Smoothed and Corrected Score Approach to Censored Quantile Regression With Measurement Errors

Yuanshan Wu, Yanyuan Ma, and Guosheng Yin*

Abstract

Censored quantile regression is an important alternative to the Cox proportional hazards model in survival analysis. In contrast to the usual central covariate effects, quantile regression can effectively characterize the covariate effects at different quantiles of the survival time. When covariates are measured with errors, it is known that naively treating mismeasured covariates as error-free would result in estimation bias. Under censored quantile regression, we propose smoothed and corrected estimating equations to obtain consistent estimators. We establish consistency and asymptotic normality for the proposed estimators of quantile regression coefficients. Compared with the naive estimator, the proposed method can eliminate the estimation bias under various measurement error distributions and model error distributions. We conduct simulation studies to examine the finite-sample properties of the new method and apply it to a lung cancer study.

KEY WORDS: Censored data; Check function; Corrected estimating equation; Measurement error; Kernel smoothing; Regression quantile; Semiparametric method; Survival analysis.

Short title: Measurement error censored quantile regression

*Yuanshan Wu is Assistant Professor, School of Mathematics and Statistics, Wuhan University, Wuhan, Hubei 430072, China (E-mail: shan@whu.edu.cn). Yanyuan Ma is Professor, Department of Statistics, University of South Carolina, Columbia, SC 29208, U.S.A. (E-mail: yanyuan.ma@stat.sc.edu). Guosheng Yin, the corresponding author, is Professor, Department of Statistics and Actuarial Science, The University of Hong Kong, Pokfulam Road, Hong Kong (E-mail: gyin@hku.hk). We thank two referees, the associate editor and editor for their many constructive comments that have led to significant improvements in the article. Wu's research was partially supported by National Natural Science Foundation of China, Ma's research by National Science Foundation and National Institute of Neurological Disorder and Stroke, and Yin's research by the Research Grants Council of Hong Kong.

1 Introduction

Mean-based regression models have been extensively studied for randomly censored survival data. For example, the Cox (1972) proportional hazards model characterizes the hazard as a function of different covariates; and the accelerated failure time (AFT) model directly formulates linear regression between the logarithm of the failure time and covariates. However, neither the Cox nor the AFT model can differentiate the covariate effect at higher or lower quantiles of survival times, as they only provide the mean effect. In particular, the AFT model concerns only the mean regression, for which the estimation procedure is typically based on the least squares or rank methods (Prentice, 1978; Buckley and James, 1979; Ritov, 1990; Tsiatis, 1990; Wei, Ying, and Lin, 1990; Lai and Ying, 1991; and Jin et al., 2003). On the other hand, quantile regression provides a robust alternative to mean-based regression models. Under this framework, we can model the median or any other quantile of the outcome or survival time (Koenker and Bassett, 1978; and Koenker, 2005). Regression parameters are often estimated by minimizing a check function, and the corresponding variance estimates are typically obtained by resampling methods, such as bootstrap. When censoring times are assumed to be fixed and known, quantile regression has been extensively studied, particularly in the field of econometrics; for example, see Powell (1984), Buchinsky and Hahn (1998), Fitzenberger (1997), and Khan and Powell (2001). In survival analysis with random censoring, censored quantile regression (CQR) has been proposed and is gaining much popularity (Ying, Jung, and Wei, 1995; Lindgren, 1997; Yang, 1999; Koenker and Geling, 2001; Bang and Tsiatis, 2002; Chernozhukov and Hong, 2002; Portnoy, 2003; Peng and Huang, 2008; and Wang and Wang, 2009).

In practice, covariates are often subject to measurement errors. The most common measurement error structure is $\mathbf{W} = \mathbf{Z} + \mathbf{U}$, where $\mathbf{W}$ is the observed surrogate, $\mathbf{Z}$ is the true but unobserved covariate, and $\mathbf{U}$ is the random measurement error. For a comprehensive coverage of various measurement error models and inference procedures with mean-based regression, see Carroll et al. (2006). In the context of quantile regression with measurement errors, Brown (1982) examined median regression and described the difficulty involved in parameter estimation. He and Liang (2000) proposed root-n consistent estimators for linear and partially linear quantile regression models. Their method assumes that the random error in the response and the measurement errors in the covariates follow a spherical symmetric distribution. Wei and Carroll (2009) proposed a novel approach to quantile regression with measurement errors by utilizing the derivative property of the quantile function when the same quantile regression structure is assumed for all the quantile levels. Recently, Wang, Stefanski, and Zhu (2012) developed a corrected-loss function for the smoothed check function, a substantial advance in this area. However, there is limited research on quantile regression with covariate measurement errors under censoring. Ma and Yin (2011) studied covariate measurement errors in CQR models based on the inverse probability weighting scheme, but their method also requires the spherically symmetric distribution. In this paper, we study the issue of covariate measurement errors in quantile regression with randomly censored data. We propose a smoothed and corrected martingale-based estimating equation, consider grid-based estimates for the quantile regression coefficients, and establish the asymptotic properties of the proposed estimator by employing empirical process theory. Our proposed method allows an abundant class of distributions for the error in the response; for example, it could be light- or heavy-tailed, symmetric or asymmetric, homoscedastic or heteroscedastic.

