Statistics and Its Interface Volume 9 (2016) 131–139

Transformed linear quantile regression with censored survival data

Rui Miao, Liuquan Sun*, and Guo-Liang Tian

Quantile regression provides a flexible method for analyzing survival data, and attracts considerable interest in survival analysis. In this article, we propose a new inference procedure for a class of power-transformed linear quantile regression models with survival data subject to conditionally independent censoring, and present a two-stage algorithm that is computationally simple and easy to implement. Consistency and asymptotic normality of the resulting estimators are established, and a simple resampling-based inference procedure is developed for variance estimation. The finite-sample behavior of the proposed methods is examined through extensive simulation studies. An application to a real data example from a health maintenance organization is provided.

AMS 2000 subject classifications: Primary 62N01, 62N02; secondary 62G20.
Keywords and phrases: Censored survival data, Box–Cox transformation, Martingale, Quantile regression, Resampling.

1. INTRODUCTION

Quantile regression offers a valuable complement to the Cox proportional hazards model (Cox, 1972) and the accelerated failure time model (Buckley and James, 1979) in survival analysis. It allows the covariate effects to vary at different tails of the survival time distribution. Such important heterogeneity in the population may be neglected by using the Cox model and the accelerated failure time model. In addition, it provides straightforward interpretation on the survival time, and allows covariate effects to vary across the location of the survival time. Recently, quantile regression has attracted considerable interest in survival analysis (Ying et al., 1995; Yang, 1999; Koenker and Geling, 2001; Portnoy, 2003; Peng and Huang, 2008; Yin et al., 2008; Wang and Wang, 2009; Qian and Peng, 2010).

For survival data, Powell (1984, 1986) first studied censored quantile regression for fixed censoring, where the censoring time is always observed (see also Tang et al., 2012). Ying et al. (1995) developed a semiparametric estimation procedure for a censored median regression model, and Honore et al. (2002) provided parallel extension of Powell's approach under random censoring, where the censoring time is assumed to be independent of the survival time and covariates. Under the usual independent censoring assumption that the censoring time is independent of the survival time conditional on covariates, Portnoy (2003) proposed censored regression quantiles and developed a recursively reweighted estimation procedure (see also Neocleous et al., 2006); Peng and Huang (2008) suggested a novel quantile regression method by utilizing the martingale feature associated with censored data; Wang and Wang (2009) presented a locally weighted censored quantile regression approach by adopting the redistribution-of-mass idea and employing a local reweighting scheme (see also Wang et al., 2013); Huang (2010) proposed an estimation procedure for censored quantile regression based on estimating integral equations; Qian and Peng (2010) developed a censored quantile regression method tailored to the partially functional effect setting with a mixture of varying and constant effects.

Transformed quantile regression models are robust and flexible, and can accommodate a wide variety of models including the accelerated failure time model as a special case. A class of power-transformed linear quantile regression models have been proposed for complete data without any censoring (e.g. Powell, 1991; Buchinsky, 1995; Machado and Mata, 2000; Mu and He, 2007). But when data are subject to random censoring, inference for quantile regression is much more involved and challenging, especially for transformed quantile regression models. Recently, Yin et al. (2008) proposed a power-transformed linear quantile regression model for randomly censored survival data. However, their methods require the unconditional independence between the survival time and the censoring time. It is well known that when the unconditional independence assumption is violated, the methods relying on such an assumption may yield biased results. Recently, Leng and Tong (2014) considered a class of power-transformed quantile regression models at a particular quantile based on Wang and Wang (2009)'s method. Note that a locally weighted Kaplan-Meier estimator may have larger bias and slower convergence rate due to the curse of dimensionality. Thus, when the dimensionality of the covariates is high, Wang and Wang (2009)'s and Leng and Tong (2014)'s estimates can have large biases and sampling standard errors (Tong, 2014).

*Corresponding author.

In this article, we propose a new inference procedure for a class of power-transformed linear quantile regression models for survival data subject to conditionally independent censoring. The estimation procedure consists of two sequential steps. First, for a given transformation parameter, we can easily obtain the estimates of the regression coefficients by utilizing the martingale-based framework (Peng and Huang, 2008). Second, we can estimate the transformation parameter based on a cumsum process of residuals (Mu and He, 2007). The resulting estimators are uniformly consistent and asymptotic normal. Compared to existing procedures, the proposed method enjoys several distinctive advantages. First, our proposed procedure does not require the local Kaplan-Meier estimator of the conditional distribution function of the survival time as in Wang and Wang (2009) and Leng and Tong (2014), and thus there is no curse of dimensionality. Second, the proposed procedure does not require iteration as in Yin et al. (2008). Third, the proposed algorithm is computationally simple and easy to implement based on existing software.

The rest of the paper is organized as follows. In Section 2, we introduce relevant notation, formulate the model, and propose the estimation procedure for the model parameters. In Section 3, asymptotic properties of the proposed estimators are established. Some numerical results from simulation studies for evaluating the proposed methods are reported in Section 4. Section 5 applies the proposed method to the health maintenance organization (HMO) data, and some concluding remarks are given in Section 6. The technical proofs are relegated to the Supplementary Material (http://www.intlpress.com/SII/p/2016/9-2/SII-9-2-MIAO-supplement.pdf).

2. MODEL AND ESTIMATION PROCEDURES

Let T be the survival time, C be the censoring time, and $\mathbf{Z}$ be the $p \times 1$ vector of covariates. Define $X = T \wedge C$ and $\delta = I(T \le C)$, where $\wedge$ is the minimum operator and $I(\cdot)$ is the indicator function. The observed data consist of n independent and identically distributed replicates of $(X, \delta, \mathbf{Z})$, denoted by $\{(X_i, \delta_i, \mathbf{Z}_i), i = 1, ..., n\}$. It is assumed that C is independent of T conditional on $\mathbf{Z}$.

Let $H_\gamma(\cdot)$ be a family of monotonic transformation indexed by a parameter γ, which includes the log-transformation and the Box–Cox transformation as special cases. Given the covariate $\mathbf{Z}$ and $\tau \in (0, 1)$, the τth conditional quantile of a random variable, say Y, is defined as $Q_\tau(Y|\mathbf{Z}) = \inf\{t : \Pr(Y \le t|\mathbf{Z}) \ge \tau\}$. The proposed power-transformed linear quantile regression models take the form

$$Q_\tau(H_\gamma(T)|\mathbf{Z}) = \mathbf{Z}^T \boldsymbol{\beta}(\tau), \tag{1}$$

where $\boldsymbol{\beta}(\tau)$ is a vector of unknown regression coefficients, which represents the effects of covariates on the τth quantile of $H_\gamma(T)$ and may change with τ. The equivariance property of quantiles to monotone transformations implies that

$$Q_\tau(T|\mathbf{Z}) = H_\gamma^{-1}\{\mathbf{Z}^T \boldsymbol{\beta}(\tau)\},$$

where $H_\gamma^{-1}(\cdot)$ is the inverse transformation of $H_\gamma(\cdot)$.

