Extremal linear quantile regression with Weibull-type tails

Fengyang He

Department of Statistics, Jinan University, Guangzhou, China

Huixia Judy Wang

Department of Statistics, George Washington University, Washington DC, USA

Tiejun Tong

Department of Mathematics, Hong Kong Baptist University, Hong Kong

Abstract: **This study examines the estimation of extreme conditional quantiles for distributions with Weibull-type tails. We propose two families of estimators for the Weibull tail-coefficient, and construct an extrapolation estimator for the extreme conditional quantiles based on a quantile regression and extreme value theory. The asymptotic results of the proposed estimators are established. This work fills a gap in the literature on extreme quantile regressions, where many important Weibull-type distributions are excluded by the assumed strong conditions. A simulation study shows that the proposed extrapolation**

method provides estimations of the conditional quantiles of extreme orders that are more efficient and stable than those of the conventional method. The practical value of the proposed method is demonstrated through an analysis of extremely high birth weights.

Key words and phrases: Asymptotic normality, Extrapolation method, Extreme conditional quantiles, Linear quantile regression, Weibull-type distributions.

1. Introduction

Weibull-type distributions with a common extreme value index at zero form a rich family of light-tailed distributions, including, for example, the Gaussian, gamma, Weibull, and extended Weibull distributions. As noted in Beirlant and Teugels (1992), these distributions are conventionally used in the area of non-life insurance. Recently, de Wet et al. (2016) mentioned that these distributions may have a wide range of applications in other fields, such as hydrology, meteorology, and environmental sciences.

There is an extensive body of literature on the analysis of univariate Weibull-type tails, including the works of Berred (1991), Broniatowski (1993), Girard (2004), Gardes and Girard (2005, 2008), Diebolt et al. (2008), Goegebeur et al. (2010), and Goegebeur and Guillou (2011). In contrast, few studies have investigated the extremal behavior of Weibull-type tails under a regression setting. Among those that have, de Wet et al. (2016) used kernel statistics to estimate the tail coefficient of a Weibull-type distribution and the extreme conditional quantiles. Gardes and Girard (2016) focused only on estimating the tail-coefficient of a Weibull-type distribution, based on a kernel estimator of extreme conditional quantiles. It is well known that a nonparametric quantile regression (QR) is not stable on the boundary of the predictor support, and that estimations are challenging for multiple predictors, owing to the "curse of dimentionality;" see Daouia et al. (2013). This motivates us to investigate the extremal behavior of Weibull-type tails under a linear regression setting. To the best of our knowledge, there is no existing literature on extreme quantile estimations of Weibull-type tails under linear regression models.

Several studies have examined tail index regressions and extremal quantiles under a regression setup. Assuming Pareto-type distributions that correspond to positive extreme value indices, Wang and Tsai (2009) studied the tail index regression model by employing the logarithmic function to link the tail index to the linear predictor. Chernozhukov (2005) considered the extremal quantiles in a linear regression framework, and derived the asymptotic properties under three types of tail distributions corresponding to the extreme value index $\xi < 0$, $\xi = 0$, and $\xi > 0$, respectively. However, for condition R1 to hold, the case $\xi = 0$ is excluded for the simple location-scale shift model; see Example 3.2 in Chernozhukov (2005). Therefore, the results in Chernozhukov (2005) are not applicable to general models with Weibull-type distributions.

In this study, we develop new theory and methods with which to examine the extremal behavior of Weibull-type tails. We reconsider the important condition R1 in Chernozhukov (2005) in order to make it applicable for Weibull-type tails. Furthermore, we propose two families of estimators for the Weibull tail-coefficient based on a linear regression of quantiles, and construct an estimator for the extreme conditional quantiles using the extrapolation method. The proposed estimators do not suffer from the "curse of dimentionality," and can be readily applied to a wide range of studies with multiple predictors.

The remainder of this paper is organized as follows. In Section 2, we introduce the linear QR model, as well as several regularity assumptions that are needed to establish the asymptotic results of the new method. In Section 3, we propose two families of estimators for the Weibull tail-coefficient, and construct an efficient extrapolation estimator for the extreme conditional quantiles. The asymptotic results of the proposed estimators are also derived in this section. Miscellaneous issues are discussed in Section 4, including identifying Weibull-type tails, a comparison of the asymptotic efficiency of different estimators, and the validation of the model and the technical assumptions. In Section 5, we conduct a simulation study to e-valuate the finite-sample performance of the proposed estimators, and then compare the results with those of the conventional method. In Section 6, we illustrate the usefulness of the new method by using it to examine ex-tremely high birth weights of live infants born in the United States. All technical proofs are provided in the online Supplementary Material.

2. Model and assumptions

Let $\{(\mathbf{X}_i, Y_i), i = 1, \ldots, n\}$ be independent copies of the random vector $(\mathbf{X}, Y)$, where $\mathbf{X} = (1, X_2, \ldots, X_d)'$ is a d-dimensional covariate and Y is a one-dimensional response variable. For convenience, let $\mathbf{X}_{-1} = (X_2, \ldots, X_d)'$ denote the covariate $\mathbf{X}$ without the first component, $\mathcal{X}$ denote the support of $\mathbf{X}$, and $F_Y(y|\mathbf{x})$ be the continuous conditional distribution function of Y, given $\mathbf{X} = \mathbf{x}$. Denote $\bar{F}_Y(y|\mathbf{x}) = 1 - F_Y(y|\mathbf{x})$ and let $q_Y(\tau|\mathbf{x}) = \inf\{y : \bar{F}_Y(y|\mathbf{x}) \leq \tau\}$ be the $(1-\tau)$th conditional quantile of Y, given $\mathbf{X} = \mathbf{x}$, also referred to as the τth right-tailed conditional quantile.

In this study, we consider the following linear QR model:

$$q_Y(\tau|\mathbf{x}) = \mathbf{x}'\beta(\tau), \qquad \text{for all } \tau \in (0, \tau_U], \text{ for some } 0 < \tau_U < 1, \ \mathbf{x} \in \mathcal{X}, \tag{2.1}$$

where $\beta(\tau)$ is a vector of quantile coefficients. For any given τ, $\beta(\tau)$ can be estimated by

$$\hat{\beta}(\tau) = \arg\min_{\beta \in \mathbb{R}^d} \sum_{i=1}^{n} \rho_\tau(Y_i - \mathbf{X}_i'\beta), \tag{2.2}$$

where $\rho_\tau(u) = u\{\mathrm{I}(u > 0) - \tau\}$ is the asymmetric L_1 "check" function. Let τ_n be an intermediate quantile level in the sense that $\tau_n \to 0$ and $n\tau_n \to \infty$. It was shown in Chernozhukov (2005) that at the intermediate quantile level, the asymptotic normal theory still holds for $\hat{\beta}(\tau_n)$ and, hence, for the conventional conditional quantile estimator $\hat{q}_n(\tau_n|\mathbf{x}) = \mathbf{x}'\hat{\beta}(\tau_n)$ of $q_Y(\tau_n|\mathbf{x})$. Our main interest is the estimation of conditional quantiles at the extreme quantile level $\psi_n \to 0$, which satisfies $\psi_n \to 0$ and $\ln \psi_n / \ln \tau_n \to \kappa \in (1, \infty)$ as $\tau_n \to 0$. This allows ψ_n to go to zero at an arbitrarily fast rate; see Section 4.3. Therefore, the corresponding quantile $q_Y(\psi_n|\mathbf{x})$ is further in the right tail and more extreme than $q_Y(\tau_n|\mathbf{x})$. In such a case, the conventional quantile estimator $\hat{q}_n(\psi_n|\mathbf{x}) := \mathbf{x}'\hat{\beta}(\psi_n)$ for $q_Y(\psi_n|\mathbf{x})$ is often unreliable, owing to the sparsity of data in the extreme tails. As a result, obtaining precise estimates of the extreme quantiles remains a challenging task. Extreme value theory provides a valuable mathematical tool for solving this problem.

