2$, $\mathrm{J}_\tau(s)=r_1(\tau)(1-s)^{r_1(\tau)-1}$ where $r_1(\tau)=\log(1/2)/\log(1-\tau)$ in Extremile, which is the density function of the Beta $(1,r_1(\tau))$ distribution. Therefore, we propose a generalized Extremile (GE) as follows:

$\mathrm{J}_\tau(s) = (1+\alpha_\tau)(1-s)^{\alpha_\tau}$. (3.2)

$\mathrm{J}_\tau(s)$ is the density function of Beta $(1, \alpha_\tau + 1)$. The corresponding distribution function is $\mathrm{G}_\tau(u) = 1-(1-u)^{\alpha_\tau+1}$, which is increasing in α_τ. Therefore, $\alpha_\tau : (0,1/2) \to (0,+\infty)$ should be a decreasing function of τ to meet condition **C1**(iii). With this choice of α_τ, when $\tau = 0.5$, $\mathrm{J}_{0.5}(s) = 1$ for all $s \in (0,1)$ so that $\xi_\tau(Y \mid \boldsymbol{x}) = \mathrm{E}(Y \mid \boldsymbol{X} = \boldsymbol{x})$. Extremile is a special case with $\alpha_\tau = -\log(2-2\tau)/\log(1-\tau)$.

When α_τ is an integer, we can obtain that $\xi_\tau(Y \mid \boldsymbol{x}) = \mathrm{E}\left\{\min(Y_{\boldsymbol{x}}^1, \ldots, Y_{\boldsymbol{x}}^{1+\alpha_\tau})\right\}$, where $Y_{\boldsymbol{x}}^i$ is the *i*th sample drawn from the conditional distribution of *Y* given $\boldsymbol{X} = \boldsymbol{x}$. The choice $\alpha_\tau = 0.5\tau^{-1} - 1$ in equation (3.2) is attractive because the role of τ in $\xi_\tau(Y \mid \boldsymbol{x})$ is then to take $0.5\tau^{-1}$ independent copies of $Y_{\boldsymbol{x}}$ (see Table 1 to follow). Figure 2 further shows the influence of different τ on $\mathrm{J}_\tau(\cdot)$.

From Figure 1b, the tails of $\mathrm{J}_\tau(s)$-based GEs (GE1 and GE2 are GE with $\alpha_\tau = 0.5\tau^{-1} - 1$ and another alternative choice $0.5\pi\cot(\pi\tau)$) are smaller than that of Extremile, and therefore $\omega_\tau\xi_\tau(Y \mid \boldsymbol{x})$ based on Extremile for the normal distribution is larger than that based on GEs. Interestingly, Extremile is greater than QR, while GEs are less than QR.

3.3 Coherent risk measure with truncated Cauchy density function.

The Cauchy distribution is a practically important probability distribution. A distinguishing characteristic is that the tail of its probability density function decays at a significantly slower rate compared to that of the normal distribution, which classifies it as a prototypical heavy tailed distribution. We construct a new coherent risk measure with truncated density of the Cauchy $(0, \alpha_\tau^{-1})$ distribution over $(0,1)$ with parameter α_τ^{-1} as follows:

$$\mathrm{J}_\tau(s) = \frac{\alpha_\tau^{-1}}{\alpha_\tau^{-2} + s^2} \times \frac{1}{\arctan(\alpha_\tau)}. \quad (3.3)$$

In this case, $\mathrm{G}_\tau(s) = f(\alpha_\tau s)/f(\alpha_\tau)$, where $f(t) = \arctan(t)$ and $t > 0$. $\mathrm{G}_\tau(s)$ is increasing with respect to α_τ, because $tf'(t)/f(t) = t/\{(1+t^2)\arctan(t)\}$ can be shown to be decreasing in *t*. When $\tau = 0.5$, the $\xi_\tau(Y \mid \boldsymbol{x})$ with $\mathrm{J}_\tau(s)$ in (3.3) is, again, the conditional expectation of *Y* given $\boldsymbol{x}$. In (3.3), α_τ^{-1} is a scaling parameter, specifically equal to half the width at half the maximum value of the density. Thus, τ is the parameter that controls the scale. For instance, $\tau = 0.5(1+1/\alpha_\tau^{-1})^{-1}$ under $\alpha_\tau = 0.5\tau^{-1} - 1$. Therefore, the value of τ can be matched to the scale parameter α_τ^{-1} (see Table 1). Furthermore, the smaller the τ, the greater the weight of the tail (see Figure 2).

The $\mathrm{J}_\tau(s)$ in (3.3) and the risk measure $\xi_\tau(Y \mid \boldsymbol{x})$ with truncated Cauchy distribution (TCRM) based on $\mathrm{J}_\tau(s)$ are shown in Figure 1c (TCRM1, TCRM2 and TCRM3 are TCRM with $\alpha_\tau = 0.5\tau^{-1} - 1$, $0.5\pi\cot(\pi\tau)$ and $-\log(2-2\tau)/\log(1-\tau)$, respectively). The results show that the value of $\omega_\tau \xi_\tau(Y \mid \boldsymbol{x})$ is smaller than that of QR.

3.4 Comparing risk measurement tools

In this subsection, we explore the relationship between GES, GE, TCRM and VaR using the Fréchet distribution. The Fréchet distribution is one of the extreme value distributions, commonly used for financial risk alongside the Weibull and Gumbel distributions. Risk measurement mainly focuses on the tail situation. In Sections 3.1-3.3, focus is on the small quantile situation. This section considers another aspect (opportunity), namely the high quantile $\tau \to 1$.

Theorem 3.1. Suppose the conditions in Theorem 2.1 hold and the conditional distribution of Y given $\boldsymbol{X} = \boldsymbol{x}$ is the Fréchet distribution with the distribution function $\exp(-y^{-1/\gamma(\boldsymbol{x})})$ on support $[0, +\infty)$ and $\gamma(\boldsymbol{x}) \in (0,1)$. Then, if also in parts (ii) and (iii), $\lim_{\tau\to 1}(1-\tau)\alpha_{1-\tau} = \theta > 0$, we have

(i) Generalized ES:

$$\lim_{\tau\to 1} \frac{\xi_\tau(Y \mid \boldsymbol{x})}{\mathrm{VaR}_\tau(Y \mid \boldsymbol{x})} = (1+a)B(1-\gamma(\boldsymbol{x}), 1+a),$$

(ii) Generalized Extremile:

$$\lim_{\tau\to 1} \frac{\xi_\tau(Y \mid \boldsymbol{x})}{\mathrm{VaR}_\tau(Y \mid \boldsymbol{x})} = \theta^{\gamma(\boldsymbol{x})}\Gamma(1-\gamma(\boldsymbol{x})),$$

(iii) TCRM:

$$\lim_{\tau\to 1} \frac{\xi_\tau(Y \mid \boldsymbol{x})}{\mathrm{VaR}_\tau(Y \mid \boldsymbol{x})} = \theta^{\gamma(\boldsymbol{x})}\sec(\gamma(\boldsymbol{x})\pi/2),$$

where $B(\cdot,\cdot)$ is the beta function and $\Gamma(\cdot)$ is the gamma function.

The following results can be found from Theorem 3.1:

(1) The larger $\gamma(\cdot)$ is, the thicker the tail will be for the Fréchet distribution. From Theorem 3.1(i), ES ($a = 0$) is larger than VaR according to $(1+a)B(1-\gamma(\boldsymbol{x}), 1+a) = \{1-\gamma(\boldsymbol{x})\}^{-1} > 1$. Moreover, function $(1+a)B(1-\gamma(\boldsymbol{x}), 1+a)$ increases as $a > 0$ increases, so GES1 ($a = 1$) and GES2 ($a = 2$) behave more conservatively than ES.