Define $F_T(t|\mathbf{Z}) = \Pr(T \le t|\mathbf{Z})$, $\Lambda_T(t|\mathbf{Z}) = -\log\{1 - \Pr(T \le t|\mathbf{Z})\}$, $N_i(t) = I(X_i \le t, \delta_i = 1)$, and $M_i(t) = N_i(t) - \Lambda_T(t \wedge X_i|\mathbf{Z}_i)$, $i = 1, ..., n$. Let $\boldsymbol{\beta}_0(\tau)$ and γ_0 be the true values of $\boldsymbol{\beta}(\tau)$ and γ, respectively. Since $M_i(t)$ is the martingale process associated with the counting process $N_i(t)$ (Fleming and Harrington, 1991), we have

$$\begin{aligned} E\Big\{n^{-1}\sum_{i=1}^n \mathbf{Z}_i \Big[N_i\left(H_{\gamma_0}^{-1}\{\mathbf{Z}_i^T\boldsymbol{\beta}_0(\tau)\}\right) \\ - \Lambda_T(H_{\gamma_0}^{-1}\{\mathbf{Z}_i^T\boldsymbol{\beta}_0(\tau)\} \wedge X_i \,|\, \mathbf{Z}_i)\Big]\Big\} = \mathbf{0}. \end{aligned} \tag{2}$$

Note that $F_T(H_{\gamma_0}^{-1}\{\mathbf{Z}_i^T\boldsymbol{\beta}_0(\tau)\} \,|\, \mathbf{Z}_i) = \tau$. Then

$$\begin{aligned} &\Lambda_T(H_{\gamma_0}^{-1}\{\mathbf{Z}_i^T\boldsymbol{\beta}_0(\tau)\} \wedge X_i \,|\, \mathbf{Z}_i) \\ &= \int_0^\tau I[\, X_i \ge H_{\gamma_0}^{-1}\{\mathbf{Z}_i^T\boldsymbol{\beta}_0(\mu)\}]\, \mathrm{d}G(\mu), \end{aligned}$$

where $G(\mu) = -\log(1-\mu)$ for $0 \le \mu < 1$. In view of (2), for a given γ, we specify the following estimating equation for $\boldsymbol{\beta}_0(\tau)$:

$$\mathbf{S}_n\{\boldsymbol{\beta}(\tau);\gamma\} = \mathbf{0}, \tag{3}$$

where

$$\begin{aligned} \mathbf{S}_n\{\boldsymbol{\beta}(\tau);\gamma\} = & \frac{1}{n}\sum_{i=1}^n \mathbf{Z}_i\Big\{N_i(H_\gamma^{-1}\{\mathbf{Z}_i^T\boldsymbol{\beta}(\tau)\}) \\ & - \int_0^\tau I[\, X_i \ge H_\gamma^{-1}\{\mathbf{Z}_i^T\boldsymbol{\beta}(\mu)\}]\, \mathrm{d}G(\mu)\Big\}. \end{aligned}$$

Because of the stochastic integration representation of $\mathbf{S}_n\{\boldsymbol{\beta}(\tau);\gamma\}$, following Peng and Huang (2008), we define the estimator $\hat{\boldsymbol{\beta}}(\tau;\gamma)$ as a right-continuous step function, which is the solution to the above estimating equation (3) and only jumps on a grid $S_{L(n)} = \{0 = \tau_0 < \tau_1 < \ldots < \tau_{L(n)} = \tau_U < 1\}$, where τ_U is a constant subject to certain identifiability constraints provided in the next section. In this paper, we denote $L \equiv L(n)$ for simplicity. Note that $Q_0(T|\mathbf{Z}) = 0$ implies $H_\gamma^{-1}\{\mathbf{Z}^T\boldsymbol{\beta}_0(0)\} = 0$. Thus, we set $H_\gamma^{-1}\{\mathbf{Z}_i^T\hat{\boldsymbol{\beta}}(0;\gamma)\} = 0$ $(i = 1, ..., n)$ for a given γ. By using the grid method, $\hat{\boldsymbol{\beta}}(\tau_j;\gamma)$ $(j = 1, ..., L)$ can be obtained sequentially by solving the following monotone estimating equation for $\boldsymbol{\beta}(\tau_j)$:

$$\begin{aligned} \frac{1}{n}\sum_{i=1}^n \mathbf{Z}_i\Big\{&N_i(\, H_\gamma^{-1}\{\mathbf{Z}_i^T\boldsymbol{\beta}(\tau_j)\}) \\ &- \sum_{k=0}^{j-1} I[\, X_i \ge H_\gamma^{-1}\{\mathbf{Z}_i^T\hat{\boldsymbol{\beta}}(\tau_k;\gamma)\}] \\ &\times \{G(\tau_{k+1}) - G(\tau_k)\}\Big\} = \mathbf{0}. \end{aligned} \tag{4}$$

As in Peng and Huang (2008), some simple algebra manipulations show that solving equation (4) is equivalent to finding the minimizer of the following L_1-type convex objective function:

$$
\begin{aligned}
&\ell_j(h,\gamma)\\
&= \sum_{i=1}^{n} \left|\delta_i\, H_\gamma\,(X_i) \,-\, \delta_i\, h^T \mathbf{Z}_i\right| + \left|R^* \,-\, h^T \sum_{i=1}^{n} (-\delta_i \mathbf{Z}_i)\right|\\
&+ \left|R^* \,-\, h^T \sum_{i=1}^{n} \Big(2\mathbf{Z}_i \sum_{k=0}^{j-1} I[\, X_i \,\geq\, H_\gamma^{-1}\, \{\mathbf{Z}_i^T \hat{\boldsymbol{\beta}}(\tau_k;\gamma)\}]\right.\\
&\left.\times \{G(\,\tau_{k+1}\,) - G(\,\tau_k\,)\}\Big)\right|,
\end{aligned}
$$

where R^* is an extremely large positive number selected to bound $|\, h^T \sum_{i=1}^{n} (-\delta_i \mathbf{Z}_i)|$ and

$$
\begin{aligned}
&|\, h^T \sum_{i=1}^{n} \big(2\mathbf{Z}_i \sum_{k=0}^{j-1} I[\, X_i \,\geq\, H_\gamma^{-1}\, \{\mathbf{Z}_i^T \hat{\boldsymbol{\beta}}(\tau_k;\gamma)\}]\\
&\times \{G(\,\tau_{k+1}\,) - G(\,\tau_k\,)\}\big)|
\end{aligned}
$$

from above for all h's in the compact parameter space for $\boldsymbol{\beta}_0(\tau_j)$. In fact, the built-in rq function in R package quantreg can be employed to find the minimizer of $\ell_j(h)$ (e.g., Peng and Huang, 2008).

To estimate γ, following Mu and He (2007) and Yin et al. (2008), we define a discrepancy measure based on the cumsum process of residuals, which can distinguish the right transformation from a wrong alternative. Let $\hat{\gamma}$ be the minimizer of

$$
(5)\qquad R_n(\gamma) = \frac{1}{n} \sum_{i=1}^{n} \int_{\nu}^{\tau_U} D_n^2(\mathbf{Z}_i, \tau, \gamma)\, \mathrm{d}\tau \quad \text{for } \gamma \in \Upsilon,
$$

where $0 < \nu < \tau_U$, Υ denotes the parameter space for γ, and

$$
\begin{aligned}
D_n(\boldsymbol{z},\tau,\gamma) \;&=\; \frac{1}{n} \sum_{i=1}^{n} I(\mathbf{Z}_i \leq \boldsymbol{z}) \Big\{ N_i(H_\gamma^{-1}\, \{\mathbf{Z}_i^T \hat{\boldsymbol{\beta}}(\tau;\gamma)\})\\
&- \int_0^{\tau} I[\, X_i \,\geq\, H_\gamma^{-1}\, \{\mathbf{Z}_i^T \hat{\boldsymbol{\beta}}(\mu;\gamma)\}]\, \mathrm{d}G(\mu) \Big\}.
\end{aligned}
$$

Here $I(\mathbf{Z}_i \leq \boldsymbol{z})$ means that each of the components of $\mathbf{Z}_i$ is not larger than the corresponding component of $\boldsymbol{z}$. Although $R_n(\gamma)$ is not differentiable with respect to γ, the built-in R function optimize can be used to find $\hat{\gamma} = \arg\min_\gamma R_n(\gamma)$, since $R_n(\gamma)$ is a function of a single parameter γ. When $\hat{\gamma}$ is available, $\boldsymbol{\beta}_0(\tau)$ can be estimated by $\hat{\boldsymbol{\beta}}(\tau) \equiv \hat{\boldsymbol{\beta}}(\tau;\hat{\gamma})$.