In this paper, we propose studying the linear QR model in (2.1) with Weibull-type tails using extreme value theory. To start with, let u be a random variable with the survival function $\bar{F}_u(z) := \mathrm{P}(u > z)$ and the upper endpoint $s_u^* = \infty$. Without loss of generality, we assume that $\bar{F}_u(\cdot)$ is continuous, differentiable, and strictly decreasing. Recall that $\bar{F}_u$ has a Weibull-type tail if there exists $\theta > 0$ such that, for all $\zeta > 0$,

$$\lim_{z\to\infty} \frac{\ln \bar{F}_u\,(\zeta z)}{\ln \bar{F}_u(z)} = \zeta^{1/\theta}. \tag{2.3}$$

The parameter θ is also referred to as the Weibull tail-coefficient; this controls the tail behavior such that a larger value of θ results in a slower decay of $\bar{F}_u$ to zero. Weibull-tailed distributions cover a wide class of light-tailed distributions in the Gumbel maximum domain, including the Gaussian ($\theta = 1/2$), exponential, gamma, logistic ($\theta = 1$), and Weibull distributions; see Section 4.3 for a more specific discussion.

For convenience, we denote the cumulative hazard function by $H_u(z) := -\ln \bar{F}_u(z)$, and the quantile function by $q_u(\tau) := \bar{F}_u^{-1}(\tau) = H_u^{-1}(\ln(1/\tau))$, for all $\tau \in (0,1)$. By (2.3), $H_u(\cdot)$ is a regularly varying function, with index $1/\theta$: that is,

$$\lim_{z\to\infty} \frac{H_u(\zeta z)}{H_u(z)} = \zeta^{1/\theta}, \text{ for all } \zeta > 0, \tag{2.4}$$

which we denote by $H_u(\cdot) \in \mathcal{RV}_\infty(1/\theta)$. Note that (2.4) also holds locally uniformly on $\zeta > 0$. By Proposition 0.1 in Resnick (1987), we have $H_u^{-1}(\cdot) \in$

$\mathcal{RV}_\infty(\theta)$. Hence, there exists a slowly varying function $l(\cdot)$, such that

$$H_u^{-1}(z) = z^\theta l(z), \text{ for } z > 0, \tag{2.5}$$

where $l(\cdot)$ satisfies that $\lim_{z\to\infty} l(\zeta z)/l(z) = 1$, for all $\zeta > 0$. In addition, because $H_u(z)$ is differentiable so that $H_u^{-1}(z)$ is differentiable, we can obtain that $\partial H_u^{-1}(z)/\partial z \in \mathcal{RV}_\infty(\theta - 1)$.

Throughout the paper, we use $a(t) \sim b(t)$ to represent $a(t)/b(t) \to 1$ when t tends to a constant or to infinity. To establish the asymptotic results of the estimators proposed in Section 3, we require the following regularity assumptions.

(C1) There exists a bounded vector $\beta_r \in \mathbb{R}^d$ and a survival function $\bar{F}_u$ of the Weibull-type tail with tail-coefficient θ, such that (i) $U = Y - \mathbf{X}'\beta_r$, with $s_U^* = \infty$; and (ii) $H_U(z|\mathbf{x}) \sim K(\mathbf{x})H_u(z)$ uniformly on $\mathbf{x} \in \mathcal{X}$ as $z \uparrow s_U^*$, where $s_U^* = \inf\left\{y : \bar{F}_U(y|\mathbf{x}) \le 0\right\}$ is the upper endpoint, and $H_U(z|\mathbf{x}) = -\ln \bar{F}_U(z|\mathbf{x})$, with $\bar{F}_U(z|\mathbf{x})$ being the conditional survival function of U, given $\mathbf{X} = \mathbf{x}$. Furthermore, $\bar{F}_U(z|\mathbf{x})$ is assumed to be continuous and strictly decreasing with respect to z, and $K(\cdot) > 0$ is a continuous bounded function on the support $\mathcal{X}$.

(C2) For any $k \in (0,1) \cup (1,\infty)$, $H_U^{-1}\left(-\ln(k\tau)|x\right)/H_U^{-1}\left(-\ln\tau|x\right) - 1 \sim \theta \ln k/\ln\tau$ as $\tau \to 0$.

(C3) $\mathcal{X}$ is a compact set in $\mathbb{R}^d$, and $\mathrm{E}(\mathbf{XX}')$ is a positive-definite matrix.

(C4) Under (C1) – (C3), we assume that

$$\frac{\partial H_U^{-1}(-\ln\tau|\mathbf{x})}{\partial\tau} \sim \frac{\partial H_u^{-1}(-\ln\tau/K(\mathbf{x}))}{\partial\tau} \quad \text{uniformly on } \mathbf{x}\in\mathcal{X}.$$

(C5) The slowly varying function $l(\cdot)$ in (2.5) satisfies the following: (i) there exist a constant $\varrho \le 0$ and a regularly varying function $b(z) \in \mathcal{RV}_\infty(\varrho)$ by (2.3.8) in de Haan and Ferreira (2006), and $b(z) \to 0$ as $z \to \infty$, such that locally uniformly on $\lambda \ge 1$,

$$\ln\left(\frac{l(\lambda z)}{l(z)}\right) = b(z)D_\varrho(\lambda)\left(1+o(1)\right), \text{as z} \to \infty,$$

where $D_\varrho(\lambda) = \int_1^\lambda t^{\varrho-1}dt$; (ii) $l(z) = c\exp\{\int_1^z \varepsilon(t)/tdt\}$, where $c > 0$ and $\varepsilon : (0,\infty) \to \mathbb{R}$ is a continuous function, with $\varepsilon(t) \to 0$ as $t \to \infty$.

Remark 1. Condition (C1) implies that for any $\mathbf{x} \in \mathcal{X}$, the conditional cumulative hazard function $H_U(\cdot|\mathbf{x})$ and the univariate cumulative hazard $H_u(\cdot)$ are tail equivalent up to a constant. Under (C1), for large z, we can write $H_U(z|\mathbf{x}) = K(\mathbf{x})H_u(z)(1+\alpha(z|\mathbf{x}))$, where $\alpha(z|\mathbf{x}) \to 0$ as $z \to \infty$ uniformly on $\mathbf{x} \in \mathcal{X}$. Noting too that $H_u^{-1}(\cdot) \in \mathcal{RV}_\infty(\theta)$, we thus have

$$\begin{aligned} H_U^{-1}(-\ln\tau|\mathbf{x}) &= H_u^{-1}\left(\frac{-\ln\tau}{K(\mathbf{x})\left(1+\alpha\left(H_U^{-1}(-\ln\tau|\mathbf{x})\,|\mathbf{x}\right)\right)}\right) \\ &\sim H_u^{-1}(-\ln\tau/K(\mathbf{x})), \end{aligned} \tag{2.6}$$

and $H_u^{-1}(-\ln\tau/K(\mathbf{x})) \sim H_u^{-1}(-\ln(k\tau)/K(\mathbf{x}))$ as $\tau \to 0$, for any $k \in (0,1)\cup(1,\infty)$. This leads to $H_U^{-1}(-\ln(k\tau)|\mathbf{x}) \sim H_u^{-1}(-\ln(k\tau)/K(\mathbf{x}))$ and $H_U^{-1}(-\ln(k\tau)|\mathbf{x}) \sim H_U^{-1}(-\ln\tau|\mathbf{x})$. Condition (C2) further assumes that $H_U^{-1}(-\ln(k\tau)|\mathbf{x})/H_U^{-1}(-\ln\tau|\mathbf{x}) - 1$ and $\theta\ln k/\ln\tau$ are asymptotically e-quivalent; that is, they converge to zero at the same rate. The rationality of (C2) is discussed in Section 4.3.

Conditions (C1), (C3), and (C4) can be regarded as adaptations of conditions R1−R3 in Chernozhukov (2005) to Weibull-type tails. Condition (C5)(i) is essentially the same as that in de Wet et al. (2016) and Girard (2004). The latter is the second-order condition on $l(\cdot)$, with the second-order parameter $\rho \le 0$ that controls the convergence rate of $l(\lambda z)/l(z)$ toward one. The closer ρ is to zero, the slower is the convergence rate. Hence, condition (C5)(i) plays a crucial role in deriving the asymptotic results of our proposed estimators. Condition (C5)(ii) is essentially the same as condition (A.2) in Gardes and Girard (2016), which is a special case of the Karamata representation; see Theorem B.1.6 in de Haan and Ferreira (2006) for regularly varying functions. The function $\varepsilon(\cdot)$ in (C5)(ii) determines the speed of the convergence of the slowly varying function $l(\cdot)$.

3. Proposed estimators

In this section, we propose an extrapolation estimator for extreme conditional quantiles. We also develop two types of estimators for the Weibull tail-coefficient based on the regression quantiles.