The rest of the article is organized as follows. In Section 2, we describe the CQR model with measurement errors, develop a corrected estimating equation based on a kernel smoothing approximation, and establish the asymptotic properties of the resultant estimator. Section 3 contains simulation studies for the evaluation of the finite sample performance of the proposed method. A data set concerning lung cancer is analyzed in Section 4 and some concluding remarks are provided in Section 5. The assumptions that we imposed in the paper were listed and discussed in the Appendix and the detailed proofs of theorems are deferred to the online Supplementary Material.

2 CQR Model With Measurement Errors

2.1 Model Specification

Let T denote the transformed failure time under a known monotone transformation, e.g., the logarithm function. Let C denote the censoring time under the same transformation. Let $\mathbf{Z}$ be a p-vector of covariates, $X = T \wedge C$ be the observed time, and $\Delta = I(T \leq C)$ be the censoring indicator, where $a \wedge b$ is the minimum of a and b, and $I(\cdot)$ is the indicator function. Assume that T and C are conditionally independent given covariate $\mathbf{Z}$.

For $\tau \in (0, 1)$, the conditional τth quantile function of survival time T given covariate $\mathbf{Z}$ is defined as $Q_T(\tau|\mathbf{Z}) = \inf\{t\colon P(T \leq t|\mathbf{Z}) \geq \tau\}$. The quantile regression model associated with covariate $\mathbf{Z}$ has the form

$$Q_T(\tau|\mathbf{Z}) = \mathbf{Z}^{\mathrm{T}}\boldsymbol{\beta}(\tau), \tag{2.1}$$

where $\boldsymbol{\beta}(\tau)$ is an unknown p-vector of regression coefficients, representing the effect of $\mathbf{Z}$ on the τth quantile of the transformed survival time.

In reality covariate $\mathbf{Z}$ may be measured with errors, so that we do not directly observe $\mathbf{Z}$ but its surrogate $\mathbf{W}$. We assume the classical error structure

$$\mathbf{W} = \mathbf{Z} + \mathbf{U},$$

where $\mathbf{U}$ is a p-variate random vector with mean $\mathbf{0}$ and covariance matrix $\boldsymbol{\Sigma}$. The case that some covariates are error-free is accommodated in our model by setting the relevant terms in $\boldsymbol{\Sigma}$ to be zero. We further make the typical surrogacy assumption that (T, C) and $\mathbf{W}$ are conditionally independent given covariate $\mathbf{Z}$. For ease of exposition, we assume $\boldsymbol{\Sigma}$ to be known provisionally, since $\boldsymbol{\Sigma}$ can easily be estimated with replicated observations or validation data.

2.2 Approximately Corrected Estimating Equation

We first introduce notation: $F_T(t|\mathbf{Z}) = P(T \leq t|\mathbf{Z})$, $\Lambda_T(t|\mathbf{Z}) = -\log\{1 - P(T \leq t|\mathbf{Z})\}$, $N(t) = \Delta I(X \leq t)$ and $M(t) = N(t) - \Lambda_T(t \wedge X|\mathbf{Z})$. Following the argument in Fleming and

Harrington (1991), it is easy to show that evaluated at $\boldsymbol{\beta}_0(\tau)$, the true value of $\boldsymbol{\beta}(\tau)$, $M(t)$ is a martingale process associated with the counting process $N(t)$. Furthermore, because $E\{M(t)|\mathbf{Z}\} = 0$ at $\boldsymbol{\beta}_0(\tau)$ for any t, we have

$$E\left\{\mathbf{Z}\left(N\{\mathbf{Z}^{\mathrm{T}}\boldsymbol{\beta}_0(\tau)\} - \Lambda_T[\{\mathbf{Z}^{\mathrm{T}}\boldsymbol{\beta}_0(\tau)\} \wedge X|\mathbf{Z}]\right)\right\} = \mathbf{0} \tag{2.2}$$

for $\tau \in (0, 1)$. Under model (2.1), after some algebraic manipulations, we obtain that

$$\Lambda_T[\{\mathbf{Z}^{\mathrm{T}}\boldsymbol{\beta}_0(\tau)\} \wedge X|\mathbf{Z}] = \int_0^\tau I\{X \geq \mathbf{Z}^{\mathrm{T}}\boldsymbol{\beta}_0(u)\}\mathrm{d}H(u), \tag{2.3}$$

where $H(u) = -\log(1-u)$ for $0 \leq u < 1$.