For a given search interval of γ such as $[a,b]$, the built-in R function optimize determines grid points by itself, such as $x_1, ..., x_K$. Then for given τ_j $(j = 1, \ldots, L)$, the proposed estimation procedure can be summarized as follows, which is robust and effective in the simulation studies in Section 4.

Step 1. For each grid point x_k $(k = 1, .., K)$, find $\hat{\boldsymbol{\beta}}(\tau_0; x_k)$ satisfies $H_{x_k}^{-1}\{\hat{\boldsymbol{\beta}}(\tau_0; x_k)\} = 0$.

Step 2. For each x_k $(k = 1, .., K)$ and $j = 1$, minimize $\ell_j(h, x_k)$ using the rq function to obtain $\hat{\boldsymbol{\beta}}(\tau_j; x_k)$.

Step 3. For each x_k $(k = 1, .., K)$, set $j = j + 1$, and go to Step 2 to obtain $\hat{\boldsymbol{\beta}}(\tau_j; x_k)$ sequentially till $j = L$.

Step 4. Put $\hat{\boldsymbol{\beta}}(\tau_j; x_k)$ $(j = 1, \ldots, L)$ into (5), we can obtain $R_n(x_k)$ $(k = 1, .., K)$, and find the optimal $\hat{\gamma} = \arg\min_{x_k, k=1,\ldots,K} R_n(x_k)$. Finally, we use the rq function to obtain $\hat{\boldsymbol{\beta}}(\tau) \equiv \hat{\boldsymbol{\beta}}(\tau; \hat{\gamma})$.

3. ASYMPTOTIC PROPERTIES

Define $F(t|\mathbf{Z}) = \Pr(X \leq t|\mathbf{Z})$, $\tilde{F}(t|\mathbf{Z}) = P(X \leq t, \delta = 1|\mathbf{Z})$, $f(t|\mathbf{Z}) = \mathrm{d}F(t|\mathbf{Z})/\mathrm{d}t$, and $\tilde{f}(t|\mathbf{Z}) = \mathrm{d}\tilde{F}(t|\mathbf{Z})/\mathrm{d}t$. In order to study the asymptotic properties of the proposed estimators, we need the following regularity conditions:

(R1) The covariate space is bounded; that is, $\sup_i \|\mathbf{Z}_i\| < \infty$.

(R2) The transformation $H_\gamma(t)$ is strictly increasing with respect to t and twice-continuously differentiable in $t \in (0, \infty)$ and $\gamma \in \Omega_{\gamma_0}$, where Ω_{γ_0} is a neighborhood of γ_0.

(R3) If $H_{\gamma_1}^{-1}\{\mathbf{Z}^T \boldsymbol{\beta}_1\} = H_{\gamma_2}^{-1}\{\mathbf{Z}^T \boldsymbol{\beta}_2\}$, then $\gamma_1 = \gamma_2$ and $\boldsymbol{\beta}_1 = \boldsymbol{\beta}_2$.

(R4) (a) Each component of $E[ZN(H_{\gamma_0}^{-1}\{\mathbf{Z}^T \boldsymbol{\beta}_0(\tau)\})]$ is a Lipschitz function of τ; (b) $f(t|\mathbf{Z})$ and $\tilde{f}(t|\mathbf{Z})$ are continuous differentiable in t and $\mathbf{Z}$.

(R5) (a) $\tilde{f}\{H_\gamma^{-1}(\mathbf{Z}^T \boldsymbol{\beta})|\mathbf{Z}\} > 0$ for any $\boldsymbol{\beta} \in \Omega_{\beta_0}$ and $\gamma \in \Omega_{\gamma_0}$; (b) Each component of

$$
\begin{aligned}
&E[\mathbf{Z}^{\otimes 2} f\{H_\gamma^{-1}(\mathbf{Z}^T\boldsymbol{\beta})|\mathbf{Z}\} \frac{\partial H_\gamma^{-1}(t)}{\partial t}|_{t=Z^T\beta}]\\
&\times (E[\mathbf{Z}^{\otimes 2} \tilde{f}\{H_\gamma^{-1}(\mathbf{Z}^T\boldsymbol{\beta})|\mathbf{Z}\} \frac{\partial H_\gamma^{-1}(t)}{\partial t}|_{t=Z^T\beta}])^{-1}
\end{aligned}
$$

is uniformly bounded in $\boldsymbol{\beta} \in \Omega_{\beta_0}$ and $\gamma \in \Omega_{\gamma_0}$, where Ω_{β_0} is a neighborhood containing $\{\boldsymbol{\beta}_0(\tau) : \tau \in (0, \tau_U]\}$.

(R6) $\inf_{\tau \in [\nu, \tau_U]} \mathrm{eigmin} E[\mathbf{Z}^{\otimes 2} \tilde{f}\{H_{\gamma_0}^{-1}(\mathbf{Z}^T \boldsymbol{\beta}_0(\tau))|\mathbf{Z}\} \frac{\partial H_{\gamma_0}^{-1}(t)}{\partial t}] > 0$ for any $\nu \in (0, \tau_U)$, where $t = \mathbf{Z}^T \beta_0(\tau)$ and $\mathrm{eigmin}(\cdot)$ denotes the minimum eigenvalue of a matrix.

Conditions (R1), (R4), (R5) and (R6) are standard for quantile regression methods in analyzing failure time data, which are analogous to those in Peng and Huang (2008). Conditions (R2) and (R3) ensure the unique parameterization of the transformation, and the identifiability of the transformation and regression parameters, which are the same as those in Yin et al. (2008). For example, when H_γ is the Box–Cox transformation and $\mathbf{Z}$ includes one continuous covariate with a nonzero effect, Condition (R3) can be replaced by the linear independence of $\mathbf{Z}$ (see Yin et al., 2008). The asymptotic properties of $\hat{\gamma}$ and $\hat{\boldsymbol{\beta}}(\tau)$ are shown in the following theorems, whose proofs are given in the Supplementary Material.