For ease of notation, we denote $q_U(\tau|\mathbf{x}) = \bar{F}_U^{-1}(\tau|\mathbf{x})$, for all $\tau \in (0,1)$. By (2.6) and condition (C1), we have

$$q_Y(\tau|\mathbf{x}) = q_U(\tau|\mathbf{x}) + \mathbf{x}'\beta_r = q_u\left(\tau^{1/K(\mathbf{x})}\right)(1+\alpha(\tau)) + \mathbf{x}'\beta_r,$$

for some $\alpha(\tau) \to 0$ as $\tau \to 0$. Therefore,

$$\varpi\left(s, \tau^{1/K(\mathbf{x})}\right) = \frac{q_Y(s\tau|\mathbf{x})}{q_Y(\tau|\mathbf{x})} \frac{q_u\left(\tau^{1/K(\mathbf{x})}\right)}{q_u\left((s\tau)^{1/K(\mathbf{x})}\right)} - 1 \to 0,$$

for all $s > 0$ as $\tau \to 0$.

3.1 Estimation of extreme conditional quantiles

Let $\tau \in (0,1)$ be sufficiently small. Then, by (2.5) and similar arguments to those used in the proof of Lemma 2 in Gardes and Girard (2016), for any given $s \in (0,1]$, we have

$$\begin{aligned}
\ln q_Y(s\tau|\mathbf{x}) - \ln q_Y(\tau|\mathbf{x}) &= \ln\left(\frac{q_u\left((s\tau)^{1/K(\mathbf{x})}\right)}{q_u\left(\tau^{1/K(\mathbf{x})}\right)}\right) + \ln\left[1 + \varpi\left(s, \tau^{1/K(\mathbf{x})}\right)\right] \\
&= \ln\left(\frac{H_u^{-1}\left(-\ln(s\tau)/K(\mathbf{x})\right)}{H_u^{-1}\left(-\ln(\tau)/K(\mathbf{x})\right)}\right) + \ln\left[1 + \varpi\left(s, \tau^{1/K(\mathbf{x})}\right)\right] \\
&= \theta\left[\ln_{-2}(s\tau) - \ln_{-2}(\tau)\right] + T(s, \tau|\mathbf{x}), \qquad (3.1)
\end{aligned}$$

where $\ln_{-2}(z) := \ln[\ln(1/z)]$ and $T(s,\tau|\mathbf{x}) = \ln[l(-\ln(s\tau)/K(\mathbf{x}))/l(-\ln(\tau)/K(\mathbf{x}))] + \ln[1+\varpi(s,\tau^{1/K(\mathbf{x})})] \to 0$ as $\tau \to 0$. Then, for any $s \in (0,1]$,

$$\frac{q_Y\left(s\tau|\mathbf{x}\right)}{q_Y\left(\tau|\mathbf{x}\right)} - \left(\frac{\ln(s\tau)}{\ln\tau}\right)^{\theta} \to 0. \tag{3.2}$$

Suppose $\hat{\theta}_n$ is some consistent estimator of θ (see Section 3.2). Then, we can estimate $q_Y(\psi_n|\mathbf{x})$ by the following extrapolation estimator:

$$\hat{q}_{n,\mathrm{E}}\left(\psi_n|\mathbf{x}\right) = \hat{q}_n(\tau_n|\mathbf{x})\left(\ln\psi_n/\ln\tau_n\right)^{\hat{\theta}_n}, \tag{3.3}$$

where $\hat{q}_n(\tau_n|\mathbf{x}) = \mathbf{x}'\hat{\beta}(\tau_n)$, and $\hat{\beta}(\tau_n)$ is defined in (2.2) at the intermediate quantile level τ_n.

3.2 Estimation of the Weibull tail-coefficient

In this section, we propose several estimators for the Weibull tail-coefficient θ. For any given $r \in (0,1)$, let $s_j = r^{j-1}$, for $j = 1,\ldots,J$, where J is a positive integer. By (3.1) and the fact that $\ln(1+u) \sim u$ as $u \to 0$, it follows that

$$\ln q_Y\left(s_{j+1}\tau|\mathbf{x}\right) - \ln q_Y\left(s_j\tau|\mathbf{x}\right) - \left\{\frac{\ln(1/r)}{\ln(1/\tau)}\right\}\theta \to 0, \text{ as } \tau \to 0.$$

Let $\mathbf{x} \in \mathcal{X}$ be a given covariate vector. Based on the conventional conditional quantile estimation at the intermediate quantile levels, namely, $\hat{q}_n(s_j\tau_n|\mathbf{x}) = \mathbf{x}'\hat{\beta}(s_j\tau_n)$, for $j = 1,\ldots,J$, we can construct a weighted estimator of θ, as follows:

$$\hat{\theta}_{n,\mathrm{P}}(\mathbf{x}) = \frac{\ln(1/\tau_n)}{\ln(1/r)} \sum_{j=1}^{J-1} \omega_j \left[\ln \hat{q}_n(s_{j+1}\tau_n|\mathbf{x}) - \ln \hat{q}_n(s_j\tau_n|\mathbf{x})\right],$$

where $\{\omega_j\}_{j=1}^{J-1}$ is a sequence of nonnegative weights summing to one. The estimator $\hat{\theta}_{n,P}(\mathbf{x})$ follows a similar spirit to the refined Pickand estimator introduced in Daouia et al. (2013) for the conditional extreme value index.

Similarly to Daouia et al. (2013), we consider two special cases of $\hat{\theta}_{n,P}(\mathbf{x})$. The first case uses constant weights $\omega_1 = \cdots = \omega_{J-1} = 1/(J-1)$, yielding

$$\hat{\theta}^{\mathrm{c}}_{n,\mathrm{P}}(\mathbf{x}) = \frac{\ln(1/\tau_n)}{(J-1)\ln(1/r)} \left[\ln \hat{q}_n\left(r^{J-1}\tau_n|\mathbf{x}\right) - \ln \hat{q}_n(\tau_n|\mathbf{x})\right].$$

In the second case, we consider linear weights $\omega_j = 2(J-j)/\{(J-1)J\}$, for $j = 1, \dots, J-1$, which results in

$$\hat{\theta}^{\mathrm{l}}_{n,\mathrm{P}}(\mathbf{x}) = \frac{2\ln(1/\tau_n)}{J(J-1)\ln(1/r)} \sum_{j=1}^{J-1} \left[\ln \hat{q}_n(s_j\tau_n|\mathbf{x}) - \ln \hat{q}_n(\tau_n|\mathbf{x})\right].$$

For comparison, we also introduce an estimator analogous to that proposed in Gardes and Girard (2016):

$$\hat{\theta}_{n,\mathrm{H}}(\mathbf{x}) = \frac{\ln(1/\tau_n)}{\sum_{j=1}^{J} \ln(1/s_j)} \sum_{j=1}^{J} \left[\ln \hat{q}_n(s_j\tau_n|\mathbf{x}) - \ln \hat{q}_n(\tau_n|\mathbf{x})\right], \qquad (3.4)$$

where $\{s_j : 0 < s_J < \cdots < s_1 \leq 1\}$ is a decreasing sequence. The estimator $\hat{\theta}_{n,H}(\mathbf{x})$ is an adaptation of the Hill estimator (Hill, 1975) for univariate heavy-tailed data; see also Daouia et al. (2011) and Wang et al. (2012) for Hill-type estimators under a regression setup.

Remark 2. From a theoretical point of view, we can use $\hat{\theta}_n(\mathbf{x})$ to estimate the coefficient θ at any given $\mathbf{x} \in \mathcal{X}$. However, given the sample data $\{\mathbf{x}_i\}_{i=1}^n$, our experience suggests that $\hat{\theta}_n(\bar{\mathbf{x}})$, with $\bar{\mathbf{x}} = \sum_{i=1}^n \mathbf{x}_i/n$, is often more stable than $\hat{\theta}_n(\mathbf{x})$ when $\mathbf{x}$ is not in the centroid of the design space. This is mainly because there are often more data around $\bar{\mathbf{x}}$, and the conventional conditional quantile estimator at $\bar{\mathbf{x}}$ is less susceptible to quantile crossing issues; see Koenker (2005, Chap. 2.5).

3.3 Asymptotic results

Here, we establish the asymptotic results of the proposed estimators. Throughout, we assume that $\tau_n \to 0$ and $n\tau_n \to \infty$ as $n \to \infty$. For any $s > 0$, define

$$\tilde{q}_n\left(s|\mathbf{x}\right) = \sqrt{n\tau_n}\ln\left(1/\tau_n\right)\left(\frac{\hat{q}_n(s\tau_n|\mathbf{x})}{q_Y(s\tau_n|\mathbf{x})} - 1\right).$$

Let $\xrightarrow{d}$ and $\stackrel{\mathrm{d}}{=}$ denote "convergence in distribution" and "equality in distribution," respectively.