(2) For GE and TCRM, the larger θ is, the larger GE and TCRM are relative to VaR for $\forall\gamma(\boldsymbol{x})\in(0,1)$, that is, the more conservative they are. For the same $\gamma(\boldsymbol{x})\in(0,1)$, GE>TCRM according to $\Gamma(1-\gamma(\boldsymbol{x}))>\sec(\gamma(\boldsymbol{x})\pi/2)$. GE and TCRM are larger than VaR for $\theta\geq1$.

(3) Extremile $\theta=\ln2$, and $(\ln2)^{\gamma(x)}\Gamma(1-\gamma(\boldsymbol{x}))>1$ for $\forall\gamma(\boldsymbol{x})\in(0,1)$. Therefore, under the Fréchet distribution, Extremile is more conservative than VaR. For GE with $\alpha_\tau=0.5(1-\tau)^{-1}-1$, then $\theta=0.5<\ln2(\approx0.69)$, therefore, Extremile is also more conservative than this GE. When $\gamma(\boldsymbol{x})<0.13168$, GE<VaR; otherwise GE>VaR.

(4) When $\theta<1$, there exists a unique $\gamma_0(\boldsymbol{x})$ such that when $\gamma(\boldsymbol{x})<\gamma_0(\boldsymbol{x})$, TCRM<VaR; when $\gamma(\boldsymbol{x})>\gamma_0(\boldsymbol{x})$, TCRM>VaR. Specifically, for TCRM with $\alpha_\tau=0.5(1-\tau)^{-1}-1$ and $\theta=0.5$, when $\gamma(\boldsymbol{x})<\gamma_0(\boldsymbol{x})=0.5$, TCRM<VaR; otherwise, TCRM>VaR.

Next, we take other specific distributions as examples to compare the commonly used and newly proposed risk measurement tools. We assume that the conditional distribution of Y given $\boldsymbol{x}$ follows the following six commonly used distributions, which are the t distribution (t(3) and t(1.2), which are in the domain of attraction of the Fréchet distribution with 1/3 and 5/6, respectively); standard Normal distribution and exponential distribution (Normal(0,1) and Exp(1), which are in the domain of attraction of the Gumbel distribution); Uniform distribution and Beta distribution (U(0,1) and Beta(2,3), which are in the domain of attraction of the Weibull distribution). The high quantiles τ from 0.90 to 0.98 are considered. We take $a=1$ for GES, $\alpha_\tau=0.5\tau^{-1}-1$ for GE and TCRM. Figure 3 shows that:

(1) For all six distributions, the order of their values is GES>ES>Extremile>GE>TCRM, which is consistent with the size of the $\mathrm{J}_\tau(s)$ tail weight (Figure 4). Based on this, the appropriate $\mathrm{J}_\tau(s)$ can be selected according to risk preference, that is, the larger the $\mathrm{J}_\tau(s)$ value at the tail, the greater the risk value (more conservative).

(2) For the distributions in the domain of attraction of the Fréchet distribution (t(3) and t(1.2)), the results conform to the conclusion in Theorem 3.1, that is, there is the relationship GES>ES>Extremile>GE>TCRM, and Extremile is greater than VaR, while TCRM is smaller than VaR under t(3) with $\gamma(\mathrm{x})=1/c<0.5$ and greater than VaR under t(1.2) with $\gamma(\mathrm{x})=5/6>0.5$.

(3) For distributions in the domain of attraction of the Gumbel distribution (Normal(0,1) and Exp(1)), the order of their values is GES>ES>Extremile>VaR>GE>TCRM, while GES>ES>VaR>Extremile>GE>TCRM under distributions in the domain of attraction of the Weibull distribution (U(0,1) and Beta(2,3)). The numerical relationship between Extremile and VaR is different here.

Beyond comparing magnitudes, selecting among GQR specifications should be guided by three criteria: the required level of conservatism (e.g., GES for strict regulation, TCRM for capital efficiency), the empirical tail behavior of the loss distribution (Theorem 3.1), and interpretability.

4 Estimation of GQR

Note that $F(Q_{Y|\boldsymbol{x}}(s)\mid \boldsymbol{x})=s$, where $F(\cdot\mid \boldsymbol{x})$ is the conditional distribution of Y given $\boldsymbol{X}=\boldsymbol{x}$. Set $y=Q_{Y|\boldsymbol{x}}(s)$, then $s=F(y\mid \boldsymbol{x})$. As others have noted in related work, we can rewrite model (2.1) as:

$$\begin{aligned}\xi_\tau(Y\mid \boldsymbol{x}) &= \int_{-\infty}^{+\infty} y f(y\mid \boldsymbol{x})\mathrm{J}_\tau\{F(y\mid \boldsymbol{x})\}dy = \mathrm{E}[Y\mathrm{J}_\tau\{F(Y\mid \boldsymbol{x})\}\mid \boldsymbol{x}]\\ &= \int_0^{+\infty}\left[1-\mathrm{G}_\tau\{F(y\mid \boldsymbol{x})\}\right]dy - \int_{-\infty}^{0}\mathrm{G}_\tau\{F(y\mid \boldsymbol{x})\}dy.\end{aligned} \quad (4.1)$$

We first present a method for estimating $\xi_\tau(Y\mid \boldsymbol{x})$ with univariate $\boldsymbol{X}$ to introduce the main idea clearly. Let $\{Y_i,\boldsymbol{X}_i\}_{i=1}^n$ be independent and identically distributed samples from $(Y,\boldsymbol{X})$ in model (2.1). From the last line of equation (4.1), we estimate $\xi_\tau(Y\mid x)$ with univariate $\boldsymbol{X}$ as:

$$\hat{\xi}_\tau(Y\mid x)=\int_0^{+\infty}\left[1-\mathrm{G}_\tau\{\hat{F}(y\mid x)\}\right]dy-\int_{-\infty}^{0}\mathrm{G}_\tau\{\hat{F}(y\mid x)\}dy. \quad (4.2)$$

In this paper, we use kernel conditional distribution estimation to estimate $F(y\mid x)$ as:

$$\hat{F}(y\mid x)=\sum_{i=1}^{n}\mathrm{I}(Y_i\le y)\mathrm{K}_h\left(X_i-x\right)/\sum_{i=1}^{n}\mathrm{K}_h\left(X_i-x\right), \quad (4.3)$$

where $K_h(\cdot)=K(\cdot/h)/h$, $K(\cdot)$ is a kernel density function and $h>0$ is the bandwidth.

When the dimensionality of $\boldsymbol{X}$ is large, the estimation method (4.3) for $F(\cdot|\cdot)$ will face the "curse of dimensionality". Therefore, we assume that there is a *p*-dimensional unknown parameter vector $\boldsymbol{\beta}_0$ that makes the following formula true:

$$F(y\mid \boldsymbol{x})=F(y\mid \boldsymbol{x}^\top\boldsymbol{\beta}_0), \quad (4.4)$$

where $\boldsymbol{x}$ is a p-dimensional vector. This is a dimension-reduction assumption and should not be interpreted as allowing fully unrestricted covariate effects across all quantile levels. Its role is to make conditional distribution estimation feasible when the dimension of $\boldsymbol{X}$ is larger than one. For identification, the first component of $\boldsymbol{\beta}_0$ is positive and $\|\boldsymbol{\beta}_0\|_2=1$, where $\|\cdot\|_2$ denotes the Euclidean 2-norm. Condition (4.4) parallels the single-index regression model, wherein the linear combination $\boldsymbol{x}^\top\boldsymbol{\beta}_0$ achieves dimension reduction. Accordingly, Chiang and Huang (2012) and Henzi et al. (2023) termed model (4.4) the single-index conditional distribution model. Nevertheless, as the conditional distribution function $F(\cdot|\cdot)$ remains unknown, $F(y\mid \boldsymbol{x}^\top\boldsymbol{\beta}_0)$ retains a semiparametric structure, comprising the parametric component $\boldsymbol{\beta}_0$ and the nonparametric component $F(\cdot|\cdot)$.