Theorem 1. *Assume that conditions (R1)–(R6) hold, and* $\lim_{n\to\infty} \|S_L\| = 0$. *Then*

$$\hat{\gamma} \xrightarrow{\mathrm{P}} \gamma_0 \quad and \quad \sup_{\tau \in [\nu, \tau_U]} \|\hat{\boldsymbol{\beta}}(\tau) - \boldsymbol{\beta}_0(\tau)\| \xrightarrow{\mathrm{P}} 0,$$

where $0 < \nu < \tau_U$*.*

Theorem 2. *Assume that conditions (R1)–(R6) hold, and* $\lim_{n\to\infty} n^{1/2}\|S_L\| = 0$*. Then* $n^{1/2}(\hat{\gamma} - \gamma_0)$ *is asymptotically normal with mean zero and* $n^{1/2}\{\hat{\boldsymbol{\beta}}(\tau) - \boldsymbol{\beta}_0(\tau)\}$ *converges weakly to a zero-mean Gaussian process for* $\tau \in [\nu, \tau_U]$*, where* $0 < \nu < \tau_U$*.*

Note that the limit distributions of $n^{1/2}(\hat{\gamma} - \gamma_0)$ and $n^{1/2}\{\hat{\boldsymbol{\beta}}(\tau) - \boldsymbol{\beta}_0(\tau)\}$ involve unknown density functions, the nonparametric estimation of which may be unstable for finite samples. In order to obtain stable variance estimates of $\hat{\gamma}$ and $\hat{\boldsymbol{\beta}}(\tau)$, we adopt a resampling technique following Jin et al. (2001), Cai et al. (2005) and Peng and Huang (2008). Specifically, let $\{\zeta_1, ..., \zeta_n\}$ be independent and identically distributed nonnegative random variables following a known distribution with mean 1 and variance 1, such as the standard exponential distribution. Then using the estimation procedure in Section 2, we obtain $\hat{\gamma}^*$ and $\hat{\boldsymbol{\beta}}^*(\tau)$, with $\ell_j(h,\gamma)$ and $R_n(\gamma)$ replaced by $\ell_j^*(h,\gamma)$ and $R_n^*(\gamma)$, respectively, where for $j = 1, ..., L$,

$$\begin{aligned}
&\ell_j^*(h,\gamma) \\
&= \sum_{i=1}^{n} \left|\zeta_i\delta_i\, H_\gamma\,(X_i) \;-\; \zeta_i\delta_i\, h^T\mathbf{Z}_i\,\right| \\
&+ \left|R^* \;-\; h^T \sum_{i=1}^{n} (-\zeta_i\delta_i\mathbf{Z}_i)\right| \\
&+ \left|R^* \;-\; h^T \sum_{i=1}^{n} \left(2\zeta_i\mathbf{Z}_i \sum_{k=0}^{j-1} I[X_i \ge H_\gamma^{-1}\{\mathbf{Z}_i^T\hat{\boldsymbol{\beta}}^*(\tau_k;\gamma)\}]\right.\right. \\
&\left.\left. \times \{G(\,\tau_{k+1}\,) \;-\; G(\,\tau_k\,)\}\right)\right|, \\
&R_n^*(\gamma) = \frac{1}{n}\sum_{i=1}^{n}\int_{\nu}^{\tau_U} \{D_n^*(\mathbf{Z}_i,\tau,\gamma)\}^2\,\mathrm{d}\tau\,,
\end{aligned}$$

and

$$\begin{aligned}
D_n^*(\boldsymbol{z},\tau,\gamma) = \frac{1}{n}\sum_{i=1}^{n} I(\mathbf{Z}_i \le \boldsymbol{z})\Big\{&\zeta_i N_i(H_\gamma^{-1}\,\{\mathbf{Z}_i^T\hat{\boldsymbol{\beta}}^*(\tau;\gamma)\}) \\
&- \int_0^{\tau} \zeta_i I[\,X_i \;\ge\; H_\gamma^{-1}\,\{\mathbf{Z}_i^T\hat{\boldsymbol{\beta}}^*(\mu;\gamma)\}]\mathrm{d}G(\mu)\Big\}.
\end{aligned}$$

In the Supplementary Material, we show that the asymptotic distributions of $n^{1/2}(\hat{\gamma}-\gamma_0)$ and $n^{1/2}\{\hat{\boldsymbol{\beta}}(\tau)-\boldsymbol{\beta}_0(\tau)\}$ can be approximated by the conditional distribution $n^{1/2}(\gamma^*-\hat{\gamma})$ and $n^{1/2}\{\hat{\boldsymbol{\beta}}^*(\tau)-\hat{\boldsymbol{\beta}}(\tau)\}$ given the observed data. To estimate the variances of $\hat{\gamma}$ and $\hat{\boldsymbol{\beta}}(\tau)$, we obtain a large number of resampling estimators, say $\hat{\gamma}_k^*$ and $\hat{\boldsymbol{\beta}}_k^*(\tau)$ $(k = 1, ..., K)$, by repeatedly generating the random samples $\{\zeta_1, ..., \zeta_n\}$ with the data at their observed values. Thus, for a fixed $\tau \in [\nu, \tau_U]$, the variances of $\hat{\gamma}$ and $\hat{\boldsymbol{\beta}}(\tau)$ can be approximated by the sample variances of $\hat{\gamma}_k^*$ and $\hat{\boldsymbol{\beta}}_k^*(\tau)$ $(k = 1, ..., K)$, respectively.

4. SIMULATION STUDIES

In this section, some simulation studies are performed to evaluate the finite sample property of the proposed estimators. In these studies, the covariate vector $\mathbf{Z} = (Z_1, Z_2)^T$ was generated as $Z_1 \sim U(0,1)$ and $Z_2 \sim \text{Bernoulli}(0.5)$. We considered the Box–Cox transformation linear quantile regression model:

$$H_\gamma(T) = \frac{T^\gamma - 1}{\gamma} = \beta_1 Z_1 + \beta_2 Z_2 + \varepsilon, \tag{6}$$

where $\beta_1 = 0.5$, $\beta_2 = 1$, and $\gamma = 0$, 0.5 or 1. The error ε was simulated from $N(0, 0.25^2)$ (i.e., normal error). Then model (1) held with $\mathbf{Z} = (1, Z_1, Z_2)^T$ and $\boldsymbol{\beta}_0(\tau) = (Q_\varepsilon(\tau), \beta_1, \beta_2)^T$. Here we focus on the case of $\tau = 0.5$. The censoring time C was generated from a uniform distribution $U(0.1Z_2, V)$, with V varying to yield censoring rates of 0%, 20% and 40%, respectively. The results presented in Tables 1–4 are based on 1,000 replications with sample size $n = 200$. To obtain the standard errors of the parameter estimates, we set $K = 250$ in the resampling method with $\{\zeta_1, ..., \zeta_n\}$ generated from the standard exponential distribution. We adopted an equally spaced grid with $\|S_L\| = 0.01$. The built-in R function optimize was employed to find $\hat{\gamma} = \arg\min_\gamma R_n(\gamma)$ in the interval $[\gamma - 0.5,\ \gamma + 0.5]$, and the built-in rq function was used to obtain $\hat{\boldsymbol{\beta}}(\tau_j;\gamma)$.

Table 1 presents the simulation results on the estimates of β_1 and β_2 when γ is taken as unknown. The table includes the biases (Bias) given by the sampling mean of the estimate minus the true value, the sample standard deviation of the estimate (SD), the average of the estimated standard error (SE) based on the resampling method, and the coverage probability (CP) of the 95% confidence interval based on a normal approximation. It can be seen from the table that the proposed estimation procedures perform well for the situations considered here. Specifically, the proposed estimators are practically unbiased, and the estimated standard error based on the resampling method is close to the empirical standard error. Also the coverage probabilities of the 95% confidence intervals are reasonable.

For comparison, we conducted simulation studies using the same setup as in Table 1, when γ is taken as a known parameter, which is termed as conditional inference as in Mu and He (2007) and Yin et al. (2008). The results are presented in Table 2, which shows that the estimates of β_1 and β_2 are much more stable when γ is assumed to be known. The biases of the estimates are very small, there is a good agreement between the estimated and empirical standard errors, and the empirical coverage probabilities are maintained at around 95%. By comparing Table 1 with Table 2, we can see that conditional inference for β_1 and β_2 is much more efficient than that when γ needs to be estimated. This implies that taking γ as an unknown parameter highly inflates the variability for the regression parameter estimates.