We first present the asymptotic joint distribution of the random vector $(\tilde{q}_n(s_1|\mathbf{x}), \ldots, \tilde{q}_n(s_J|\mathbf{x}))$, for any given $\mathbf{x} \in \mathcal{X}$ and a positive sequence $s_j \in (0, 1]$, for $j = 1, \ldots, J$.

Theorem 1. *Suppose conditions* (C1)−(C5) *hold. For all* $\mathbf{x} \in \mathcal{X}$, *if* $\tau_n \to 0$ *as* $n \to \infty$, *such that* $n\tau_n \to \infty$, *then*

$$\left(\tilde{q}_n\left(s_1|\mathbf{x}\right), \ldots, \tilde{q}_n\left(s_J|\mathbf{x}\right)\right)' \xrightarrow{d} \left(q_\infty\left(s_1|\mathbf{x}\right), \ldots, q_\infty\left(s_J|\mathbf{x}\right)\right)' \stackrel{d}{=} \mathrm{N}(\mathbf{0}, \Sigma_{q(x)}),$$

where $(\Sigma_{q(\mathbf{x})})_{j,j'} = \theta^2(\mathbf{x}'\Omega_1\mathbf{x})H^{-2}(\mathbf{x})(\max(s_j, s_{j'}))^{-1}$, *for* $j,\ j' = 1, \ldots, J$, $\Omega_1 = \mathcal{Q}_H^{-1}\mathcal{Q}_{\mathbf{X}}\mathcal{Q}_H^{-1}$, $\mathcal{Q}_{\mathbf{X}} = \mathrm{E}(\mathbf{X}\mathbf{X}')$, $\mathcal{Q}_H = \mathrm{E}[(H(\mathbf{X}))^{-1}\mathbf{X}\mathbf{X}']$, *and* $H(\mathbf{x}) = [K(\mu_{\mathbf{X}})/K(\mathbf{x})]^\theta$, *with* $\mu_{\mathbf{X}} = \mathrm{E}(\mathbf{X})$.

Theorems 2 and 3 present the asymptotic results of the two proposed Weibull tail-coefficient estimators: the Pickand-type estimator $\hat{\theta}_{n,P}(\mathbf{x})$, and the Hill-type estimator $\hat{\theta}_{n,H}(\mathbf{x})$ with $\mathbf{x} \in \mathcal{X}$ being a given design vector.

Theorem 2. *Suppose conditions* (C1) − (C5) *hold. Let* $s_j = r^{j-1}$, *for* $j = 1, \ldots, J$, *where* $r \in (0, 1)$. *For any* $\mathbf{x} \in \mathcal{X}$, *if* $\sqrt{n\tau_n} \max\big(1/\ln(1/\tau_n), |b(\ln(1/\tau_n))|\big) \to 0$, *and* $\sqrt{n\tau_n}\ln(1/\tau_n)\ \max_{j=1,\ldots,J} |\varpi(s_j, {\tau_n}^{1/K(\mathbf{x})})| \to 0$, *then*

$$\sqrt{n\tau_n}\left(\hat{\theta}_{n,\mathrm{P}}(\mathbf{x}) - \theta\right) \xrightarrow{d} \mathrm{N}\big(0, (\ln r)^{-2}W'\Sigma_{q(\mathbf{x})}W\big),$$

where $W = (w_0 - w_1, \ldots, w_{j-1} - w_j, \ldots, w_{J-1} - w_J)'$, *with* $w_0 = w_J = 0$.

Theorem 3. *Suppose conditions* (C1) − (C5) *hold. Let* $1 = s_1 > s_2 > \cdots > s_J > 0$ *be a positive decreasing sequence. For any* $\mathbf{x} \in \mathcal{X}$, *if* $\sqrt{n\tau_n}\ln(1/\tau_n)\max_{j=1,\ldots,J}\ |\varpi(s_j, {\tau_n}^{1/K(\mathbf{x})})| \to 0$ *and* $\sqrt{n\tau_n}\max\big(1/\ln(1/\tau_n),$

$|b(\ln(1/\tau_n))|\big) \to 0$, *then*

$$\sqrt{n\tau_n}\left(\hat{\theta}_{n,\mathrm{H}}(\mathbf{x}) - \theta\right) \xrightarrow{d} \mathrm{N}\big(0, \Lambda_J H^{-2}(\mathbf{x})\theta^2\left(\mathbf{x}'\Omega_1\mathbf{x}\right)\big),$$

where

$$\Lambda_J = \left(\sum_{j=1}^{J}\left[\{2(J-j)+1\}/s_j\right] - J^2\right)\left(\sum_{j=1}^{J}\ln\left(1/s_j\right)\right)^{-2}. \qquad (3.5)$$

For the Hill-type estimator, in practice, we can choose $s_j = 1/j$, as in Daouia et al. (2011). Consequently, $\Lambda_J = J(J-1)(2J-1)/(6\ln^2(J!))$. In this case, Λ_J is a convex function of J, and is minimized at $J = 9$, with $\Lambda_9 = 1.245$. Throughout the paper, we use $\hat{\theta}_{n,H}(\mathbf{x})$ with the "optimal" tuning parameters $s_j = 1/j$ and $J = 9$.

Finally, we establish the asymptotic normality of the proposed extrapolation estimator for the extreme conditional quantitle, $\hat{q}_{n,\mathrm{E}}(\psi_n|\mathbf{x})$, based on an asymptotically normal tail-coefficient estimator $\hat{\theta}_n$, which can be either the Pickand- or the Hill-type.

Theorem 4. *Suppose conditions* (C1) − (C5) *hold, and* $\kappa_n := \ln\psi_n/\ln\tau_n \to \kappa \in (1,\infty)$ *as* $n \to \infty$. *Let* $\hat{\theta}_n$ *be an estimator of* θ *satisfying* $\sqrt{n\tau_n}(\hat{\theta}_n - \theta) \xrightarrow{d} \mathrm{N}(0,\sigma_\theta^2)$, *with* $\sigma_\theta^2 > 0$. *Then, for any* $\mathbf{x} \in \mathcal{X}$, *if* $\sqrt{n\tau_n}\max\{|b(\ln(1/\tau_n))|, |\varpi(\psi_n/\tau_n, {\tau_n}^{1/K(\mathbf{x})})|\} \to 0$, *we have*

$$\frac{\sqrt{n\tau_n}}{\ln\kappa_n}\left(\frac{\hat{q}_{n,\mathrm{E}}(\psi_n|\mathbf{x})}{q_Y(\psi_n|\mathbf{x})} - 1\right) \xrightarrow{d} \mathrm{N}(0,\sigma_\theta^2).$$

4. Miscellaneous issues

4.1 Identifying Weibull-type tails

The expression in (3.1) suggests that if the conditional distribution of Y has a Weibull-type tail, then $\ln(q_Y(\tau|\mathbf{x}))$ will be approximately linear in $\ln_{-2}(\tau)$, with slope θ. Motivated by this, we consider a graphical tool to check the assumption of Weibull-type tail for the conditional distribution of Y. Specifically, given the sample data $\{(\mathbf{x}_i, y_i)\}_{i=1}^n$, we can obtain the conventional estimator $\hat{q}_n(\tau_j|\bar{\mathbf{x}})$ at the sample mean $\bar{\mathbf{x}}$ for a grid of small quantile levels $\tau_1, \ldots, \tau_m$. Then, we can draw a quantile plot by plotting $\ln(\hat{q}_n(\tau_j|\bar{\mathbf{x}}))$ against $\ln_{-2}(\tau_j)$, with $j = 1, \ldots, m$. If the distribution has a Weibull-type tail, the points should lie roughly on a straight line. The graphical diagnosis at one design point, $\bar{\mathbf{x}}$, is reasonable, because condition (C1) implies that, for any $\mathbf{x}$, $\mathbf{x}' \in \mathcal{X}$, $z \mapsto H_U(z|\mathbf{x})$ and $z \mapsto H_U(z|\mathbf{x}')$ are tail equivalent up to a constant. The above steps are described in further detail in the case study in Section 6.