Based on (2.2) and (2.3), when all $\mathbf{Z}_i$'s are observed, Peng and Huang (2008) proposed an estimating equation for $\boldsymbol{\beta}(\tau)$,

$$\sum_{i=1}^n \mathbf{Z}_i\left[N_i\{\mathbf{Z}_i^{\mathrm{T}}\boldsymbol{\beta}(\tau)\} - \int_0^\tau I\{X_i \geq \mathbf{Z}_i^{\mathrm{T}}\boldsymbol{\beta}(u)\}\mathrm{d}H(u)\right] = \mathbf{0}. \tag{2.4}$$

However, when the covariates $\mathbf{Z}_i$'s are measured with errors, naively treating mismeasured covariates to be error-free would cause estimation bias and thus lead to incorrect inference. In (2.4), because covariate $\mathbf{Z}_i$ lies inside the indicator function, which is discontinuous, it is difficult to build up a consistent estimator when the surrogates $\mathbf{W}_i$'s, instead of $\mathbf{Z}_i$'s, are observed. To overcome the challenge caused by discontinuity and measurement errors, we propose an approximately corrected estimating equation for (2.4) and further establish the asymptotic properties of the resultant estimators for regression quantile coefficients.

We denote the observed data $\mathcal{O} = (X, \Delta, \mathbf{W})$ and let $\mathcal{U} = (X, \Delta, \mathbf{Z})$. In view of the estimating equation (2.4), if we can find a function $g^*\{\mathcal{O}, \boldsymbol{\beta}(\tau)\}$ such that for $\tau \in (0, 1)$,

$$E[g^*\{\mathcal{O}, \boldsymbol{\beta}(\tau)\}|\mathcal{U}] = \mathbf{Z}I\{X > \mathbf{Z}^{\mathrm{T}}\boldsymbol{\beta}(\tau)\},$$

we can then follow the corrected score argument (Stefanski, 1989; Nakamura, 1990) to construct an unbiased estimating equation as

$$\sum_{i=1}^n \left[\Delta_i\mathbf{W}_i - \Delta_i g^*\{\mathcal{O}_i, \boldsymbol{\beta}(\tau)\} - \int_0^\tau g^*\{\mathcal{O}_i, \boldsymbol{\beta}(u)\}\mathrm{d}H(u)\right] = \mathbf{0}.$$

However, the cusp in the indicator function makes it difficult to find such a function. On the other hand, Horowitz (1992, 1998) proposed the smoothed maximum score estimator for the binary response model and the smoothed least absolute deviation for median regression. Motivated by the smoothing scheme, we circumvent the discontinuity stemming from the indicator function and consider a smoothing function that approaches the indicator function as $n \to \infty$. More specifically, assume that a smooth function $K(\cdot)$ satisfies $\lim_{x\to-\infty} K(x) = 0$ and $\lim_{x\to\infty} K(x) = 1$. If we consider a positive scale parameter h_n that converges to zero as sample size $n \to \infty$, $K(x/h_n)$ may provide an adequate approximation to $I(x > 0)$ as $n \to \infty$, where h_n behaves like the bandwidth in the kernel smoothing.

If we can find a function $G\{\mathcal{O}, \boldsymbol{\beta}(\tau); h_n\}$ such that

$$\begin{aligned} E[G\{\mathcal{O}, \boldsymbol{\beta}(\tau); h_n\}|\mathcal{U}] &= \{X - \mathbf{Z}^{\mathrm{T}}\boldsymbol{\beta}(\tau)\}K\left\{\frac{X - \mathbf{Z}^{\mathrm{T}}\boldsymbol{\beta}(\tau)}{h_n}\right\} \\ &\approx \{X - \mathbf{Z}^{\mathrm{T}}\boldsymbol{\beta}(\tau)\}I\{X > \mathbf{Z}^{\mathrm{T}}\boldsymbol{\beta}(\tau)\}, \end{aligned} \tag{2.5}$$

we may set

$$g\{\mathcal{O}, \boldsymbol{\beta}(\tau); h_n\} = -\frac{\partial G\{\mathcal{O}, \boldsymbol{\beta}(\tau); h_n\}}{\partial \boldsymbol{\beta}(\tau)},$$

and conclude that $E[g\{\mathcal{O}, \boldsymbol{\beta}(\tau); h_n\}|\mathcal{U}]$ is close to $\mathbf{Z}I\{X > \mathbf{Z}^{\mathrm{T}}\boldsymbol{\beta}(\tau)\}$. As a result, we can construct an approximately corrected estimating equation