From the equations (4.1) and (4.4), we can derive $\xi_\tau(Y \mid \boldsymbol{x})$ as a generalized quantile single-index regression (GQSIR) as follows:

$$\xi_\tau(Y \mid \boldsymbol{x}^\top \boldsymbol{\beta}_0) = \int_0^{+\infty} \left[1 - \mathrm{G}_\tau\{F(y \mid \boldsymbol{x}^\top \boldsymbol{\beta}_0)\}\right] dy - \int_{-\infty}^0 \mathrm{G}_\tau\{F(y \mid \boldsymbol{x}^\top \boldsymbol{\beta}_0)\} dy \equiv g_\tau(\boldsymbol{x}^\top \boldsymbol{\beta}_0), \quad (4.5)$$

where $g_\tau(\cdot)$ is an unknown function determined by the unknown conditional distribution function $F(\cdot|\cdot)$. Therefore, we can estimate $\xi_\tau(Y \mid \boldsymbol{x}^\top \boldsymbol{\beta}_0)$ as:

$$\hat{\xi}_\tau(Y \mid \boldsymbol{x}^\top \boldsymbol{\beta}) = \int_0^{+\infty} \left[1 - \mathrm{G}_\tau\{\hat{F}(y \mid \boldsymbol{x}^\top \boldsymbol{\beta})\}\right] dy - \int_{-\infty}^0 \mathrm{G}_\tau\{\hat{F}(y \mid \boldsymbol{x}^\top \boldsymbol{\beta})\} dy. \quad (4.6)$$

According to definitions (4.4) and (4.5), $\boldsymbol{\beta}_0$ is independent of τ, then $\boldsymbol{\beta}$ can be obtained by the pseudo sum of integrated squares (PSIS) inspired by Chiang and Huang (2012) and Huang and Chiang (2017) as:

$$\begin{aligned} \boldsymbol{\beta} &= \arg\min_{\boldsymbol{\beta} \in \mathbb{R}^p} \sum_{i=1}^n \int_{-\infty}^{+\infty} \left\{ \mathrm{I}(Y_i \le y) - \hat{F}(y \mid \boldsymbol{X}_i^\top \boldsymbol{\beta}) \right\}^2 d\hat{F}(y) \\ &= \arg\min_{\boldsymbol{\beta} \in \mathbb{R}^p} \sum_{i=1}^n \sum_{j=1}^n \{ \mathrm{I}(Y_i \le Y_j) - \hat{F}(Y_j \mid \boldsymbol{X}_i^\top \boldsymbol{\beta}) \}^2, \end{aligned} \quad (4.7)$$

where $\hat{F}(y) = n^{-1} \sum_{i=1}^n \mathrm{I}(Y_i \le y)$.

To establish the asymptotic normality of the proposed estimators, the following technical assumptions are introduced.

C2. The conditional distribution function $F(y \mid x)$ has continuous second-order partial derivatives with respect to x and the conditional density function $f(y \mid x)$ satisfies $0 < f(y \mid x) < \infty$ for all $y \in \mathbb{R}$ and $x \in \mathrm{I}_x$, where I_x is a bounded interval on $\mathbb{R}$. In addition, the density function $f_{\boldsymbol{X}}(\cdot)$ of $\boldsymbol{X}$ is positive and continuously differentiable on $\mathbb{R}$.

C3. The kernel function $K(\cdot)$ is even, integrable, and twice differentiable with bounded first and second derivatives such that $\int K(u) du = 1$, $\int |u^2 K(u)| du < \infty$, $\int u K(u) du = 0$ and $\int u^2 K(u) du \ne 0$.

C4. Suppose that $\inf_{\boldsymbol{x}^\top \boldsymbol{\beta}} f(\boldsymbol{x}^\top \boldsymbol{\beta}) > 0$ for all $\boldsymbol{x} \in \mathrm{I}_{\boldsymbol{x}}^p$ and $\boldsymbol{\beta} \in \mathbb{R}^p$, where $f(\boldsymbol{x}^\top \boldsymbol{\beta})$ is the density function of $\boldsymbol{X}^\top \boldsymbol{\beta}$. Moreover, the third derivative of $f(\boldsymbol{x}^\top \boldsymbol{\beta})$ and $\mathrm{E}\{F(y \mid \boldsymbol{X}^\top \boldsymbol{\beta})(\boldsymbol{x} - \boldsymbol{X})(\boldsymbol{x} - \boldsymbol{X})^\top \mid \boldsymbol{x}^\top \boldsymbol{\beta}\}$ with respect to $\boldsymbol{x}^\top \boldsymbol{\beta}$, are Lipschitz continuous in $\boldsymbol{x}^\top \boldsymbol{\beta}$ with the Lipschitz constants being independent of $(y, \boldsymbol{x}^\top \boldsymbol{\beta})$.

C5. $\boldsymbol{\Sigma}_1 = 4\mathrm{E}(\boldsymbol{A}\boldsymbol{A}^\top)$ with $\boldsymbol{A} = \int \left\{ \mathrm{I}(Y \le y) - F(y \mid \boldsymbol{X}^\top \boldsymbol{\beta}_0) \right\} \nabla F(y \mid \boldsymbol{X}^\top \boldsymbol{\beta}_0) dF(y)$ and $\boldsymbol{\Sigma}_2 = 2\mathrm{E}\int \, [\left\{ \nabla F(y \mid \boldsymbol{X}^\top \boldsymbol{\beta}_0) \right\}^2 - \{\mathrm{I}(Y \le y) - F(y \mid \boldsymbol{X}^\top \boldsymbol{\beta}_0)\} \nabla^2 F(y \mid \boldsymbol{X}^\top \boldsymbol{\beta}_0)] dF(y)$ are non-singular. Moreover, the minimum eigenvalue of $\boldsymbol{\Sigma}_2$ is positive.

Remark 4.1. Assumption **C2** imposes mild regularity conditions on the conditional distribution and density of Y given $\boldsymbol{X}$, for instance, the normal distribution satisfies this condition. Condition **C3** is a mild condition on $K(\cdot)$. For example, taking a normal density as the kernel function satisfies condition **C3**. The condition **C4** is the smoothness condition required for the uniqueness and convergence of the estimator. **C5** ensures the asymptotic normality of the estimator. Conditions **C4** and **C5** are the general conditions for establishing the consistency and asymptotic normality of the single-index conditional distribution model (4.6) (Chiang and Huang, 2012; Henzi et al., 2023).