4.2 Comparison of asymptotic variances

Theorem 4 suggests that the estimation accuracy of the proposed extreme quantile estimator $\hat{q}_{n,\mathrm{E}}(\psi_n|\mathbf{x})$ depends heavily on that of the Weibull tail-

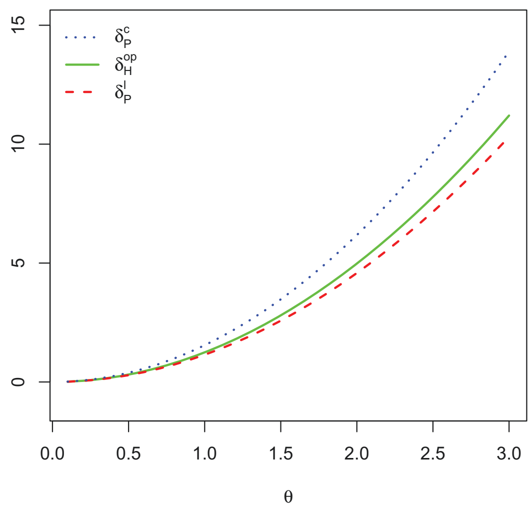

Figure 1: Plots of δ_P^{c}, δ_P^{l}, and δ_H^{op} against the Weibull tail-coefficient θ.

coefficient estimator. Define $\delta_P = W'\Sigma W/(\ln r)^2$, with $\Sigma_{j,j'} = \theta^2/(r^{j-1} \vee r^{j'-1})$, for $j,\ j' = 1, \ldots, J$, and $\delta_H = \Lambda_J \theta^2$. By Theorems 3 and 4, we have $W'\Sigma_{q(\mathbf{x})}W/(\ln r)^2 = [(\mathbf{x}'\Omega_1\mathbf{x})/H^2(\mathbf{x})]\delta_P$ and $\Lambda_J\theta^2(\mathbf{x}'\Omega_1\mathbf{x})/H^2(\mathbf{x}) = [(\mathbf{x}'\Omega_1\mathbf{x})/H^2(\mathbf{x})]\delta_H$. Therefore, to compare the asymptotic variances of $\hat{\theta}_{n,P}(\mathbf{x})$ and $\hat{\theta}_{n,H}(\mathbf{x})$, it suffices to compare δ_P and δ_H, where both are quadratic functions of θ. For convenience, denote δ_P^c and δ_P^l as special cases of δ_P for constant and linear weights, respectively, and δ_H^{op} as a special case of δ_H with the "optimal" tuning parameters $s_j = 1/j$ and $J = 9$. For the Pickand-type estimators, we select the tuning parameters J and r by searching over $\mathcal{J} = \{2, 3, \ldots, 10\}$ and $\mathcal{R} = \{0.01, 0.02, \ldots, 0.99\}$, respestively, to identify the optimal pair that gives the smallest δ_P. Figure 1 shows that the three Weibull tail-coefficient estimators have similar efficiency for small $\theta \in (0, 0.5]$, but that for larger θ, $\hat{\theta}_{n,H}$ and $\hat{\theta}^l_{n,P}$ tend to be more efficient than $\hat{\theta}^c_{n,P}$.

4.3 Model validation

In this section, we show that conditions (C1) and (C2) are very general, and that they cover a wide range of conventional regression models as special cases. We also present several important Weibull-type distributions that fulfill the conditions in (C5). For illustration, we first present two conventional regression models that satisfy condition (C1).

(M1) Consider the location shift model

$$Y = \mathbf{X}'\beta + u,$$

where u is independent of $\mathbf{X}$, and the survival function $\bar{F}_u(\cdot)$ of u has a Weibull-type tail. This model is a special case of (C1), where $X'\beta_r = \mathbf{X}'\beta$, $U \equiv Y - \mathbf{X}'\beta = u$, and $K(x) = 1$, given $\mathbf{X} = \mathbf{x}$. Moreover, $\bar{F}_U(z|\mathbf{x}) = \bar{F}_u(z)$, for any $z \in \mathbb{R}$, such that $H_U(z|\mathbf{x}) \sim K(x)H_u(z)$ uniformly on $\mathbf{x} \in \mathcal{X}$ as $z \to \infty$.

(M2) Consider the heteroscedastic model

$$Y = \mathbf{X}'\beta + (\mathbf{X}'\xi)\, u,$$

where the scale function $\mathbf{x}'\xi > 0$, for any $\mathbf{X} = \mathbf{x} \in \mathcal{X}$, u is independent of $\mathbf{X}$, and the survival function $\bar{F}_u(\cdot)$ of u has a Weibull-type tail. It is easy to see that

$$\bar{F}_Y^{-1}(\tau|\mathbf{X}) = \mathbf{X}'\beta + (\mathbf{X}'\xi)\, \bar{F}_u^{-1}(\tau).$$

Then, for $\mathbf{X}'\beta_r = \mathbf{X}'\beta$ and $U \equiv Y - \mathbf{X}'\beta = (\mathbf{X}'\xi)\, u$, we have

$$\begin{aligned} H_U(z|\mathbf{x}) &= -\ln \mathrm{P}\left((\mathbf{x}'\xi)u > z|\mathbf{x}\right) \\ &= H_u\left((\mathbf{x}'\xi)^{-1} z\right) \\ &\sim (\mathbf{x}'\xi)^{-1/\theta} H_u(z), \end{aligned}$$

as $z \to \infty$, by (2.3). Thus, condition (C1) is satisfied, with $K(\mathbf{x}) = (\mathbf{x}'\xi)^{-1/\theta}$, for any $\mathbf{x} \in \mathcal{X}$.

Next, we present some important Weibull-type distributions as examples that satisfy condition (C5).

(E1) Let u follow the Gaussian distribution $\mathrm{N}(\mu, \sigma^2)$, with $\sigma > 0$.

We have $H_u^{-1}(z) = z^{1/2} l(z)$, and an asymptotic expansion of $l(\cdot)$ as

$$l(z) = \sqrt{2}\sigma - \frac{\sigma}{2^{3/2}} \frac{\ln z}{z} + O(1/z).$$

This leads to $\theta = 1/2$, $\rho = -1$, $c = \sqrt{2}\sigma \exp(-1/4)$, and $b(z) = \varepsilon(z) = \ln z/(4z)$.

(E2) Let u follow the gamma distribution $\Gamma(\beta, \alpha)$, with $\alpha, \beta > 0$.

We have the density function $f(z) = \beta^{\alpha}\Gamma^{-1}(\alpha)z^{\alpha-1}\exp(-\beta z)$, and $H_u^{-1}(z) = zl(z)$, with

$$l(z) = \begin{cases} \frac{1}{\beta} & \text{if } \alpha = 1, \\ \frac{1}{\beta} + \frac{\alpha-1}{\beta}\frac{\ln z}{z} + O(1/z) & \text{if } \alpha \neq 1. \end{cases}$$

This leads to $\theta = 1$, $\rho = -1$, $c = \exp(\alpha - 1)/\beta$, and $b(z) = \varepsilon(z) = (1-\alpha)\ln z/z$.

(E3) Let u follow the Weibull distribution $\mathrm{W}(\alpha, \lambda)$, with $\alpha, \lambda > 0$.

We have the density function $f(z) = (\alpha/\lambda)(z/\lambda)^{\alpha-1}\exp(-(z/\lambda)^{\alpha})$, $H_u^{-1}(z) = \lambda z^{1/\alpha}$, and $l(z) = \lambda$, for all $z > 0$. This leads to $\theta = 1/\alpha$, $\rho = -\infty$, $c = \lambda$, and $b(z) = \varepsilon(z) = 0$.

(E4) Let u follow the extended Weibull distribution $\mathrm{EW}(\alpha, \beta)$, with $\alpha > 0$ and $\beta \in \mathbb{R}$.

The survival function of u is given by $\bar{F}_u(z) = r(z)\exp(-z^{\alpha})$, where $r(\cdot) \in \mathcal{RV}_{\infty}(\beta)$. In addition, $H_u^{-1}(z) = z^{1/\alpha}l(z)$, with

$$l(z) = 1 + \frac{\beta}{\alpha^2}\frac{\ln z}{z} + O(1/z).$$

This leads to $\theta = 1/\alpha$, $\rho = -1$, $c = \exp(\beta/\alpha^2)$, and $b(z) = \varepsilon(z) = -\beta(\ln z)/(\alpha^2 z)$.

(E5) Let u follow the modified Weibull distribution MW(α), with $\alpha > 0$. Let $V \sim$ W$(\alpha, 1)$ and $u = V \ln V$. Thus, $H_u^{-1}(z) = z^{1/\alpha} l(z)$, with $l(z) = \alpha \ln z$. This leads to $\theta = 1/\alpha$, $\rho = 0$, $c = \alpha$, and $b(z) = \varepsilon(z) = 1/\ln z$.