We also conducted simulation studies to examine the performance of the proposed method with skewed and het-

Table 1. Simulation results with normal error when γ is unknown

c%	$\beta_1 = 0.5$ Bias	SD(SE)	CP	$\beta_2 = 1.0$ Bias	SD(SE)	CP	γ Bias	SD(SE)	CP
					$\gamma = 0.0$				
0	0.0054	0.1596(0.1471)	0.944	0.0240	0.2181(0.1871)	0.932	-0.0067	0.2688(0.2243)	0.944
20	0.0029	0.1626(0.1515)	0.944	0.0195	0.2103(0.1863)	0.944	-0.0108	0.2580(0.2208)	0.946
40	0.0188	0.1614(0.1627)	0.942	0.0503	0.2121(0.2286)	0.954	0.0050	0.2501(0.2161)	0.946
					$\gamma = 0.5$				
0	-0.0001	0.1395(0.1349)	0.938	0.0147	0.1764(0.1542)	0.940	-0.0072	0.2590(0.2168)	0.942
20	-0.0055	0.1458(0.1440)	0.944	0.0011	0.1707(0.1863)	0.950	-0.0285	0.2418(0.2138)	0.938
40	-0.0008	0.1515(0.1593)	0.954	0.0073	0.1712(0.1879)	0.956	-0.0225	0.2403(0.2087)	0.956
					$\gamma = 1.0$				
0	-0.0021	0.1301(0.1292)	0.940	0.0077	0.1460(0.1358)	0.960	-0.0129	0.2399(0.2133)	0.966
20	-0.0039	0.1357(0.1398)	0.958	0.0048	0.1445(0.1398)	0.948	-0.0176	0.2327(0.2095)	0.966
40	-0.0052	0.1498(0.1579)	0.948	0.0008	0.1537(0.1431)	0.952	-0.0339	0.2324(0.2054)	0.948

Note: c% stands for the censoring rate.

Table 2. Simulation results with normal error when γ is known

γ	c%	$\beta_1 = 0.5$ Bias	SD	SE	CP	$\beta_2 = 1.0$ Bias	SD	SE	CP
0.0	0	0.0030	0.0789	0.0809	0.950	0.0004	0.0451	0.0462	0.952
	20	0.0020	0.0810	0.0895	0.945	0.0006	0.0466	0.0492	0.952
	40	0.0011	0.0914	0.0978	0.947	0.0030	0.0527	0.0575	0.953
0.5	0	0.0030	0.0790	0.0812	0.943	0.0011	0.0437	0.0465	0.957
	20	0.0041	0.0860	0.0890	0.945	0.0020	0.0479	0.0513	0.947
	40	0.0031	0.0993	0.1006	0.946	0.0054	0.0557	0.0584	0.943
1.0	0	0.0003	0.0785	0.0806	0.952	0.0003	0.0440	0.00465	0.953
	20	0.0000	0.0870	0.0895	0.947	0.0009	0.0500	0.0516	0.954
	40	0.0055	0.0953	0.1021	0.955	0.0032	0.0585	0.0597	0.946

Note: c% stands for the censoring rate.

Table 3. Simulation results with skewed or heteroscedastic errors when γ is unknown

c%	$\beta_1 = 0.5$ Bias	SD(SE)	CP	$\beta_2 = 1.0$ Bias	SD(SE)	CP	$\gamma = 0.5$ Bias	SD(SE)	CP
					Skewed				
0	-0.0064	0.4077(0.4115)	0.964	0.0240	0.2848(0.2923)	0.968	-0.0117	0.2387(0.2074)	0.932
20	-0.0331	0.4356(0.4557)	0.952	-0.0111	0.3113(0.3430)	0.964	-0.0163	0.2297(0.2057)	0.966
40	0.0157	0.5256(0.6121)	0.966	0.0482	0.3424(0.3183)	0.957	-0.0034	0.2190(0.2040)	0.972
					Heteroscedastic				
0	-0.0253	0.4146(0.4763)	0.962	0.0149	0.3502(0.3445)	0.956	-0.0345	0.2541(0.2240)	0.980
20	0.0352	0.5201(0.5358)	0.942	0.0962	0.4002(0.5010)	0.940	-0.0031	0.2374(0.2110)	0.952
40	0.0456	0.6040(0.6678)	0.969	0.3088	0.7811(0.8073)	0.981	0.0158	0.2137(0.2032)	0.976

Note: c% stands for the censoring rate.

eroscedastic errors. For the skewed error case, we considered model (6), and took the error ε from a shifted chi-squared distribution with 1 degree of freedom and a median of 0. For the heteroscedastic error case, we considered the following Box–Cox transformation linear quantile regression model:

$$H_\gamma(T) = \beta_1 Z_1 + \beta_2 Z_2 \xi + \varepsilon, \tag{7}$$

where $\xi \sim$ Exponential(1) independent of ε, and $\varepsilon \sim N(0,1)$. All other setups were the same as before. The simulation results in Tables 3 and 4 are reported in the same manner as those in Tables 1 and 2 but with $\gamma = 0.5$. It can be seen from the tables that the performance of the proposed method in the skewed and heteroscedastic error cases is robust and as satisfactory as that in the normal error case. That is, the proposed estimators have small

Table 4. Simulation results with skewed or heteroscedastic errors when γ is known

		$\beta_1 = 0.5$				$\beta_2 = 1.0$			
Error	c%	Bias	SD	SE	CP	Bias	SD	SE	CP
Skewed	0	0.0013	0.2654	0.2758	0.952	0.0034	0.1555	0.1636	0.955
	20	0.0017	0.1886	0.1957	0.953	0.0095	0.2914	0.3030	0.956
	40	-0.0133	0.2977	0.3092	0.952	0.0372	0.1962	0.2174	0.968
Heteroscedastic	0	0.0002	0.3047	0.3186	0.954	0.0021	0.1060	0.1098	0.954
	20	0.0002	0.3508	0.3560	0.938	0.0009	0.1835	0.1992	0.950
	40	-0.0133	0.2978	0.3092	0.952	0.0372	0.1962	0.2174	0.968

Note: c% stands for the censoring rate.

Table 5. Comparison results with normal error and unconditionally independent censoring when γ is unknown