$$\sum_{i=1}^{n}\left[\Delta_i \bar{g}\{\mathcal{O}_i, \boldsymbol{\beta}(\tau); h_n\} - \int_0^{\tau} g\{\mathcal{O}_i, \boldsymbol{\beta}(u); h_n\}\mathrm{d}H(u)\right] = \mathbf{0}, \tag{2.6}$$

where $\bar{g}\{\mathcal{O}_i, \boldsymbol{\beta}(\tau); h_n\} = \mathbf{W}_i - g\{\mathcal{O}_i, \boldsymbol{\beta}(\tau); h_n\}$. Since it is challenging to obtain the functional solution to the integral equation (2.6), we follow Peng and Huang (2008) to develop a grid-based estimation procedure for $\boldsymbol{\beta}_0(\cdot)$. Assume that τ_U is a deterministic constant in $(0, 1)$ subject to certain identifiability constraints, e.g., Assumption 4-(iii) in the Appendix. Due to the inherent nonidentifiability of the regression quantiles beyond the level τ_U, we confine estimation of $\boldsymbol{\beta}_0(\tau)$ for $\tau \in (0, \tau_U]$. We denote a partition over the interval $[0, \tau_U]$ by $\mathcal{S}_{q_n} = \{0 \equiv \tau_0 < \tau_1 < \cdots < \tau_{q_n} \equiv \tau_U\}$, where the number of grid points q_n depends on n. We consider an estimator of $\boldsymbol{\beta}_0(\tau)$ that is a right-continuous piecewise constant function and jumps only at grid points in $\mathcal{S}_{q_n}$. Noting that $\mathbf{Z}^{\mathrm{T}}\boldsymbol{\beta}_0(\tau_0) = -\infty$, we intuitively set $g\{\mathcal{O}, \widehat{\boldsymbol{\beta}}(\tau_0); h_n\} = \mathbf{W}$. For a given h_n, employing the Newton–Raphson algorithm, the estimates $\widehat{\boldsymbol{\beta}}(\tau_j)$, $j = 1, \ldots, q_n$, can be obtained sequentially by solving

$$\sum_{i=1}^{n} \left[\Delta_i \bar{g}\{\mathcal{O}_i, \boldsymbol{\beta}(\tau); h_n\} - \sum_{k=0}^{j-1} g\{\mathcal{O}_i, \widehat{\boldsymbol{\beta}}_n(\tau_k); h_n\}\{H(\tau_{k+1}) - H(\tau_k)\} \right] = \mathbf{0}. \tag{2.7}$$

2.3 Laplace and Normal Measurement Errors

Apparently, it is crucial to find the function G such that (2.5) holds. For illustration, we construct G when the measurement errors follow a multivariate Laplace and a multivariate normal distribution, respectively. Wang, Stefanski, and Zhu (2012) also considered these two types of measurement errors, as Laplace distributions are more heavy-tailed than normal distributions, and both are widely used in practice.

Assume that $\mathbf{U}$ is a p-variate Laplace distributed random vector with mean $\mathbf{0}$ and covariance matrix $\boldsymbol{\Sigma}$, denoted by $\mathbf{U} \sim L_p(\mathbf{0}, \boldsymbol{\Sigma})$, whose characteristic function is given by $\varphi(\mathbf{t}) = 1/(1 + 0.5\mathbf{t}^{\mathrm{T}}\boldsymbol{\Sigma}\mathbf{t})$ for $\mathbf{t} \in \mathbb{R}^p$ (Kotz, Kozubowski, and Podgorski, 2001). Thus,

$$\epsilon(\tau)|\mathcal{U} \sim L_1\big\{X - \mathbf{Z}^{\mathrm{T}}\boldsymbol{\beta}(\tau), \boldsymbol{\beta}(\tau)^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{\beta}(\tau)\big)\big\},$$

where $\epsilon(\tau) = X - \mathbf{W}^{\mathrm{T}}\boldsymbol{\beta}(\tau)$. Following the work of Hong and Tamer (2003) and Wang, Stefanski, and Zhu (2012), we have

$$G_L\{\mathcal{O}, \boldsymbol{\beta}(\tau); h_n\} = \epsilon(\tau) K\left\{\frac{\epsilon(\tau)}{h_n}\right\} - \frac{\boldsymbol{\beta}(\tau)^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{\beta}(\tau)}{2}\left[\frac{2}{h_n}K^{(1)}\left\{\frac{\epsilon(\tau)}{h_n}\right\} + \frac{\epsilon(\tau)}{h_n^2}K^{(2)}\left\{\frac{\epsilon(\tau)}{h_n}\right\}\right],$$

where $K^{(j)}(x) = \mathrm{d}^j K(x)/\mathrm{d}x^j$ for $j = 1, 2, 3, 4$. It is easy to show that $G_L\{\mathcal{O}, \boldsymbol{\beta}(\tau); h_n\}$ satisfies (2.5). Therefore,