In what follows, we show that (C2) holds for both the location shift model (M1) and the heteroscedastic model (M2) with Weibull-tailed errors. By (2.5) and (C5), and after some calculation, we have that

$$\frac{H_u^{-1}(-\ln(k\tau))}{H_u^{-1}(-\ln \tau)} - 1 \sim \frac{\theta \ln k}{\ln \tau} \quad \text{as} \quad \tau \to 0. \tag{4.1}$$

Note that $H_U^{-1}(-\ln \tau|\mathbf{x}) = H_u^{-1}(-\ln \tau)$, for any $\tau \in (0, 1)$ in (M1). Thus, it is clear that condition (C2) holds under (M1). Second, by $H_U^{-1}(-\ln \tau|\mathbf{x}) = (\mathbf{x}'\xi) H_u^{-1}(-\ln \tau)$ and $\mathbf{x}'\xi > 0$ in (M2), it is easy to check that condition (C2) is also fulfilled under (M2) by using (4.1).

To verify the conditions required in Theorems 1–4, we need to determine the appropriate rates of τ_n and ψ_n. Specifically, we need that as $n \to \infty$, τ_n satisfies $\tau_n \to 0$, $n\tau_n \to \infty$, $\sqrt{n\tau_n} \ln(1/\tau_n) \max_{j=1,\ldots,J} |\varpi(s_j, \tau_n^{1/K(\mathbf{x})})| \to 0$, and $\sqrt{n\tau_n} \max\left\{1/\ln(1/\tau_n), |b(\ln(1/\tau_n))|, |\varpi(\psi_n/\tau_n, \tau_n^{1/K(\mathbf{x})})|\right\} \to 0$, for all $\mathbf{x} \in \mathcal{X}$. The condition $n\tau_n \to \infty$ implies that τ_n should be of a larger order than $1/n$. In Propositions 1 and 2, provided in the online Supplementary Material, we show that under both the location shift model (M1) and the heteroscedastic model (M2), $\tau_n = k_0(\ln\ln n)/n$, for some constant $k_0 > 0$, is suitable for all five Weibull-type tail distributions in (E1) – (E5). Then, a reasonable choice of ψ_n is $\psi_n = k_1/n^{1+\nu}$ or $k_1 \ln n/n^{\nu+1}$, for some $k_1 > 0$ and $\nu > 0$, leading to $\kappa_n = \ln\psi_n/\ln\tau_n \to 1+\nu > 1$ as $n \to \infty$. This implies that any conditional quantile $q_Y(\psi_n|\mathbf{x})$ with order higher than $q_Y(\tau_n|\mathbf{x})$ can be estimated effectively by our extrapolation method, because the rate of $\psi_n = k_1/n^{1+\nu}$ or $k_1 \ln n/n^{\nu+1} \to 0$ as $n \to \infty$ can be arbitrarily fast, given a suitable ν.

5. Simulation study

In this section, we conduct a simulation study to assess the finite-sample performance of the proposed extreme quantile estimator. Consider the following data-generating process:

$$Y_i = 1 + X_{i1} + X_{i2} + X_{i3} + \frac{(X_{i1} + X_{i2})\,V_i}{2}, \quad i = 1, \ldots, n,$$

where $\{X_{ij}\}_{i=1}^n$ are independent and identically distributed (i.i.d.) random variables from the uniform distribution U(0, 1), for $j = 1, 2, 3$, and $\{V_i\}_{i=1}^n$ are generated from the following five Weibull-type distributions: $\mathrm{N}(0, 9)$, with $\theta = 0.5$; $\mathrm{W}(5, 1)$, with $\theta = 0.2$; $\mathrm{W}(1, 1)$, with $\theta = 1$; $\mathrm{MW}(2/3)$, with

$\theta = 1.5$; and MW(1/2), with $\theta = 2$. In each case, the true conditional quantile of Y is $q_Y(\psi_n|\mathbf{x}) = 1 + x_1 + x_2 + x_3 + (x_1 + x_2)\bar{F}_V^{-1}(\psi_n)/2$, for $\psi_n \in (0, 1)$ and $\mathbf{x} = (1, x_1, x_2, x_3)'$. We consider $n = 1000$ in the simulation study, and repeat the simulation 200 times for each case.

Our focus is the estimation of the extreme conditional quantiles $q_Y(\psi_n|\mathbf{x})$, where $\psi_n = 1/n^{1+\nu}$, with $\nu = 0.01$ (resulting in $\psi_n = 0.001$). For comparison, we consider the conventional QR estimator $\hat{q}_n(\psi_n|\mathbf{x}) = \mathbf{x}'\hat{\beta}(\psi_n)$, and three variations of the proposed extreme conditional quantile estimator, $\hat{q}^{\mathrm{P,c}}_{n,\mathrm{E}}(\psi_n|\mathbf{x})$, $\hat{q}^{\mathrm{P,l}}_{n,\mathrm{E}}(\psi_n|\mathbf{x})$, and $\hat{q}^{\mathrm{H}}_{n,\mathrm{E}}(\psi_n|\mathbf{x})$, based on the tail-coefficient estimators $\hat{\theta}^{\mathrm{c}}_{n,\mathrm{P}}(\bar{\mathbf{x}})$, $\hat{\theta}^{\mathrm{l}}_{n,\mathrm{P}}(\bar{\mathbf{x}})$, and $\hat{\theta}_{n,\mathrm{H}}(\bar{\mathbf{x}})$, respectively. Here, $\bar{\mathbf{x}} = (1, \bar{x}_1, \bar{x}_2, \bar{x}_3)'$ with $\bar{x}_j = R^{-1}\sum_{s=1}^{R} x_{sj}$, for $j = 1, 2, 3$, and $\{x_{sj}\}_{s=1}^{R}$ ($R = 100$) are drawn randomly from U(0, 1).

To examine the sensitivity of the proposed estimators to the choice of τ_n, we let $\tau_n = k_0(\ln\ln n)/n$, and plot the RMISE versus $k_0 \in [2, 30]$ in Figure 1 of the online Supplementary Material. Here, the RMISE is defined as the square root of the mean integrated squared error between a conditional quantile estimator and the truth $q_Y(\psi_n|\mathbf{x})$, integrated over $\mathbf{x}$ and across 200 simulations. Figure 1 yields the following observations. For the Gaussian, Weibull(5,1), and Weibull(1,1) distributions with small or modest tail-coefficients, the estimator $\hat{q}^{\mathrm{P,l}}_{n,\mathrm{E}}$ is more sensitive to the choice of k_0, and is generally more efficient than the conventional QR estimator for $k_0 \in [2, 10]$. However, the estimators $\hat{q}_{n,\mathrm{E}}^{\mathrm{P,c}}$ and $\hat{q}_{n,\mathrm{E}}^{\mathrm{H}}$ are more efficient than the QR estimator, in general, for $k_0 \in [2, 20]$. On the other hand, for the MW(2/3) and MW(1/2) distributions with larger tail-coefficients, the estimator $\hat{q}_{n,\mathrm{E}}^{\mathrm{P,l}}$ appears to be more efficient than $\hat{q}_{n,\mathrm{E}}^{\mathrm{P,c}}$ and $\hat{q}_{n,\mathrm{E}}^{\mathrm{H}}$, and all three are clearly more efficient than the QR estimator across $k_0 \in [2, 30]$.

The tuning parameter k_0 plays a similar role to the threshold value in the extreme value literature; that is, it balances the bias and the variance, and has to be properly chosen. Several methods exist for choosing the threshold-type tuning parameter; see Caeiro and Gomes (2016) for a review on this topic. In practice, we choose k_0 by adapting the procedure in Neves et al. (2015) based on path-stability. Specifically, in our simulation study, we regard the path of the tail-coefficient estimation as a function of k_0. Then, we choose the smallest value of k_0 within $[2, 30]$, starting from which, the estimation $\hat{\theta}$ becomes most stable.

Table 1 summarizes the RMISE of the conventional QR estimator and the three extrapolation estimators based on $\tau_n = k_0(\ln \ln n)/n$, with k_0 chosen by the path-stability procedure. The Hill-type estimator and the Pickand-type estimator with constant weights perform similarly, and both are clearly more efficient than the QR estimator across all five distributions considered. The Pickand-type estimator with linear weights performs best for the two MW distributions, which have larger tail-coefficients, but the method is less efficient than the other two extrapolation estimators for distributions with a tail-coefficient $\theta \leq 1$. These observations support the theoretical comparison in Section 4.2.