Theorem 4.1. Assume that Y given $\boldsymbol{X} = x$ has a finite absolute first moment and that conditions **C1-C3** hold. Suppose that $h = O(n^{-c_1})$ with $c_1 \in (1/9, 1/5]$ and $n \to \infty$. Then for given $x \in \mathrm{I}_x$, we have

$$\sqrt{nh}\left\{ \hat{\xi}_\tau(Y \mid x) - \xi_\tau(Y \mid x) - \tfrac{1}{2}\nu_2^1 h^2 B_x \right\} \xrightarrow{L} \mathrm{N}\left(0, \boldsymbol{\Sigma}_x\right),$$

where $\nu_a^b = \int_{-\infty}^{+\infty} u^a K^b(u) du$, $B_x = -\int_{-\infty}^{+\infty} \mathrm{J}_\tau\{F(y \mid x)\}\{\mathrm{F}^{''}(y \mid x) + 2\mathrm{F}^{'}(y \mid x) f_{\boldsymbol{X}}^{'}(x) / f_{\boldsymbol{X}}(x)\} dy$, $F^{'}(y \mid x) = \partial F(y \mid x) / \partial x$, $F^{''}(y \mid x) = \partial^2 F(y \mid x) / \partial x^2$, $f_{\boldsymbol{X}}^{'}(x) = \partial f_{\boldsymbol{X}}(x) / \partial x$, $\xrightarrow{L}$ stands for convergence in distribution, and

$$\boldsymbol{\Sigma}_x = \nu_0^2 f_{\boldsymbol{X}}^{-1}(x) \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \mathrm{J}_\tau\{F(y_1 \mid x)\} \mathrm{J}_\tau\{F(y_2 \mid x)\} \left\{ F(y_1 \wedge y_2 \mid x) - F(y_1 \mid x) F(y_2 \mid x) \right\} dy_1 dy_2.$$

Theorem 4.2. Assume that Y given $\boldsymbol{X} = \boldsymbol{x}$ has a finite absolute first moment and that conditions **C1-C5** are satisfied. If $h = O(n^{-c_2})$ with $c_2 \in (1/8, 1/5)$. Then, under $n \to \infty$, we have

$$\sqrt{n}(\boldsymbol{\beta} - \boldsymbol{\beta}_0) \xrightarrow{L} \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_2^{-1} \boldsymbol{\Sigma}_1 \boldsymbol{\Sigma}_2^{-1}),$$

and

$$\sqrt{nh}\left\{ \hat{\xi}_\tau(Y \mid \boldsymbol{x}^\top \boldsymbol{\beta}) - \xi_\tau(Y \mid \boldsymbol{x}^\top \boldsymbol{\beta}_0) - \tfrac{1}{2}\nu_2^1 h^2 B_{\boldsymbol{x}^\top \boldsymbol{\beta}_0} \right\} \xrightarrow{L} \mathrm{N}\left(0, \boldsymbol{\Sigma}_{\boldsymbol{x}^\top \boldsymbol{\beta}_0}\right),$$

where $B_{\boldsymbol{x}^\top \boldsymbol{\beta}_0}$ and $\boldsymbol{\Sigma}_{\boldsymbol{x}^\top \boldsymbol{\beta}_0}$ are respectively B_x and $\boldsymbol{\Sigma}_x$, which use $\boldsymbol{x}^\top \boldsymbol{\beta}_0$ instead of x.

Remark 4.2. If we only consider Y without the covariate $\boldsymbol{X}$, GQR is reduced to $\xi_\tau(Y)=\int_0^1 Q_Y(s)\mathrm{J}_\tau(s)ds$. A corresponding estimate of $\xi_\tau(Y)$ is $\hat{\xi}_\tau(Y)=n^{-1}\sum_{i=1}^{n}\tilde{Y}_i\mathrm{J}_\tau\{i/(n+1)\}$, where $\tilde{Y}_1\le\cdots\le\tilde{Y}_n$ denotes the ordered sample. For any given $\tau\in(0,1)$ and $\mathrm{E}|Y|^{\varsigma}<\infty$ for some $\varsigma>2$, then by Theorem 1(ii) of Shorack and Wellner (1986), we have

$$\sqrt{n}\left\{\hat{\xi}_\tau(Y)-\xi_\tau(Y)\right\}\xrightarrow{L}\mathrm{N}\left(0,\int_0^1\int_0^1\mathrm{J}_\tau(r)\mathrm{J}_\tau(s)(r\wedge s-rs)dF^{-1}(r)dF^{-1}(s)\right).$$

Remark 4.3. If the integral (4.6) cannot produce an explicit expression, since expressions for $\mathrm{J}_\tau(s)$ and $\hat{F}(y\,|\,\boldsymbol{x}^\top\boldsymbol{\beta})$ are known, it can be accurately approximated by the rectangular method as follows:

$$\hat{\xi}_\tau(Y\,|\,\boldsymbol{x}^\top\boldsymbol{\beta})\simeq\overline{\mathrm{Y}}\mathrm{I}(\overline{\mathrm{Y}}>0)-\underline{\mathrm{Y}}\mathrm{I}(\underline{\mathrm{Y}}>0)-\frac{\overline{\mathrm{Y}}-\underline{\mathrm{Y}}}{n}\sum_{i=0}^{n}\mathrm{G}_\tau\{\hat{F}(y_i\,|\,\boldsymbol{x}^\top\boldsymbol{\beta})\}+O(n^{-1}),\quad(4.8)$$

where $\overline{\mathrm{Y}}=\max\{Y_1,\ldots,Y_n\}$, $\underline{\mathrm{Y}}=\min\{Y_1,\ldots,Y_n\}$ and $y_i=\underline{\mathrm{Y}}+i(\overline{\mathrm{Y}}-\underline{\mathrm{Y}})/n$, $i=0,1,\ldots,n$. Because $\hat{\xi}_\tau(Y\,|\,\boldsymbol{x}^\top\boldsymbol{\beta})-\xi_\tau(Y\,|\,\boldsymbol{x}^\top\boldsymbol{\beta}_0)=O_p((nh)^{-1/2})$ in Theorem 4.2 and $h\to 0$, the error term $O(n^{-1})$ in (4.8) can be asymptotically ignored.

5 Numerical studies

In this section, we first employ a Monte Carlo simulation study to evaluate the finite-sample performance of the proposed procedures. Subsequently, we illustrate the application of the proposed methods through two real-data analyses. The versions of GQR considered here are identical to those described in Section 3.4, including choices of a and α_τ. The standard normal density is utilized as the kernel function, and the bandwidth h is determined via the cross-validation method (Li et al., 2013) in this section. All programs are implemented using R code.

5.1 Simulation example

In this subsection, we study the estimation method proposed in Section 4 for risk measures (Sections 2 and 3) involved in GQR. We generate 300 data points from the following model:

$$Y=20\sin(\pi\boldsymbol{X})+\sigma(\boldsymbol{X})\varepsilon,$$

where $\boldsymbol{X}$ is drawn from a normal distribution $N(0,1)$. Three error distributions of ε are considered: Normal(0,1), t(3) and Exp(1), where Normal(0,1) is the most commonly used, t(3) is a thick-tailed distribution and Exp(1) is one-sided. Two cases of $\sigma(\boldsymbol{X})$ are considered: homogeneity $\sigma(\boldsymbol{X})=1$ and heterogeneity $\sigma(\boldsymbol{X})=|\boldsymbol{X}|$.

The relative percentage absolute deviation (RPAD) is used to assess the performance of estimates as: $\mathrm{RPAD} = |\hat{\xi}_\tau(Y|x) - \xi_\tau(Y|x)| / |\xi_\tau(Y|x)| \times 100\%$. We take $x = -0.5$ for small values $\tau = 0.05, 0.10$ and $x = 0.5$ for large values $\tau = 0.90, 0.95$, respectively. They represent loss (negative at $\tau = 0.05, 0.10$) and gain (positive at $\tau = 0.90, 0.95$). All reported simulation results are averages over 500 simulation replications.

For extreme values $\tau \in \{0.05, 0.10, 0.90, 0.95\}$ in Tables 2 and 3, since all RPAD values are less than 10% (most are less than 5%, indeed all but one are in Table 3), the proposed estimation method performs well. It is also worth noting that the RPAD values in Table 3 are uniformly smaller than those in Table 2. This indicates that, under the heteroscedastic design considered here, the proposed estimator performs better than under the homoscedastic design. A plausible explanation is that the evaluation points are $x = -0.5$ for lower-tail risks and $x = 0.5$ for upper-tail risks. Under $\sigma(X) = |X|$, the conditional scale at these points is $\sigma(X) = 0.5$, whereas under $\sigma(X) = 1$ the conditional scale is twice as large. The smaller local noise level reduces the variability of the conditional distribution estimator, leading to smaller absolute estimation errors and hence lower RPAD values. Therefore, the heteroscedastic case in Table 3 should not be interpreted as being uniformly more difficult; rather, it is easier at the specific evaluation points used in this simulation. In practice, this suggests that the finite-sample accuracy of the proposed method depends not only on whether heteroscedasticity is present, but also on the local conditional scale at the point where the risk measure is evaluated.