$$\begin{aligned} g_L\{\mathcal{O}, \boldsymbol{\beta}(\tau); h_n\} \;=\;& \left[K\left\{\frac{\epsilon(\tau)}{h_n}\right\} + \frac{\epsilon(\tau)}{h_n}K^{(1)}\left\{\frac{\epsilon(\tau)}{h_n}\right\}\right]\mathbf{W} \\ &+ \left[\frac{2}{h_n}K^{(1)}\left\{\frac{\epsilon(\tau)}{h_n}\right\} + \frac{\epsilon(\tau)}{h_n^2}K^{(2)}\left\{\frac{\epsilon(\tau)}{h_n}\right\}\right]\boldsymbol{\Sigma}\boldsymbol{\beta}(\tau) \\ &- \left[\frac{3}{h_n^2}K^{(2)}\left\{\frac{\epsilon(\tau)}{h_n}\right\} + \frac{\epsilon(\tau)}{h_n^3}K^{(3)}\left\{\frac{\epsilon(\tau)}{h_n}\right\}\right]\frac{\boldsymbol{\beta}(\tau)^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{\beta}(\tau)}{2}\mathbf{W}. \end{aligned}$$

After plugging $g_L\{\mathcal{O}, \boldsymbol{\beta}(\tau); h_n\}$ in (2.7), we can solve for $\boldsymbol{\beta}(\tau)$.

We consider a more common case that $\mathbf{U}$ is a p-variate normal random vector with mean $\mathbf{0}$ and covariance matrix $\boldsymbol{\Sigma}$, i.e., $\mathbf{U} \sim N_p(\mathbf{0}, \boldsymbol{\Sigma})$. Note that

$$\epsilon(\tau)|\mathcal{U} \sim N\big\{X - \mathbf{Z}^{\mathrm{T}}\boldsymbol{\beta}(\tau), \boldsymbol{\beta}(\tau)^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{\beta}(\tau)\big\}.$$

Motivated by Stefanski (1989) and Wang, Stefanski, and Zhu (2012), we take the objective function $G_N\{\mathcal{O}, \boldsymbol{\beta}(\tau); h_n\}$ to be

$$G_N\{\mathcal{O}, \boldsymbol{\beta}(\tau); h_n\} = \sum_{j=0}^{\infty} \frac{\{-\boldsymbol{\beta}(\tau)^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{\beta}(\tau)\}^j}{2^j j!} \left[\epsilon(\tau) K\left\{\frac{\epsilon(\tau)}{h_n}\right\}\right]^{(2j)},$$

provided that $K(\cdot)$ is sufficiently smooth, where $\{xK(x/h_n)\}^{(0)} = xK(x/h_n)$ and

$$\left\{xK\left(\frac{x}{h_n}\right)\right\}^{(j)} = \frac{j}{h_n^{j-1}} K^{(j-1)}\left(\frac{x}{h_n}\right) + \frac{x}{h_n^j} K^{(j)}\left(\frac{x}{h_n}\right), \quad j = 1, 2, \ldots.$$

Consequently, $g_N\{\mathcal{O}, \boldsymbol{\beta}(\tau); h_n\}$ can be obtained by taking the derivative of $G_N\{\mathcal{O}, \boldsymbol{\beta}(\tau); h_n\}$. Although we can construct the approximately corrected estimating equation as (2.6) and theoretically define an estimator based on the resultant grid-based solution for $\boldsymbol{\beta}_0(\cdot)$, it is infeasible to solve the equation because G_N involves an infinite series. Following the recommendation of Stefanski (1989), we keep the first two summands in G_N as an approximation, which is found to be adequate in our simulation studies. More interestingly, using the first two summands leads to exactly the same form of the approximately corrected estimating equation as that in the Laplace measurement error model.

2.4 Asymptotic Properties

Denote $a_n = \max_{1\le j\le q_n} |\tau_j - \tau_{j-1}|$, the maximum distance between two adjacent points belonging to $\mathcal{S}_{q_n}$. The asymptotic properties of the estimator $\widehat{\boldsymbol{\beta}}(\tau)$, which solves (2.7), are summarized in the following two theorems.