6. Analysis of birth weights

To illustrate the usefulness of the proposed methods, we study the effects of various behaviors of pregnant women on extremely high quantiles of birth weights of live infants born in the United States. It is well known that a low birth weight is associated with many health problems. On the other hand, a high birth weight can also have serious adverse effects on both maternal and child health. For example, a baby born with an excessively high birth weight may be at increased risk at birth of injuries, respiratory distress syndrome, low blood sugar, jaundice, and long-term health risks such as type-2 diabetes, childhood obesity, and metabolic syndrome; see, for instance, Aye et al. (2010) and Mohammadbeigi et al. (2013).

We use the June 1997 Detailed Natality Data published by the National Center for Health Statistics, which contains the birth weights of 31912 infants born to black mothers. We let the response Y be the birth weights in grams, and consider eight covariates: X_1 is a binary variable indicating whether the mother was married; X_2 indicates whether the infant is a boy; X_3 represents the mother's age (mean 26); $X_{4,1}$, $X_{4,2}$, and $X_{4,3}$ indicate whether the mother had no prenatal visit, visited for the first time in the second trimester, and visited for the first time in the third trimester, respectively; X_5 denotes the mother's education level (0 for less than high school, 1 for high school, 2 for some college, and 3 for college graduate); X_6 indicates whether the mother smoked during pregnancy; X_7 represents the average daily number of cigarettes per day the mother smoked; and X_8 denotes the mother's weight gain during pregnancy (mean 29 pounds). The same data set was also analyzed in Abreveya (2001), Koenker and Hallock (2001), and Chernozhukov and Fernández-Val (2011). However, the former two focused on analyzing typical birth weights in the range between 2000 and 4500 grams, and the latter examined extremely low birth weights in the range between 250 and 1500 grams. In contrast, we focus on the extremely high quantiles of birth weights, over 4500 grams.

Let $\mathbf{X} = (1, X_1, X_2, X_3, X_3^2, X_{4,1}, X_{4,2}, X_{4,3}, X_5, \ldots, X_8, X_8^2)^T$, where X_3, X_3^2 and X_8, X_8^2 are centered at zero. We consider the following linear quantile regression model: $q_Y(\tau|\mathbf{X}) = \mathbf{X}'\beta(\tau)$, $\tau \in (0, 1)$.

To examine whether the conditional distribution of Y has a Weibull-type tail, we follow the suggestion in Section 4.1 and plot $\ln(\hat{q}_n(\tau|\bar{\mathbf{x}}))$ against $\ln_{-2}(\tau)$ for $\tau \in \{0.01, 0.0095\ , \ldots, 0.001\}$ in Figure 2. The plot suggests that there is a strong linear relationship between $\ln(\hat{q}_n(\tau|\bar{\mathbf{x}}))$ and $\ln_{-2}(\tau)$. Hence, our proposed method is appropriate for analyzing the data. Similarly to the

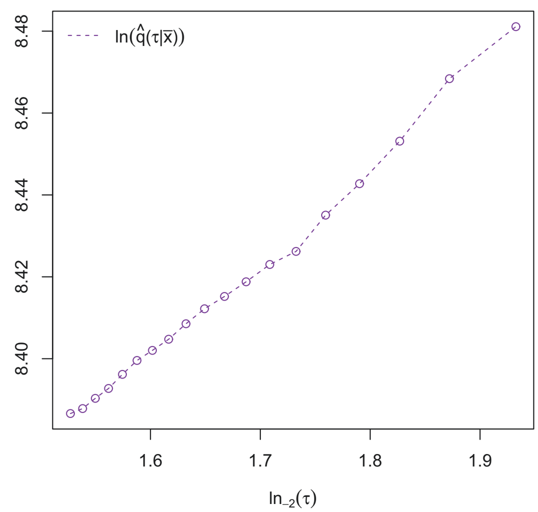

Figure 2: Diagnosis of the Weibull-type tail for the birth-weight data.

simulation study, we choose J and r by following the grid search method discussed in Section 4.2, and let $\tau_n = k_0(\ln\ln n)/n$. Figure 3 shows the path of the three tail-coefficient estimators against $k_0 \in [2, 100]$. Note that we exclude $k_0 = 1$, because this results in a small τ_n such that the tail-coefficient is estimated to be zero. Using the path-stability procedure in Neves et al. (2015), the adaptive k_0 is chosen as 45, 63, and 40 for

$\hat{\theta}^{\rm c}_{n,\rm P}(\bar{\mathbf{x}})$, $\hat{\theta}^{\rm l}_{n,\rm P}(\bar{\mathbf{x}})$, and $\hat{\theta}_{n,\rm H}(\bar{\mathbf{x}})$, respectively, and the corresponding estimates are $\hat{\theta}^{\rm c}_{n,\rm P}(\bar{\mathbf{x}}) = 0.225$, $\hat{\theta}^{\rm l}_{n,\rm P}(\bar{\mathbf{x}}) = 0.166$, and $\hat{\theta}_{n,\rm H}(\bar{\mathbf{x}}) = 0.247$, respectively. Figure 3 shows that the path of $\hat{\theta}^{\rm c}_{n,\rm P}(\bar{\mathbf{x}})$ is relatively more stable than those of $\hat{\theta}^{\rm l}_{n,\rm P}(\bar{\mathbf{x}})$ and $\hat{\theta}_{n,\rm H}(\bar{\mathbf{x}})$ when $k_0 \in [40, 100]$.

Figure 4 plots the estimated extremely high conditional quantiles of the birth weights of baby girls and boys born to black mothers, of the average profile, from the conventional QR and the proposed extrapolation estimators against the percentile level $100(1-\psi_n)$, where $\psi_n = k_1/n^{1.01}$ with $k_1 \in \{0.1, 0.2, \ldots, 0.9, 1, 2, \ldots, 50\}$, and from the three extrapolation estimators based on $\hat{\theta}^{\rm c}_{n,\rm P}(\bar{\mathbf{x}})$, $\hat{\theta}^{\rm l}_{n,\rm P}(\bar{\mathbf{x}})$, and $\hat{\theta}_{n,\rm H}(\bar{\mathbf{x}})$, denoted by EC, EL, and EH, respectively.

The following observations are derived from Figure 4. First, the estimates from the conventional QR method are not monotonically increasing with the quantile level, whereas such monotonicity is ensured by the extrapolation estimators. Second, for $100(1-\psi_n)$ ranging over $[99.8588, 99.9576]$, both the QR and the extrapolation estimators suggest that the quantiles of the birth weights of boys are higher than those of girls. However, for extremely high percentiles $100(1-\psi_n) > 99.9831$, the QR estimates suggest an opposite relationship, namely, that girls have higher birth weights than boys. This result is surprising, because we often found that male infants

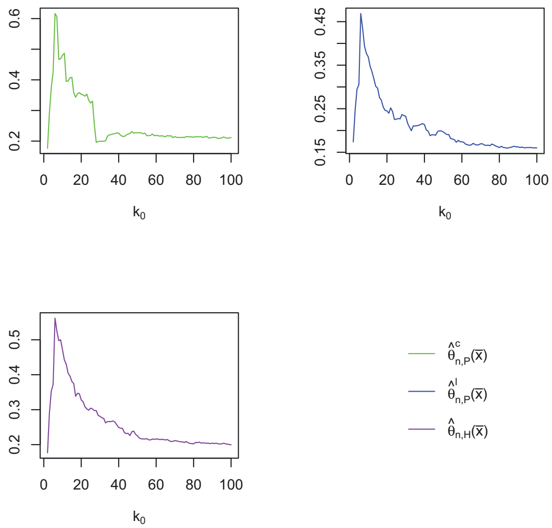

Figure 3: Three estimators of the Weibull tail-coefficient θ versus k_0 for the high birth weight.

are heavier than female infants, in general. Based on the QR, the 99.98th percentile of the birth weight of an infant girl born to an average mum is estimated to be 5269.218 grams, and the 99.99th percentile is estimated to be 5674.657 grams. Further investigation shows that these high estimates from the QR are mainly affected by one infant girl who has an extremely high birth weight of 6776 grams, and was born to a mother whose first prenatal visit was during the second trimester. In contrast, the proposed estimators are based on extrapolations from the $(1-\psi_n)$th quantile and, thus, are less susceptible to the extreme measurements of individual subjects.