		$\beta_1 = 0.5$			$\beta_2 = 1.0$			γ		
c%	Method	Bias	SD(SE)	CP	Bias	SD(SE)	CP	Bias	SD(SE)	CP
						$\gamma = 0.0$				
0	Our	0.0030	0.1480(0.1492)	0.946	0.0124	0.2126(0.1867)	0.938	-0.0170	0.2642(0.2251)	0.952
	YZL	0.0072	0.0857(0.0838)	0.932	0.0407	0.1739(0.1551)	0.936	0.0346	0.2339(0.1992)	0.942
20	Our	0.0023	0.1493(0.1549)	0.954	0.0103	0.2048(0.1862)	0.946	-0.0183	0.2530(0.2202)	0.952
	YZL	0.0036	0.0894(0.0957)	0.964	0.0116	0.1687(0.1595)	0.958	0.0035	0.2109(0.1941)	0.970
40	Our	0.0203	0.1522(0.1647)	0.950	0.0367	0.2088(0.1894)	0.936	-0.0006	0.2454(0.2125)	0.946
	YZL	-0.0072	0.1005(0.1293v	0.968	0.0012	0.1570(0.1823)	0.979	0.0394	0.1841(0.1908)	0.968
						$\gamma = 0.5$				
0	Our	0.0048	0.1363(0.1368)	0.946	0.0136	0.1759(0.1538)	0.926	-0.0042	0.2577(0.2178)	0.932
	YZL	0.0097	0.0814(0.0815)	0.940	0.0441	0.1481(0.1370)	0.942	0.0574	0.2259(0.2044)	0.940
20	Our	0.0059	0.1388(0.1466)	0.956	0.0116	0.1742(0.1557)	0.940	-0.0065	0.2511(0.2139)	0.938
	YZL	0.0058	0.0883(0.0956)	0.960	0.0106	0.1452(0.1419)	0.953	0.0113	0.2012(0.1987)	0.986
40	Our	-0.0017	0.1550(0.1613)	0.960	0.0152	0.1723(0.1622)	0.954	-0.0082	0.2360(0.2073)	0.944
	YZL	-0.0107	0.1062(0.1186)	0.962	-0.0147	0.1373(0.1568)	0.972	0.1989	0.2076(0.2074)	0.988
						$\gamma = 1.0$				
0	Our	0.0003	0.1478(0.1376)	0.948	-0.0009	0.1290(0.1313)	0.934	-0.0220	0.2478(0.2133)	0.950
	YZL	0.0112	0.0792(0.0798)	0.928	0.0420	0.1360(0.1239)	0.932	0.0692	0.2327(0.2054)	0.930
20	Our	0.0005	0.1322(0.1410)	0.956	0.0003	0.1478(0.1376)	0.948	-0.0210	0.2378(0.2095)	0.948
	YZL	0.0050	0.0869(0.0936)	0.956	0.0102	0.1319(0.1302)	0.962	0.0139	0.2090(0.2004)	0.958
40	Our	-0.0007	0.1588(0.1576)	0.948	0.0172	0.1489(0.1440)	0.958	0.0035	0.2210(0.2052)	0.958
	YZL	-0.0050	0.1075(0.1157)	0.958	-0.0067	0.1308(0.1447)	0.967	-0.0052	0.1913(0.2027)	0.984

Note: c% stands for the censoring rate, and YZL stands for the method of Yin et al. (2008).

biases, reasonable variance estimates and the empirical coverage probabilities. In addition, Table 4 indicates that the conditional inference is more stable and accurate when γ is assumed to be known.

An additional simulation study was conducted for comparison with the method of Yin et al. (2008) (denoted by YZL). Note that the YZL's method needs the censoring time C to be independent of T and $\mathbf{Z}$. Thus, we considered model (6) with the same setup as Table 1, except that two situations for the censoring time C were investigated: (I) the unconditionally independent censoring with C generated from a uniform distribution $U(0.5, V)$; (II) the conditionally independent censoring with C assumed to follow a uniform distribution $U(0, 3Z_1V)$, where V was selected to give a censoring rate of $0\%, 20\%$ or 40%. As in Yin et al. (2008), under each configuration, we generated 500 simulated data sets of sample size $n = 300$. To obtain the estimated standard error, we used the bootstrap method and the resampling method with 400 resampled data sets for the YZL's method and the proposed method, respectively. Tables 5 and 6 summarize the estimation results under situations (I) and (II), respectively. It can be seen from Table 5 that, under the unconditionally independent censoring, both methods provide reasonable and comparable estimates, and the YZL estimator seems a little more efficient than our proposed estimator. This is because the former utilizes the unconditionally independent censoring assumption in estimation. However, when such unconditionally independent assumption is violated, it can be seen from Table 6 that the YZL's method may lead to biases, especially when the censoring rate is high (say 40%). Our proposed estimators are essentially unbiased in all settings. Thus, the proposed method is more flexible and robust than the YZL's method. We also considered other setups and the results were similar to those given above.

Table 6. Comparison results with normal error and conditionally independent censoring when γ is unknown

		Our						YZL					
		$\beta_1 = 0.5$		$\beta_2 = 1.0$		γ		$\beta_1 = 0.5$		$\beta_2 = 1.0$		γ	
γ	c%	Bias	SD	Bias	SD	Bias	SD	Bias	SD	Bias	SD	Bias	SD
	0	0.0062	0.1316	0.0216	0.2356	-0.0124	0.3001	-0.0004	0.0637	0.0143	0.1652	-0.0026	0.2316
0.0	20	-0.0031	0.1252	0.0075	0.2280	-0.0285	0.2806	0.0275	0.0475	-0.2346	0.0872	-0.3018	0.1350
	40	0.0016	0.1246	0.0325	0.2343	-0.0002	0.2727	0.1199	0.0652	-0.2811	0.1096	-0.2653	0.1697
	0	0.0010	0.1071	0.0122	0.1805	-0.0067	0.2797	0.0069	0.0574	0.0368	0.1534	0.0399	0.2484
0.5	20	-0.0038	0.1084	0.0137	0.1785	-0.0003	0.2691	0.0541	0.0496	-0.1631	0.1020	-0.2099	0.1729
	40	-0.0070	0.1140	0.0175	0.1837	-0.0070	0.2672	0.1516	0.0663	-0.1974	0.1123	-0.1496	0.1826
	0	0.0013	0.0958	0.0035	0.1580	-0.0181	0.2736	0.0034	0.0558	0.0194	0.1352	0.0187	0.2594
1.0	20	0.0060	0.1067	0.0205	0.1534	0.0109	0.2651	0.0029	0.1626	0.0195	0.2103	-0.1846	0.1804
	40	-0.0052	0.1154	0.0022	0.1469	-0.0202	0.2497	0.1632	0.0662	-0.1779	0.1123	-0.1252	0.2108

Note: c% stands for the censoring rate, and YZL stands for the method of Yin et al. (2008).

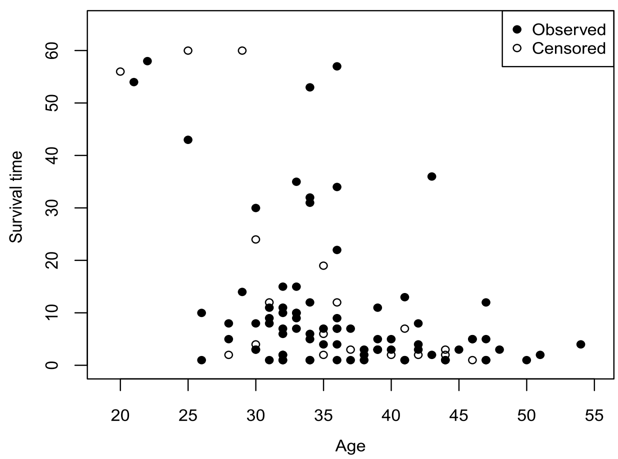

Figure 1. The scatter plot of the survival times against Age.

5. AN APPLICATION

For illustration purposes, we applied the proposed methods to the HMO data from Hosmer and Lemeshow (1999). In this data set, there were 100 HIV positive subjects who were followed until death due to AIDS or AIDS-related complications, until the end of the study, or until the subject was lost to follow-up. The outcome variable was survival time after a confirmed diagnosis of HIV, and 20% of subjects were censored. Two covariates of interest were Age (denoted by Z_1): the age of the subject at entry (in years), and Drug (denoted by Z_2): history of prior IV drug use (0 = No, 1 = Yes). As discussed in Hosmer and Lemeshow (1999), the censoring time C was assumed to be independent of the survival time T conditional on $\mathbf{Z} = (Z_1, Z_2)^T$. Here we are focus on the covariate effects on the quantiles of the survival time.