5.2 Real-data example 1: Investment portfolio

In this subsection, GQR is applied to investment portfolios to illustrate its practical application in finance. The 10 stocks in the portfolio, with reference to the Blackrock U.S. Flexible Equity Fund (BR), are MSFT, AMZN, META, V, NVDA, CIEN, ICE, APD, CAH, WFC. The 250 trading days in 2023 are used as a fixed fitting sample, while the 252 trading days in 2024 are used as an out-of-sample test sample. Thus, this empirical exercise is a fixed train-test comparison rather than a rolling-window portfolio backtest. We use it to illustrate how different risk measures generate different portfolio weights and out-of-sample performance. A rolling-window re-training analysis would be a useful extension for a more complete dynamic portfolio study. The data for the 10 stocks are downloaded from Yahoo Finance (https://hk.finance.yahoo.com). BR has performed well with returns of 22.46% and 15.21% in 2023 and 2024, respectively. For specific information about BR see https://www.blackrock.com/cn/products/228610/bgf-us-flexible-equity-fund-a2-usd.

The weight $\boldsymbol{\alpha}$ of the specific portfolio is chosen to minimize the GQR for a specified value of τ :

$$\min_{\boldsymbol{\alpha}} \omega_\tau \xi_\tau(\boldsymbol{\alpha}^\top \boldsymbol{Y}), \quad s.t. \boldsymbol{\alpha}^\top \mathbf{1} = 1, \boldsymbol{\alpha} \geq 0,$$

where $\xi_\tau(\boldsymbol{\alpha}^\top \boldsymbol{Y})$ is defined and estimated in Remark 4.2, $\mathbf{1}$ is a 10×1 dimensional vector with all entries equal to one, and $\boldsymbol{Y} = (Y_1, \ldots, Y_{10})^\top$ is the logarithmic return of the above 10 stocks. We set

$\tau = 0.05$, a conventional lower-tail level in financial risk management, to make the different risk measures comparable under the same tail probability. This choice is used for a controlled comparison and is not meant to imply that the same τ is necessarily optimal for all risk measures.

The optimal $\boldsymbol{\alpha}$ for the different criteria is reported in Table 4. BR is excluded from Table 4 because the fund company discloses only its top 10 holdings, so the complete portfolio weights are unavailable. Table 4 reports the portfolio allocations implied by the different risk criteria. The table shows that the choice of risk measure materially affects the selected exposures. For example, QR assigns relatively large weights to NVDA and ICE, ES places larger weights on ICE, APD and WFC, Extremile concentrates more heavily on CAH, while TCRM assigns relatively large weights to META, NVDA and WFC. These weights are useful for interpreting the performance comparison in Table 5, but they should not be viewed as performance measures themselves.

Portfolio performance is evaluated using two complementary metrics: (1) the Sharpe Ratio (SR), calculated as annualized return divided by return volatility (standard deviation), and (2) the Percentage of Days (PD) with excess returns relative to the benchmark (BR). Both indicators are larger-is-better measures. Analysis of Table 5 reveals that the TCRM strategy demonstrates superior performance across both evaluation dimensions, achieving top-ranked SR and PD values among all seven methods examined. Notably, TCRM generates a 52.27% absolute return, outperforming BR's 15.21% by a margin of 3706 basis points. Under this fixed train-test design and the two reported evaluation metrics, TCRM performs best among the methods considered. However, the results also show that the newly proposed measures do not uniformly dominate QR: in this example, GES and GE have lower SR and PD values than QR. This illustrates that the choice of $J_\tau(s)$ matters in applications and should be guided by the investor's risk preference, the tail behavior of returns, and the evaluation criterion.

5.3 Real-data example 2: Beijing multi-site air quality dataset

We apply the proposed GQSIR in Section 4 to the analysis of a Beijing multi-site air quality dataset (Chen, 2017). This dataset includes air pollutant ($PM_{2.5}$) data from 12 nationally controlled air quality monitoring sites. The air quality data, comprising 1080 entries, are sourced from the Beijing Municipal Environmental Monitoring Center. The Chinese standard requires the 24-hour average $PM_{2.5}$ concentration to be below 75 micrograms per cubic meter. A 24-hour average concentration up to 35 micrograms per cubic meter is classified as optimal, a value up to 75 micrograms per cubic meter as good, and a value above 75 micrograms per cubic meter as polluted. The dataset can be obtained from https://archive.ics.uci.edu/dataset/501/beijing+multi+site+air+quality+data.

The official air quality statistics in China are predicated on daily $PM_{2.5}$ values. Nevertheless, it is known that the observed $PM_{2.5}$ levels are affected by meteorological conditions (Zhang et al., 2017). Secondary formation of fine particulate matter can be promoted by interactions among high humidity, high temperature and calm wind. Therefore, the meteorological data for each air quality site are obtained from the nearest weather station of the China Meteorological

Administration. The meteorological variables are temperature (TEMP), pressure (PRES), dew point temperature (DEWP) and wind speed (WSPM). Zhang et al. (2017) used a nonparametric mean regression model to analyze the dataset, which is a special case of GQSIR with $\tau = 0.5$.

The histogram of $PM_{2.5}$ in Figure 5 reveals a significant right skew. Moreover, people tend to be more concerned about high $PM_{2.5}$ levels rather than the average. Therefore, it is more appropriate to analyze this dataset using a non-mean regression model such as GQSIR. This section focuses on daily data from the winter of 2016/17 (December 2016 to February 2017), as winter typically exhibits the highest average $PM_{2.5}$ levels compared to other seasons. In addition, to remove scale differences among covariates, the four meteorological variables were standardized.

Inspired by the functions (4.2) and (4.3) in Zhang et al. (2017), we calculate the $\text{Average PM}_{2.5} = 1080^{-1}\sum_{i=1}^{1080}\hat{\xi}_\tau(Y \mid \boldsymbol{X}_i^\top \boldsymbol{\beta})$, where Y is $PM_{2.5}$, $\boldsymbol{X}$ is (TEMP, PRES, DEWP, WSPM), $\hat{\xi}_\tau(Y \mid \cdot)$ is obtained by (4.6) and $\boldsymbol{\beta} = (0.370, 0.275, -0.814, 0.354)^\top$ is obtained by (4.7). Table 6 reports the Average $PM_{2.5}$ values for τ ranging from 0.01 to 0.99. The resulting patterns are consistent with the analysis in Section 3. Moreover, Table 6 shows that the median (83, QR) and mean (98, Zhang et al. (2017)) of $PM_{2.5}$ in Beijing are greater than the standard value 75.

To extract more information from the data, we use the GQSIR non-mean regression models for further analysis. Table 7 provides a calibration of the different risk measures against the official $PM_{2.5}$ thresholds. The intervals show the values of τ for which the sample-average fitted conditional risk index falls into the optimal, good, and polluted categories. The results show that different choices of $J_\tau(s)$ lead to different mappings between τ and the official air-quality categories. For GES, the interval corresponding to the good category is very narrow, indicating that this specification moves quickly from the optimal category to the polluted category as τ increases. For TCRM, the transition to the polluted category occurs at a relatively small value of τ. In contrast, Extremile and GE provide a more gradual transition across the three categories, and their results are close to each other. This suggests that, for this dataset, Extremile and GE give more interpretable calibrations of the $PM_{2.5}$ risk levels.