Theorem 1 *Under Assumptions 1–4 in the Appendix, if $a_n = o(1)$, then $\sup_{\tau\in[\nu,\tau_U]} \|\widehat{\boldsymbol{\beta}}(\tau) - \boldsymbol{\beta}_0(\tau)\| \to 0$ in probability for any $\nu \in (0, \tau_U]$ as $n \to \infty$.*

Theorem 2 *Under Assumptions 1–5 in the Appendix, if $a_n = o(n^{-1/2})$, then $n^{1/2}\{\widehat{\boldsymbol{\beta}}(\tau) - \boldsymbol{\beta}_0(\tau)\}$ converges weakly to a mean zero Gaussian random field over $\tau \in [\nu, \tau_U]$ for any $\nu \in (0, \tau_U]$ as $n \to \infty$.*

Both the consistency and weak convergence of the proposed estimator only hold for quantile levels bounded away from zero due to the data sparsity when τ is close to zero. A much finer partition with a step size of order $o(n^{-1/2})$ is required to establish the weak convergence property. The proofs of both theorems rely heavily on empirical process theory, which are provided in the Supplementary Material.

It is crucial to select the smoothing parameter h_n. Without loss of generality, assume that $\mathbf{Z}$ includes the intercept as its first element. Noting that $E[g\{\mathcal{O}, \boldsymbol{\beta}(\tau); h_n\}|\mathcal{U}]$ is close to $\mathbf{Z}I\{X > \mathbf{Z}^{\mathrm{T}}\boldsymbol{\beta}(\tau)\}$ and utilizing only the intercept term, we can get the smoothed and corrected function

$$\mathcal{M}(\mathcal{O}, \boldsymbol{\beta}(\tau); h_n) = \Delta\bar{g}_1\{\mathcal{O}, \boldsymbol{\beta}(\tau); h_n\} - \int_0^{\tau} g_1\{\mathcal{O}, \boldsymbol{\beta}(u); h_n\}\mathrm{d}H(u)$$

for the martingale $M\{\mathbf{Z}^{\mathrm{T}}\boldsymbol{\beta}(\tau)\}$, where $\bar{g}_1$ and g_1 are the first elements of $\bar{g}$ and g, respectively. In practice, we recommend a d-fold cross-validation method to choose h_n. We randomly divide the data into d nonoverlapping and approximately equal-sized subgroups. For the jth subgroup $\mathcal{D}_j$, we fit the proposed procedure using the data excluding $\mathcal{D}_j$, denoted by $\mathcal{D}_{(-j)}$, and calculate the loss function

$$\mathcal{L}_j(h_n) = \frac{1}{|\mathcal{D}_j|}\sum_{k \in \mathcal{D}_j}\int_0^{\tau_U} |\mathcal{R}(h_n, \mathbf{W}_k^{\mathrm{o}}, \widehat{\boldsymbol{\beta}}_{(-j)}(\tau))|\mathrm{d}\tau,$$

where $|\mathcal{D}_j|$ denotes the cardinality of the set $\mathcal{D}_j$,

$$\mathcal{R}(h_n, \mathbf{w}^{\mathrm{o}}, \boldsymbol{\beta}(\tau)) = \frac{1}{|\mathcal{D}_{(-j)}|}\sum_{i \in \mathcal{D}_{(-j)}} I(\mathbf{W}_i^{\mathrm{o}} \le \mathbf{w}^{\mathrm{o}})\mathcal{M}(\mathcal{O}_i, \boldsymbol{\beta}(\tau); h_n),$$

$\mathbf{W}_i^{\mathrm{o}}$ denotes the error-free elements of $\mathbf{Z}_i$, and $\mathbf{W}_i^{\mathrm{o}} \le \mathbf{w}^{\mathrm{o}}$ means every entry of $\mathbf{W}_i^{\mathrm{o}}$ is not larger than the counterpart of $\mathbf{w}^{\mathrm{o}}$. The loss function is based on a cumulative sum of martingale residuals, with further correction on measurement errors. Here, $\widehat{\boldsymbol{\beta}}_{(-j)}(\tau)$ for $\tau \in [0, \tau_U]$ is obtained using the proposed procedure on the data $\mathcal{D}_{(-j)}$. Finally, we select the bandwidth by minimizing the total loss $\mathcal{L}(h_n) = \sum_{j=1}^{d} \mathcal{L}_j(h_n)$.

3 Simulation Studies

We conducted extensive simulation studies to assess the performance of the proposed method with finite samples. We generated survival time $\tilde{T}$ from the log-transformed linear model with heteroscedastic errors,

$$\log \tilde{T} = -0.5 + Z + (1 + 0.2Z)\epsilon,$$

where the model error ϵ was from the standard normal distribution, and Z was generated from the uniform distribution, $\mathrm{Unif}(0, \sqrt{12})$. The corresponding CQR model (2.1) given $\mathbf{Z} = (1, Z)^{\mathrm{T}}$ takes the form of