Figure 1 shows the scatter plot of the survival times against Age. It suggests that some transformations might be needed to achieve linearity. We applied model (1) to the data, where $H_\gamma(t)$ was taken to be the Box–Cox transformation:

$$H_\gamma(t) = \begin{cases} (t^\gamma - 1)/\gamma & \text{if } \gamma \neq 0, \\ \ln(t) & \text{if } \gamma = 0. \end{cases}$$

We considered the value of τ from 0.2 to 0.5 in steps of 0.01, and used 300 resampled data sets for variance estimation. This quantile region is of interest since lower quantiles of the survival time have an immediate concern to HIV subjects, and have significant biomedical implications in the short term. The search range of the optimal γ was taken in the interval $[-2, 2]$. The application of the proposed method gave the estimate of the Box–Cox transformation parameter as $\hat{\gamma} = 1.5435$ with estimated standard error of 0.7977 (p-value = 0.0529), which implies that the transformation parameter is marginally significantly different from zero, and the Box–Cox transformation seems reasonable to analyze this data. Figure 2 displays the estimated quantile regression coefficients $\hat{\boldsymbol{\beta}}(\tau)$ with the pointwise 95% confidence band. This means that both Drug and Age are significant across the quantiles. Such varying effects would not have been identified by the original Cox model or classic linear regression model. In addition, both Drug and Age have negative associations with the survival time. In particular, older subjects are more likely to die than younger subjects, and those with a prior history of IV drug use tend to die sooner than those who do not have a history of IV drug use. These results are consistent with those obtained by the Cox model (Hosmer and Lemeshow, 1999, p. 105). However, the quantile regression models provide substantially more information, and thus presents a global view of the relationship between the survival time and the covariates.

To examine the effects of the covariates in the original scale of the outcome, we may evaluate their marginal effects. Following Mu and He (2007) and Yin et al. (2008), if the jth covariate of $\tilde{\mathbf{Z}}$ is continuous, then its marginal effect is defined as

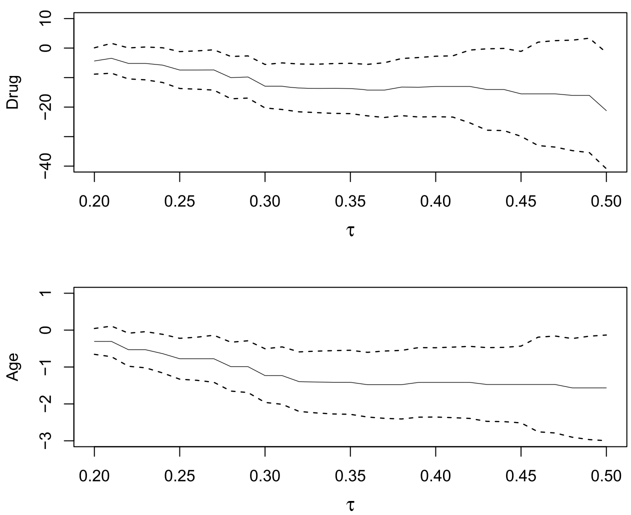

Figure 2. Estimates of quantile regression coefficients with the pointwise 95% confidence band. The solid lines are the estimates, and the dashed lines are their pointwise 95% confidence bands.

$$\frac{\partial H_\gamma^{-1}\{\mathbf{Z}^T\boldsymbol{\beta}(\tau)\}}{\partial Z_j}\bigg|_{\tilde{\mathbf{Z}}} = \begin{cases}\beta_j(\tau)(\gamma\tilde{\mathbf{Z}}^T\boldsymbol{\beta}(\tau)+1)^{1/\gamma-1}, & \gamma \neq 0;\\ \beta_j(\tau)\exp\{\tilde{\mathbf{Z}}^T\boldsymbol{\beta}(\tau)\}, & \gamma = 0,\end{cases}$$

where $\beta_j(\tau)$ is the jth component of $\boldsymbol{\beta}(\tau)$. If the jth covariate is discrete taking values 0 and 1, then its marginal effect is given by

$$H_\gamma^{-1}\{\tilde{\mathbf{Z}}^T\boldsymbol{\beta}(\tau)\}|_{Z_j=1} - H_\gamma^{-1}\{\tilde{\mathbf{Z}}^T\boldsymbol{\beta}(\tau)\}|_{Z_j=0}.$$

Figure 3 presents the estimated marginal effects of Drug (=1) and Age (=35) with the pointwise 95% confidence band based the proposed method and the YZL's method, respectively. It can be seen that the marginal effects of Drug and Age are quite different between the two methods. Our proposed method provides tighter confidence intervals than the YZL's method, and thus is more efficient than the YZL's method. In addition, the YZL's method would overestimate the marginal effects.

6. DISCUSSION

In this article, we proposed an estimation procedure for a class of power-transformed linear quantile regression models with censored survival data. The implementation of the proposed method involves only minimizing two convex objective functions that guarantee a unique solution. The asymptotic properties of the proposed estimators were derived, and the resampling approach was used to estimate the asymptotic covariance. The simulation results showed that the proposed estimation approach performs well. An application to the HMO data was provided to illustrate our method. When the dimensionality of the covariates is high, and some transformations for the response variable might be needed to achieve linearity, the proposed method should be used in practice.

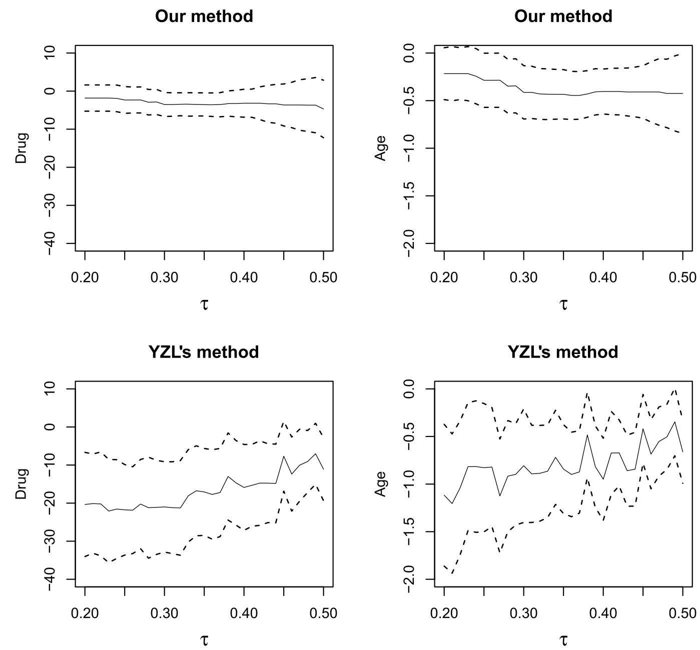

Figure 3. The estimated marginal effects given Drug (=1) and Age (=35). The solid lines are the estimates, and the dashed lines are their pointwise 95% confidence bands. Our method: the proposed method; YZL's method: the method of Yin et al. (2008).

Note that the proposed estimation procedure may not always ensure the monotonicity of $\mathbf{Z}^T\hat{\boldsymbol{\beta}}(\tau)$. However, a simple modification of the quantile prediction can be made by $\sup_{u\le\tau} H_{\hat{\gamma}}^{-1}\{\mathbf{Z}^T\hat{\boldsymbol{\beta}}(u)\}$, which is a nondecreasing function of τ, and is asymptotically equivalent to $H_{\hat{\gamma}}^{-1}\{\mathbf{Z}^T\hat{\boldsymbol{\beta}}(\tau)\}$ (e.g., Peng and Huang, 2008).