In a recent winter (December 2023 to February 2024), the average $PM_{2.5}$ concentration was 38, which was close to the optimal threshold of 35 and much lower than 98 (the winter of 2016/17). Moreover, the average annual concentration of $PM_{2.5}$ in Beijing's atmospheric environment in 2023 was 32 micrograms per cubic meter, and the average annual concentration of $PM_{2.5}$ in Beijing in the first three quarters of 2024 (January-September) was 29. These data are from the Beijing Municipal Ecology and Environment Bureau (https://sthjj.beijing.gov.cn/bjhrb/index/index.html).

6 Conclusion

This paper introduces a novel family of coherent risk measures based on equation (2.1), termed Generalized Quantile Regression (GQR), which can also be used as regression models for fitting tail features. Although certain conditions are imposed on $\mathrm{J}_{\tau}(s)$, the flexibility and adaptability of this function, depending on both τ and s, enable GQR to encompass many classical and recently proposed risk measures as special cases. Sections 2 and 3 introduce a series of innovative coherent risk metrics, providing readers with a foundational framework to construct customized models and risk assessment tools grounded in GQR principles. The estimation theory developed in Section 4 complements the proposed risk-measure construction by providing nonparametric estimators for univariate covariates and single-index estimators for multivariate covariates.

Risk measurement often requires real-time analysis of time-series data that arrive continuously and rapidly. Such data can be viewed as streaming data, namely a dynamically expanding dataset that grows over time. The proposed method could be extended to streaming data analysis, enabling the development of corresponding online learning algorithms (Jiang and Yu, 2024). To apply GQR in streaming settings, a local polynomial interpolation method (Chen et al., 2024) can be employed to obtain the online update estimator $\hat{F}(y \mid \boldsymbol{x})$ in equation (4.3), which in turn allows for real-time updating of the estimator $\hat{\xi}_{\tau}(Y \mid \boldsymbol{x})$ as new data become available.

Acknowledgments

This research is supported by the Humanities and Social Sciences Research Planning Fund of the Ministry of Education (Grant No. 25YJA910003), the National Social Science Fund of China (Grant No. 25BTJ041), the National Key R&D Program of China (Grant No. 2024YFA1013502), the National Natural Science Foundation of China (Grant Nos. U23A2064 and 12531013), the Natural Science Foundation of Zhejiang Province (Grant No. LY24A010004), and the Chern Institute of Mathematics Visiting Scholar Program.

Disclosure Statement

The authors do not have conflicts of interest with regard to this word.

Supplementary material

The Supplementary Material contains the proofs of the theorems and the algorithm.

References

Acerbi, C. (2002). Spectral measures of risk: a coherent representation of subjective risk aversion. *Journal of Banking & Finance 26*, 1505–1518.

Acerbi, C. and D. Tasche (2002). Expected shortfall: a natural coherent alternative to value at risk. *Economic Notes 31*, 379–388.

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

Chen, S. (2017). Beijing multi-site air quality. UCI machine learning repository. DOI: https://doi.org/10.24432/C5RK5G.

Chen, Y., S. Fang, and L. Lin (2024). Renewable composite quantile method and algorithm for nonparametric models with streaming data. *Statistics and Computing 34*, 43.

Chetverikov, D., Y. Liu, and A. Tsyvinski (2025). Weighted-average quantile regression. *Journal of Econometrics 252*, 106115.

Chiang, C.-T. and M.-Y. Huang (2012). New estimation and inference procedures for a single-index conditional distribution model. *Journal of Multivariate Analysis 111*, 271–285.

Daouia, A., I. Gijbels, and G. Stupfler (2019). Extremiles: a new perspective on asymmetric least squares. *Journal of the American Statistical Association 114*, 1366–1381.

Daouia, A., I. Gijbels, and G. Stupfler (2022). Extremile regression. *Journal of the American Statistical Association 117*, 1579–1586.

Debrauwer, D., I. Gijbels, and K. Herrmann (2025). On a general class of functionals: statistical inference and application to risk measures. *Electronic Journal of Statistics 19*, 2456–2510.

Dowd, K. and J. Cotter (2007). Exponential spectral risk measures. *The IUP Journal of Financial Economics V*, 57–66.

Henzi, A., G.-R. Kleger, and J. Ziegel (2023). Distributional (single) index models. *Journal of the American Statistical Association 118*, 489–503.

Huang, M.-Y. and C.-T. Chiang (2017). An effective semiparametric estimation approach for the sufficient dimension reduction model. *Journal of the American Statistical Association 112*, 1296–1310.

Jiang, R. and K. Yu (2024). Unconditional quantile regression for streaming data sets. *Journal of Business & Economic Statistics 42*, 1143–1154.

Li, Q., J. Lin, and J. Racine (2013). Optimal bandwidth selection for nonparametric conditional distribution and quantile functions. *Journal of Business & Economic Statistics 31*, 57–65.

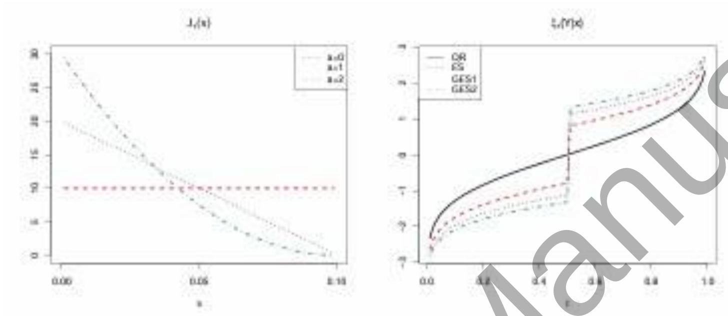
Shorack, G. and J. Wellner (1986). Empirical processes with applications in statistics. *New York: John Wiley & Sons*.

Wang, J., Y. Wu, C. Hai, R. Zhang, H. Liu, and Q. Wang (2025). Spectral risk-averse optimal operation of active distribution networks incorporating multimode switching of hybrid open points. *International Journal of Electrical Power & Energy Systems 172*, 111255.

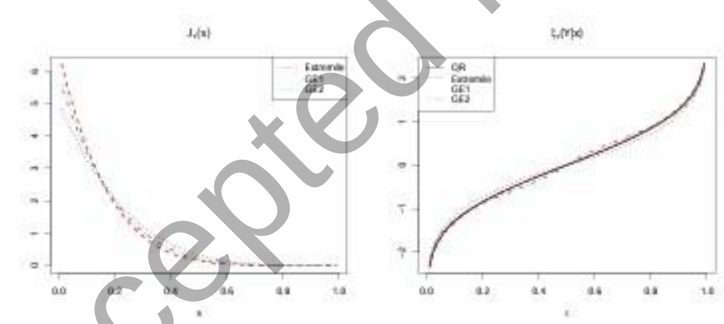
Wang, S. (2000). A class of distortion operators for pricing financial and insurance risks. *The Journal of Risk and Insurance 67*, 15–36.

Zhang, S., B. Guo, A. Dong, J. He, Z. Xu, and S. Chen (2017). Cautionary tales on air-quality improvement in beijing. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*.

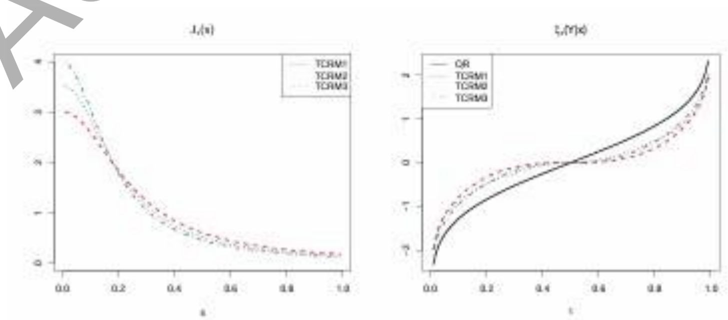
Figure 1: Plots of the weight functions $J_\tau(s)$ in (3.1)-(3.3) under $\tau = 0.1$ and of $\xi_\tau(Y \mid \boldsymbol{x})$ with $Q_{Y|\boldsymbol{x}}(\cdot)$ the quantile of the standard normal distribution.