$$Q_T(\tau|\mathbf{Z}) = \beta_0(\tau) + \beta_1(\tau)Z,$$

where $T = \log \tilde{T}$, $\beta_0(\tau) = -0.5 + Q_\epsilon(\tau)$, $\beta_1(\tau) = 1 + 0.2Q_\epsilon(\tau)$, and $Q_\epsilon(\tau)$ is the τth quantile of ϵ. We further assumed that Z was measured with errors in the form of $W = Z + U$, where U is the measurement error and W is the surrogate of Z. We generated the measurement errors from three different distributions, respectively; that is, Laplace: $U \sim L_1(0, 0.5^2)$, normal: $U \sim N(0, 0.5^2)$, and uniform: $U \sim \mathrm{Unif}(-\sqrt{3}/2, \sqrt{3}/2)$. These choices of measurement error distributions correspond to a signal-to-noise ratio of 0.8. The censoring time C, dependent on Z, was generated from $\mathrm{Unif}(c_1, c_2)$ if $Z < \sqrt{12}/2$ and from $\mathrm{Unif}(c_1 + 1, c_2)$ otherwise. For each scenario, c_1 and c_2 were chosen to yield a censoring rate of around 20%. Note that although the proposed method is developed to handle the Laplace or normal measurement errors, we also considered uniform measurement errors to examine the robustness of our approach. We chose the bandwidth $h_n = 1$, while sensitivity analysis with different values of h_n is given at the end of this section. We set the smoothing function $K(\cdot)$ as the standard normal distribution function, and adopted an equally spaced grid over interval $[0.1, 0.78]$ with a step 0.02. The naive estimator was obtained by directly regressing on $\mathbf{W} = (1, W)^{\mathrm{T}}$.

Our proposed estimator, which solves the estimating equation (2.7) coupled with treating the measurement error as Laplace, was obtained through the Newton–Raphson algorithm by taking the naive estimator as the initial value. We set sample size $n = 200$, and simulated 500 replicated data sets under each configuration. Following the convention in quantile regression, we used bootstrap with 200 bootstrap samples to obtain the variances of the parameter estimates.

In Table 1, the column labeled “Est” is the median of the estimates, “SE” is the rescaled (i.e., multiplied by $\Phi^{-1}(0.75)$, where $\Phi(\cdot)$ is the standard normal cumulative distribution function) median absolute deviation of the estimates, which is a robust estimate for the standard error (van der Vaart, 1998, Example 21.11), “ESE” is the average of the bootstrap rescaled median absolute deviation, “CP” is the coverage probability of 95% confidence intervals. With sample size $n = 200$, the proposed estimation method performs reasonably well under the standard normal distribution for the model error ϵ, coupled with three different distributions for the measurement error U. The biases are essentially negligible except for those of the lower quantile levels near 0.2 due to sparse event information observed at the initial follow-up time. The estimated standard errors using the bootstrap method agree well with the sampling standard errors, and the coverage probabilities of 95% confidence intervals are around the nominal level. We specifically point out that the performance is similar in all three measurement error cases, even though strictly speaking, our implementation here is only valid for Laplace measurement errors and serves as an approximation for normal measurement errors.

We also explored different distributions for the model error ϵ; for example, an extreme value distribution with the cumulative distribution function $F_\epsilon(x) = 1 - \exp(-e^x)$ and Student’s t distribution with two degrees of freedom, while keeping the rest of the data generation scheme the same as before. The corresponding simulation results are respectively presented in Tables 2 and 3, from which we can draw similar conclusions. When the sample size is small, some nonconvergent cases might be encountered using the Newton–Raphson algorithm. Often, the nonconvergent issue would disappear with a larger sample size. An alternative solution is to minimize the L_2-norm of the estimating function that would bring the estimating equation value as close to zero as possible. More detailed discussions on numerical issues are given in the Supplementary Material.

To evaluate the overall performance of the proposed method as well as its comparison with the naive method, we present the biases of the estimated quantile intercept and slope coefficients across $\tau \in [0.2, 0.78]$ under different model error and covariate measurement error distributions in Figures 1 and 2, respectively. It can be seen that the proposed method can effectively correct the biases caused by measurement errors, whereas the naive method indeed produces serious biases, especially for the quantile slope coefficients, which are typically of more interest in practice. Moreover, Figures 3 and 4 exhibit the mean squared errors (MSEs) of the estimated quantile intercept and slope coefficients, respectively. The MSEs under the proposed method are much smaller than those under the naive method, which further demonstrates that the proposed method is a viable approach to CQR with measurement errors.

When $\boldsymbol{\Sigma}$ is unknown, it may be estimated from replicated data, for which case Table 4 shows that the proposed method also performs well. Obviously, when the censoring rate becomes heavier, the range of estimable quantile levels shrinks. It is evident from Table 5 that the performance of the proposed method is satisfactory even with a censoring rate of 50%. Furthermore, when the symmetric Laplace measurement error is misspecified as an asymmetric distribution, for example, $U \sim \text{Exp}(1/\lambda) - \lambda$ with $\lambda = 0.5$, the conclusions remains the same.