In practice, the choice of an appropriate power transformation $H_\gamma(\cdot)$ may be based on prior data or the desiring interpretation of the regression coefficients. For a given power transformation, following Lin et al. (1993), we can use a residual-based procedure for checking the adequacy of the model. Of course, it would be desirable to develop some data-driven methods for the model checking. This is a challenging problem and requires further research efforts.

ACKNOWLEDGEMENTS

The authors would like to thank the Editor, Professor Heping Zhang, an Associate Editor and three referees for their constructive and insightful comments and suggestions that greatly improved the article. The second author's research was partly supported by the National Natural Science Foundation of China (Grant Nos. 11231010 and 11171330), Key Laboratory of RCSDS, CAS (No. 2008DP173182) and BCMIIS.

Received 12 June 2014

REFERENCES

Amemiya, T. (1985). *Advanced Econometrics*. Harvard University Press, Cambridge, MA.

Andersen, P. K., Borgan, Ø., Gill, R. D. and Keiding, N. (1998). *Statistical Models Based on Counting Processes*, 2nd ed. Springer, New York. MR1198884

Arcones, M. A. and Giné, E. (1993). Limit theorems for U-processes. *Ann. Probab.* **21** 1494–1542. MR1235426

Buchinsky, M. (1995). Quantile regression, Box–Cox transformation model, and the U.S. wage structure, 1963–1987. *J. Econometrics* **65** 109–154.

Buckley, J. and James, I. (1979). Linear regression with censored data. *Biometrika* **66** 429–436.

Cai, T., Tian, L. and Wei, L. J. (2005). Semiparametric Box–Cox power transformation models for censored survival observations. *Biometrika* **92** 619–632. MR2202650

Cox, D. R. (1972). Regression models and life-tables (with discussion). *J. R. Statist. Soc. B* **34** 187–220. MR0341758

Fleming, T. R. and Harrington, D. P. (1991). *Counting Processes and Survival Analysis*. Wiley, New York. MR1100924

Gill, R. D. and Johansen, S. (1990). A survey of product-integration with a view toward application in survival analysis. *Ann. Statist.* **4** 1501–1555. MR1074422

Hosmer, D. W. and Lemeshow, S. (1999). *Applied Survival Analysis: Regression Modeling of Time to Event Data*. Wiley, New York. MR1674644

Honoré, B., Khan, S. and Powell, J. L. (2002). Quantile regression under random censoring. *J. Econometrics* **109** 67–105. MR1899693

Huang, Y. (2010). Quantile calculus and censored regression. *Ann. Statist.* **38** 1607–1637. MR2662354

Jin, Z., Ying, Z. and Wei, L. J. (2001). A simple resampling method by perturbing the minimand. *Biometrika* **88** 381–390. MR1844838

Koenker, R. and Geling, O. (2001). Reappraising medfly longevity: A quantile regression survival analysis. *J. Amer. Statist. Assoc.* **96** 458–468. MR1939348

Lai, T. L. and Ying, Z. (1988). Stochastic integrals of empirical-type processes with applications to censored regression. *J. Multivariate Anal.* **27** 334–358. MR0970959

Leng, C. and Tong, X. (2014). Censored quantile regression via Box–Cox transformation under conditional independence. *Statist. Sinica* **24** 221–249. MR3183682

Lin, D. Y., Wei, L. J. and Ying, Z. (1993). Checking the Cox model with cumulative sums of martingale-based residuals. *Biometrika* **80** 557–572. MR1248021

Linz, P. (1985). *Analytical and Numerical Methods for Volterra Equations.* Vol. 7 of SIAM Studies in Applied Mathematics. SIAM, Philadelphia. MR0796318

Machado, J. A. F. and Mata, J. (2000). Box–Cox quantile regression and the distribution of firm sizes. *J. Appl. Econometrics* **15** 253–274.

Mu, Y. (2005). *Power Transformation Towards Linear or Partially Linear Quantile Regression Models*. Ph.D. Dissertation of University of Illinois at Urbana-Champaign. MR2708202

Mu, Y. and He, X. (2007). Power transformation toward a linear regression quantile. *J. Amer. Statist. Assoc.* **102** 269–279. MR2293308

Neocleous, T., Branden, K. V. and Portnoy, S. (2006). Correction to censored regression quantiles by S. Portnoy, **98** (2003), 1001–1012. *J. Amer. Statist. Assoc.* **101** 860–861. MR2041488

Pakes, A. and Pollard, D. (1989). Simulation and the asymptotics of optimization estimators. *Econometrica* **57** 1027–1057. MR1014540

Peng, L. and Huang, Y. (2008). Survival analysis with quantile regression models. *J. Amer. Statist. Assoc.* **103** 637–649. MR2435468

Polyanin, A. D. and Manzhirov, A. V. (2008). *Handbook of Integral Equations*, 2nd ed. Chapman and Hall/CRC, Boca Raton, FL. MR2404728

Portnoy, S. (2003). Censored regression quantiles. *J. Amer. Statist. Assoc.* **98** 1001–1012. MR2041488

Powell, J. L. (1984). Least absolute deviations estimation for the censored regression model. *J. Econometrics* **25** 303–325. MR0752444

Powell, J. L. (1986). Censored regression quantiles. *J. Econometrics* **32** 143–155. MR0853049

Powell, J. L. (1991). Estimation of monotonic regression models under quantile restrictions. In *Nonparametric and Semiparametric Methods in Econometrics* (ed. by W. Barnett, J. Powell and G. Tauchen). Cambridge University Press, Cambridge. MR1174980

Qian, J. and Peng, L. (2010). Censored quantile regression with partially functional effects. *Biometrika* **97** 839–850. MR2746155

Serfling, R. J. (2002). *Approximation Theorems of Mathematical Statistics*, 2nd ed. Wiley, New York. MR0595165

Tang, Y., Wang, H., He, X. and Zhu, Z. (2012). An informative subset-based estimator for censored quantile regression. *TEST* **21** 635–655. MR2992086

Van der Vaart, A. W. (1998). *Asymptotic Statistics*. Cambridge University Press, Cambridge. MR1652247

Van der Vaart, A. W. and Wellner, J. A. (1996). *Weak Convergence and Empirical Processes: With Applications to Statistics*. Springer-Verlag, New York. MR1385671

Wang, H. and Wang, L. (2009). Locally weighted censored quantile regression. *J. Amer. Statist. Assoc.* **104** 1117–1128. MR2562007

Wang, H., Zhou, J. and Li, Y. (2013). Variable selection for censored quantile regression. *Statist. Sinica* **23** 145–167. MR3076162

Yang, S. (1999). Censored median regression using weighted empirical survival and hazard functions. *J. Amer. Statist. Assoc.* **94** 137–145. MR1689219

Ying, Z., Jung, S. H. and Wei, L. J. (1995). Survival analysis with median regression models. *J. Amer. Statist. Assoc.* **90** 178–184. MR1325125

Yin, G., Zeng, D. and Li, H. (2008). Power-transformed linear quantile regression with censored data. *J. Amer. Statist. Assoc.* **103** 1214–1224. MR2462894

Rui Miao
Institute of Applied Mathematics
Academy of Mathematics and Systems Science
Chinese Academy of Sciences
Beijing 100190
People's Republic of China
E-mail address: ruimiao@amss.ac.cn

Liuquan Sun
Institute of Applied Mathematics
Academy of Mathematics and Systems Science
Chinese Academy of Sciences
Beijing 100190
People's Republic of China
E-mail address: slq@amt.ac.cn

Guo-Liang Tian
Department of Statistics and Actuarial Science
The University of Hong Kong
Pokfulam Road, Hong Kong
People's Republic of China
E-mail address: gltian@hku.hk