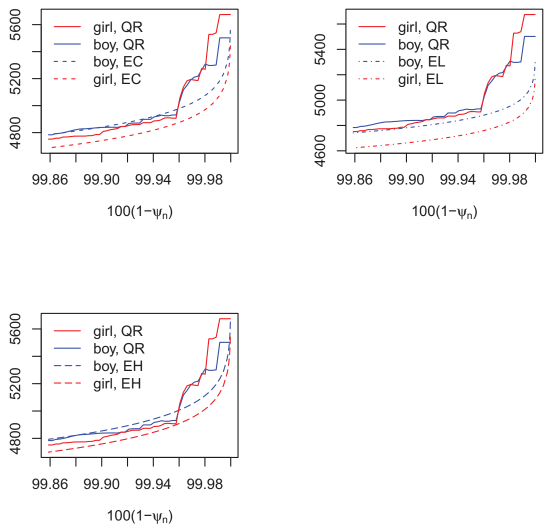

Figure 4: Estimation of the extremely high conditional quantile of the birth weights of baby girls and boys born to black mothers of the average profile, using the conventional QR and three extrapolation estimators.

Supplementary Material

The online Supplementary Material includes four sections. In Section S1, we provide seven lemmas that are needed to derive the asymptotic results of the proposed estimators. In Section S2, we provide two propositions that are used in Section 4.3. Technical proofs of all four theorems are presented in Section S3. In Section S4, we present Figure 1, which plots the RMISE of different estimators versus k_0 for the simulation study.

Acknowledgments

This research was partly supported by the National Natural Science Foundation of China grants No.11671338 and No.11690012, the National Science Foundation (NSF) grant DMS-1712760, the IR/D program from the NSF, and the OSR-2015-CRG4-2582 grant from KAUST. The authors thank the editor, the associate editor, and two referees for their constructive comments. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF.

References

Abreveya, J. (2001). The effects of demographics and maternal behavior on the distribution of birth outcomes. *Empirical Economics 26*, pp. 247–257.

Aye, S. S., Miller, V., Saxena, S. and Farhan, M. (2010). Management of large-for-gestational-age pregnancy in non-diabetic women. *The Obstetrician and Gynaecologist 12*, pp. 250–256.

Beirlant, J. and Teugels, J. L. (1992). Modeling large claims in non-life insurance. *Insurance: Mathematics and Economics 11*, pp. 17–29.

Berred, M. (1991). Record values and the estimation of the Weibull tail-coefficient. *Comptes Rendus de l'Académie des Sciences, Série I: Mathématique 312*, pp. 943–946.

Broniatowski, M. (1993). On the estimation of the Weibull tail coefficient. *Journal of Statistical*

Planning and Inference 35, pp. 349–365.

Caeiro, F. and Gomes, M. I. (2016). *Threshold Selection in Extreme Value Analysis* in Dey, D. K. and Yan, J. (Eds.), *Extreme Value Modeling and Risk Analysis: Methods and Applications* (pp. 69–86). Boca Raton: CRC Press.

Chernozhukov, V. (2005). Extremal quantile regression. *The Annals of Statistics 33*, pp. 806–839.

Chernozhukov, V. and Fernández-Val, I. (2011). Inference for extremal conditional quantile models, with an application to market and birthweight risks. *The Review of Economic Studies 78*, pp. 559–589.

Daouia, A., Gardes, L. and Girard, S. (2013). On kernel smoothing for extremal quantile regression. *Bernoulli 19*, pp. 2557–2589.

Daouia, A., Gardes, L., Girard, S. and Lekina, A. (2011). Kernel estimators of extreme level curves. *TEST 20*, pp. 311–333.

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

de Wet, T., Goegebeur, Y., Guillou, A. and Osmann, M. (2016). Kernel regression with Weibull-type tails. *Annals of the Institute of Statistical Mathematics 68*, pp. 1135–1162.

Diebolt, J., Gardes, L., Girard, S. and Guillou, A. (2008). Bias-reduced estimators of the Weibull tail-coefficient. *TEST 17*, pp. 311–331.

Gardes, L. and Girard, S. (2005). Estimating extreme quantiles of Weibull tail distributions. *Communications in Statistics - Theory and Methods 34*, pp. 1065–1080.

Gardes, L. and Girard, S. (2008). Estimation of the Weibull tail-coefficient with linear combination of upper order statistics. *Journal of Statistical Planning and Inference 138*, pp. 1416–1427.

Gardes, L. and Girard, S. (2016). On the estimation of the functional Weibull tail-coefficient. *Journal of Multivariate Analysis 146*, pp. 29–45.

Girard, S. (2004). A Hill type estimator of the Weibull tail-coefficient. *Communications in Statistics - Theory and Methods 33*, pp. 205–234.

Goegebeur, Y., Beirlant, J. and de Wet, T. (2010). Generalized kernel estimators for the Weibull-tail coefficient. *Communications in Statistics - Theory and Methods 39*, pp. 3695–3716.

Goegebeur, Y. and Guillou, A. (2011). A weighted mean excess function approach to the estimation of Weibull-type tails. *TEST 20*, pp. 138–162.

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

Koenker, R. (2005). *Quantile Regression.* Cambridge: Cambridge University Press.

Koenker, R. and Hallock, K. (2001). Quantile regression: an introduction. *Journal of Economic Perspectives 15*, pp. 143–156.

Mohammadbeigi, A., Farhadifar F., Soufi zadeh, N., Mohammadsalehi, N., Rezaiee, M. and

Aghaei, M. (2013). Fetal Macrosomia: risk Factors, Maternal, and Perinatal Outcome. *Annals of Medical and Health Sciences Research 3*, pp. 546–550.

Neves, M. M., Gomes, M. I., Figueiredo, F. and Gomes, D. P. (2015). Modeling extreme events: sample fraction adaptive choice in parameter estimation. *Journal of Statistical Theory and Practice 9*, pp. 184–199.

Resnick, S. I. (1987). *Extreme Values, Regular Variation and Point Processes.* New York: Springer-Verlag.

Wang, H., Li, D. and He, X. (2012). Estimation of high conditional quantiles for heavy-tailed distributions. *Journal of the American Statistical Association 107*, pp. 1453-1464.

Wang, H. and Tsai, C. L. (2009). Tail Index Regression. *Journal of the American Statistical Association 104*, pp. 1233–1240.

Department of Statistics, Jinan University, Guangzhou, China

E-mail: (198578hfy@163.com)

Department of Statistics, George Washington University, Washington DC, USA

E-mail: (judywang@email.gwu.edu)

Department of Mathematics, Hong Kong Baptist University, Hong Kong

E-mail: (tongt@hkbu.edu.hk)

Table 1: The root mean integrated squared errors of different estimators of $q_Y(\psi_n|\mathbf{x})$, with $\psi_n = n^{-1.01}$ and $n = 1000$. Values in parentheses are the standard errors. $\hat{q}_n$ is the conventional quantile regression estimator, and $\hat{q}_{n,\mathrm{E}}^{\mathrm{P,\,c}}$, $\hat{q}_{n,\mathrm{E}}^{\mathrm{P,\,l}}$, and $\hat{q}_{n,\mathrm{E}}^{\mathrm{H}}$ are the extrapolation estimators based on the Pickand-type tail-coefficient estimators with constant and linear weights, and the Hill-type tail-coefficient estimator, respectively. For the extrapolation estimators, $\tau_n = k_0(\ln\ln n)/n$, where k_0 is chosen using the path-stability procedure.

Distribution	$\hat{q}_{n,\mathrm{E}}^{\mathrm{P,\,c}}$	$\hat{q}_{n,\mathrm{E}}^{\mathrm{P,\,l}}$	$\hat{q}_{n,\mathrm{E}}^{\mathrm{H}}$	$\hat{q}_n$
$\mathrm{N}(0,9)$	0.6143	0.8067	0.6260	0.7753
	(0.0194)	(0.0189)	(0.0184)	(0.0233)
$\mathrm{W}(5,1)$	0.0344	0.0423	0.0342	0.0392
	(0.0009)	(0.0009)	(0.0008)	(0.0012)
$\mathrm{W}(1,1)$	0.6892	0.8264	0.6921	0.8712
	(0.0168)	(0.0178)	(0.0172)	(0.0329)
$\mathrm{MW}(2/3)$	11.264	8.394	11.150	17.041
	(0.5043)	(0.4223)	(0.4736)	(0.8208)
$\mathrm{MW}(1/2)$	54.730	39.437	52.879	84.043
	(1.6219)	(1.4245)	(1.4677)	(3.9583)