(a) Generalized ES



(b) Generalized Extremile



(c) Coherent risk measure with truncated Cauchy density function

Figure 2: Plots of $J_{\tau}(s)$ in (3.2) and (3.3) with $\alpha_{\tau} = 0.5\tau^{-1} - 1$ and different τ s according to Table 1.

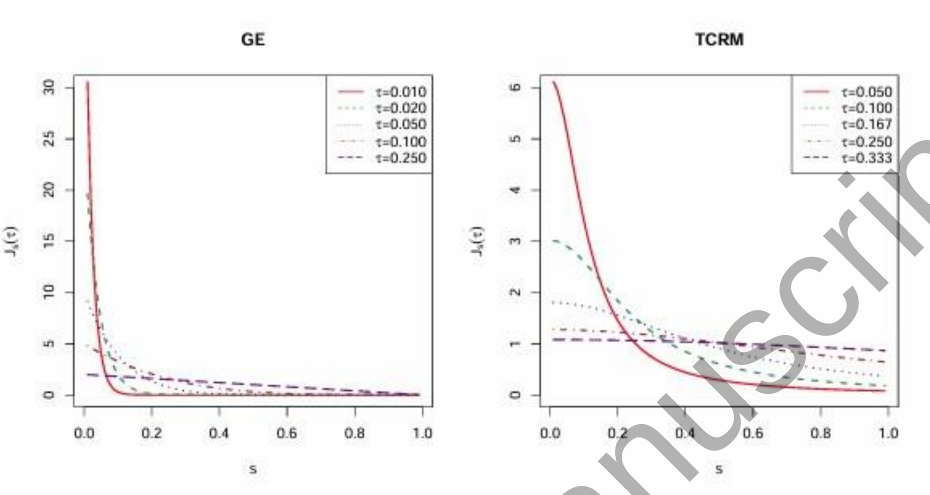


Figure 3: Several risk measures under different commonly used distributions for $\tau \in [0.90, 0.98]$.

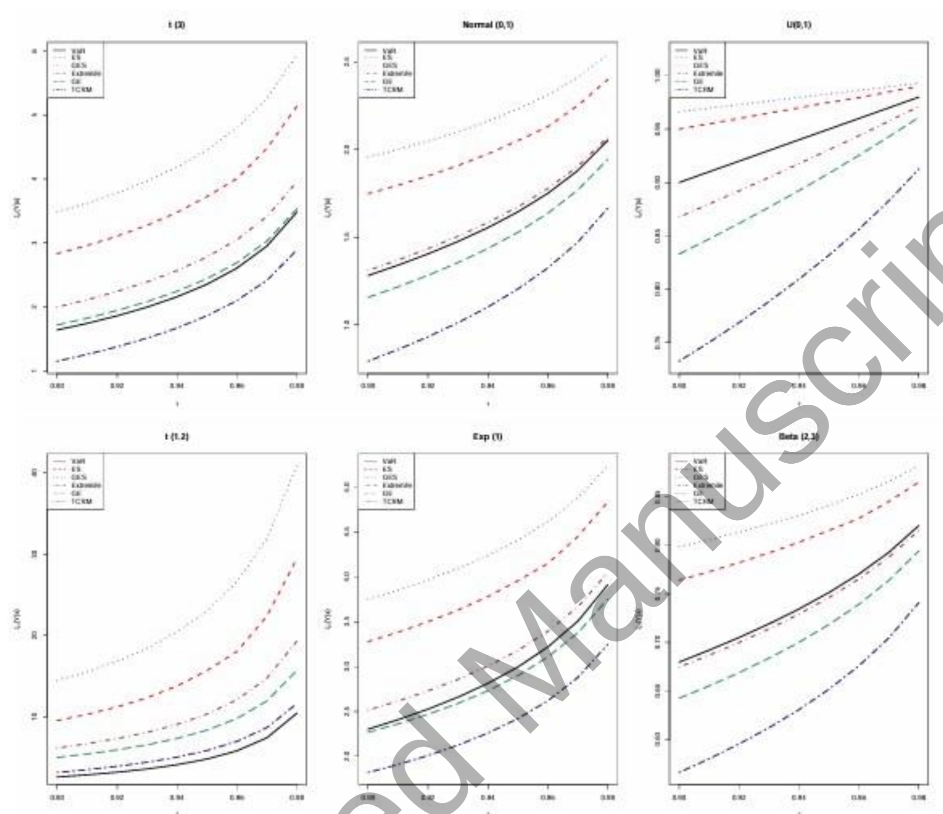


Figure 4: Several weight functions $J_{\tau}(s)$ under $\tau = 0.90, 0.95$.

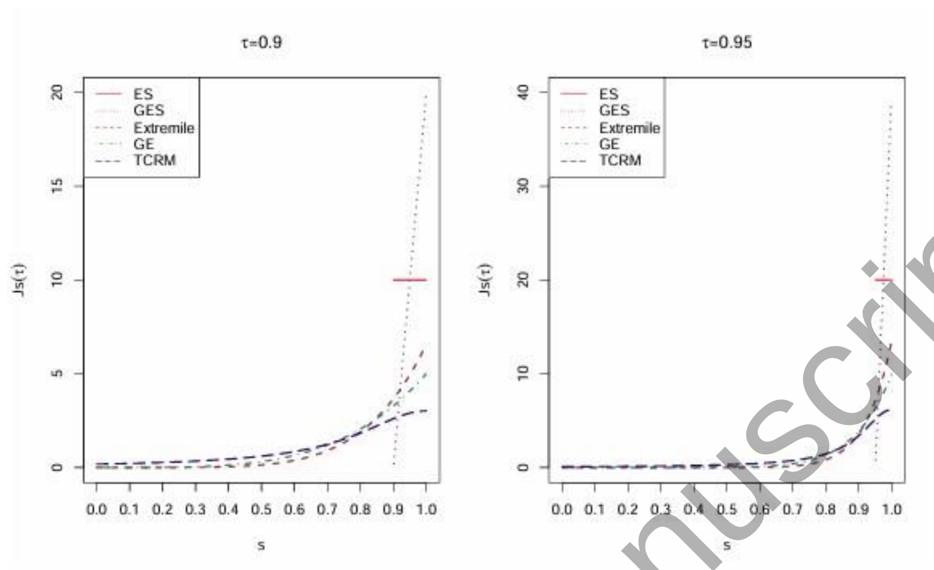


Figure 5: Histogram of $PM_{2.5}$ in the Beijing multi-site air quality dataset.

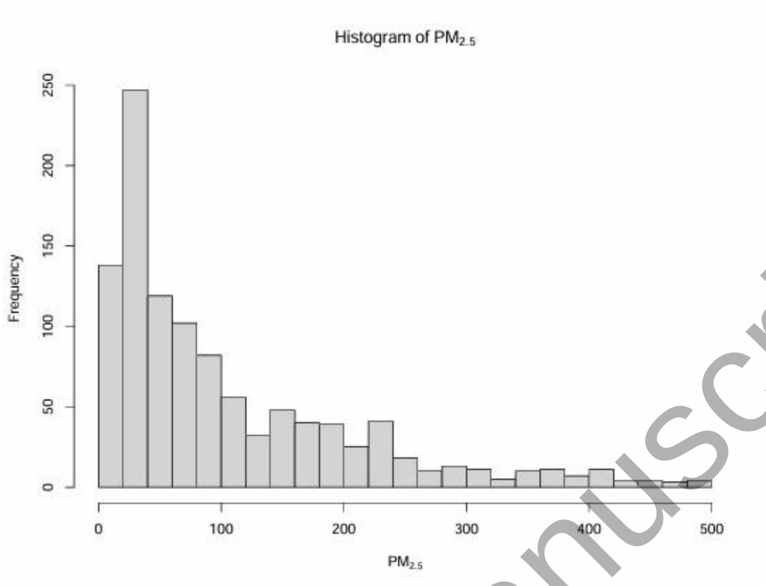


Table 1: τ values corresponding to Copies and Scale for GE and TCRM, respectively.

	Copies	2	5	10	25	50
GE	τ	0.250	0.100	0.050	0.020	0.010
	Scale	2	1	1/2	1/4	1/9
TCRM	τ	0.333	0.250	0.167	0.100	0.050

Table 2: The mean and standard deviation (in parentheses) of 500 replicates of RPADs (%) for $\tau = 0.05, 0.10, 0.90, 0.95$ when the error follows Normal(0,1), t(3) and Exp(1) with $\sigma(\boldsymbol{X}) = 1$.

Error	Method	τ =0.05	τ =0.1	τ =0.9	τ =0.95
Normal(0,1)	ES	3.14 (2.08)	2.61 (1.86)	2.73 (1.80)	3.36 (2.03)
	GES	3.59 (2.20)	2.97 (2.01)	3.17 (1.95)	3.90 (2.14)
	Extremile	2.54 (1.76)	2.17 (1.53)	2.19 (1.52)	2.66 (1.73)
	GE	2.36 (1.66)	2.08 (1.45)	2.06 (1.45)	2.45 (1.63)
	TCRM	2.15 (1.15)	1.95 (1.39)	1.90 (1.34)	2.18 (1.49)
t(3)	ES	6.99 (4.63)	5.05 (3.72)	5.73 (3.97)	7.74 (5.25)
	GES	8.77 (5.68)	6.40 (4.35)	7.15 (4.71)	9.33 (5.43)
	Extremile	4.97 (3.46)	3.65 (2.57)	4.04 (2.62)	5.55 (3.63)
	GE	4.33 (3.07)	3.27 (2.28)	3.57 (2.31)	4.83 (3.19)
	TCRM	3.59 (2.52)	2.78 (1.94)	2.97 (1.97)	3.96 (2.60)
Exp(1)	ES	0.94 (0.78)	0.81 (0.99)	4.43 (2.84)	5.64 (3.22)
	GES	0.95 (0.76)	0.96 (0.78)	5.25 (3.07)	6.72 (3.50)
	Extremile	1.03 (0.81)	1.18 (0.90)	3.27 (2.19)	4.30 (2.72)
	GE	1.07 (0.83)	1.25 (0.96)	2.94 (1.99)	3.86 (2.52)
	TCRM	1.19 (0.92)	1.45 (1.19)	2.44 (1.65)	3.21 (2.16)

Table 3: The mean and standard deviation (in parentheses) of 500 replicates of RPADs (%) for $\tau = 0.05, 0.10, 0.90, 0.95$ when the error follows Normal(0,1), t(3) and Exp(1) with $\sigma(\boldsymbol{X}) = |\boldsymbol{X}|$.

Error	Method	τ =0.05	τ =0.1	τ =0.9	τ =0.95
Normal(0,1)	ES	1.07 (1.19)	1.65 (2.00)	1.16 (1.06)	1.68 (2.09)
	GES	1.16 (1.27)	1.91 (2.28)	1.20 (1.12)	2.32 (1.94)
	Extremile	0.92 (1.04)	1.36 (1.63)	1.00 (0.86)	1.67 (1.39)
	GE	0.87 (0.99)	1.28 (1.51)	0.94 (0.81)	1.54 (1.31)
	TCRM	0.78 (0.90)	1.16 (1.34)	0.84 (0.74)	1.37 (1.18)
t(3)	ES	4.09 (4.20)	3.42 (4.45)	2.87 (2.79)	4.45 (3.27)
	GES	4.15 (4.01)	4.25 (5.55)	2.86 (2.83)	5.45 (4.11)
	Extremile	2.25 (3.22)	2.36 (3.31)	2.40 (1.86)	3.22 (2.34)
	GE	1.91 (2.83)	2.07 (2.88)	2.19 (1.65)	2.81 (2.08)
	TCRM	1.43 (2.26)	1.67 (2.33)	1.83 (1.33)	2.31 (1.72)
Exp(1)	ES	0.43 (0.43)	0.51 (0.51)	1.91 (1.68)	3.32 (2.53)
	GES	0.43 (0.43)	0.50 (0.52)	2.00 (1.82)	3.83 (3.04)
	Extremile	0.41 (0.41)	0.58 (0.69)	1.54 (1.22)	2.44 (1.84)
	GE	0.42 (0.41)	0.74 (0.62)	1.41 (1.10)	2.18 (1.65)
	TCRM	0.47 (0.41)	0.81 (0.71)	1.18 (0.89)	1.81 (1.35)

Table 4: The values of $\boldsymbol{\alpha}$ based on fitting data for different methods with $\tau = 0.05$.

Method	MSFT	AMZN	META	V	NVDA	CIEN	ICE	APD	CAH	WFC
QR	0.140	0.014	0.015	0.088	0.230	0.066	0.251	0.110	0.082	0.004
ES	0.072	0.018	0.000	0.140	0.020	0.100	0.190	0.190	0.063	0.207
GES	0.244	0.120	0.042	0.110	0.110	0.033	0.150	0.031	0.150	0.010
Extremile	0.006	0.058	0.099	0.120	0.011	0.110	0.090	0.018	0.368	0.120
GE	0.068	0.110	0.030	0.137	0.063	0.022	0.130	0.150	0.140	0.150
TCRM	0.061	0.050	0.190	0.021	0.140	0.047	0.100	0.028	0.100	0.263

Table 5: The SRs and PDs (%) based on testing data for different methods with $\tau = 0.05$.

Method	BR	QR	ES	GES	Extremile	GE	TCRM
SR	18.03	41.99	33.63	36.41	39.32	38.41	47.31
PD	-	53.17	49.21	51.98	50.79	50.40	56.75

Table 6: Average $PM_{2.5}$ with $\tau = 0.01$ to 0.99 for Beijing multi-site air quality dataset.

Method	$\tau = 0.01$	0.05	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.95	0.99
QR	9	19	27	41	55	68	83	99	117	145	193	230	309
Extremile	10	21	30	47	63	80	98	113	132	158	199	238	311
GE	11	24	36	54	70	84	98	109	124	146	184	222	298
TCRM	17	37	53	77	91	97	98	99	104	119	155	192	272

Table 7: Intervals of τ (from 0.01 to 0.99) corresponding to Average $PM_{2.5}$ for different methods in the Beijing multi-site air quality dataset, where 499 is the maximum observed $PM_{2.5}$ value.

Average $PM_{2.5}$	QR	ES	GES	Extremile	GE	TCRM
(0,35)	(0.01,0.16)	(0.01,0.32)	(0.01,0.49)	(0.01,0.13)	(0.01,0.10)	(0.01,0.05)
(35,75)	(0.16,0.45)	(0.32,0.50)	(0.49,0.50)	(0.13,0.37)	(0.10,0.34)	(0.05,0.19)
(75,499)	(0.45,0.99)	(0.50,0.99)	(0.50,0.99)	(0.37,0.99)	(0.34,0.99)	(0.19,0.99)