We investigated the sensitivity of the proposed method to the smoothing parameter h_n when the data were generated from the log-transformed CQR model with heteroscedastic model errors for $\epsilon \sim N(0, 1)$ and covariate measurement errors $U \sim N(0, 0.5^2)$. As shown in Figure 5, the biases and MSEs vary slightly with different values of h_n, demonstrating the estimation stability and, more strikingly, both of them are always much smaller than those from the naive method. We also explored the situation where multiple covariates are subject to measurement errors while others are measured precisely. For normal measurement errors, we conducted simulations to compare our infinite series correction function with the integral correction function proposed by Wang, Stefanski, and Zhu (2012). In addition, we examined the simulation and extrapolation (SIMEX) method in He, Yi, and Xiong (2007) under the AFT model, and the simulation results in the Supplemental Material demonstrate the comparability of our proposed method with other alternatives.

4 Application

As an illustration, we applied the proposed estimation and inference procedure for CQR with measurement errors to a lung cancer study. This data set contains 280 lung cancer patients, whose survival times were recorded with a censoring rate of 64.3%. One of the main objectives of this study was to assess the association of patient survival with certain biomarker expression in the tumor cell cytoplasm. The reading of the biomarker expression was performed by pathologists and could be subjective. As a result, the readings of the biomarker expression for each patient were considered imprecise measurements. To reduce the possible subjectivity of the evaluation, for some patients, two readings of the biomarker expression were provided by two different pathologists (replicates). However, neither of the two measurements of biomarker expression can be considered precise. Our main interest lies in investigating the potential of the biomarker as a new prognostic marker and therapeutic target for lung cancer. Other confounding covariates of interest include tumor histology (there were 61% of patients with adenocarcinoma coded as 1, and 39% squamous cell carcinoma coded as 0), age (ranging from 34 to 90 years with mean 66 years), and sex (52% female coded as 1, 48% male coded as 0). In our analysis, we standardized the patients' ages by subtracting their mean and dividing their standard deviation. Half of the patients in the data set had duplicated readings of the biomarker expression and the averaged value of the two expression readings was considered as the surrogate variable in our analysis. Based on the duplicates, we were able to calculate the variance of the measurement error.

We selected the smoothing parameter h_n through the ten-fold cross-validation procedure proposed in Section 2 and obtained the optimal $h_n = 1.88$. Figure 6 displays the proposed quantile regression estimates of covariate effects and the corresponding 95% pointwise confidence intervals for $\tau \in [0.1, 0.5]$ on the basis of 200 bootstrap samples. As the censoring rate is high, we can only estimate regression quantiles up to $\tau_U = 0.5$. We observe that in general patients with a tumor histology of adenocarcinoma had a significantly better survival rate than those with squamous cell carcinoma, and younger patients could be expected to live longer at a lower risk of death. We did not find any significant covariate effects for patients' sex on their survival for all of the considered regression quantiles. There was no significant effect of the biomarker expression detected on the survival for the regression quantiles that we considered. However, there was a trend that a lower level of biomarker expression tended to be associated with a longer survival time, which nevertheless requires a confirmative study. The naive estimates of covariate effects, ignoring the measurement errors, show large volatilities.

Li and Ryan (2004) proposed a first-order bias correction method for the Cox proportional hazards model with covariate measurement errors. For comparison, we also analyzed the lung cancer data using the method of Li and Ryan (2004) as well as the SIMEX method of He, Yi, and Xiong (2007). The corresponding results are summarized in Table 7, from which we can see that patients with a tumor histology of adenocarcinoma or younger patients could be expected to experience significantly longer survivals whereas patients' sex and biomarker expression were not significantly associated with their survival times. These results in general agree with those drawn by the smoothed and corrected method for CQR. Nevertheless, the Cox model in Li and Ryan (2004) or the AFT model in He, Yi, and Xiong (2007) cannot provide the dynamic covariate effects as the quantile level varies.

For model checking, we consider the cumulative residuals over the precisely measured covariates,

$$\mathcal{T}(\mathbf{w}^{\mathrm{o}},\tau) = n^{-1/2}\sum_{i=1}^{n} I(\mathbf{W}_i^{\mathrm{o}} \leq \mathbf{w}^{\mathrm{o}})\mathcal{M}(\mathcal{O}_i, \widehat{\boldsymbol{\beta}}_n(\tau); h_n).$$

where $\mathbf{W}^{\mathrm{o}} = (\text{Histology, Age, Sex})^{\mathrm{T}}$ represent the error-free covariates excluding the biomarker expression. The null distribution of $\mathcal{T}(\mathbf{w}^{\mathrm{o}},\tau)$ can be approximated by the zero-mean process

$$\mathcal{T}^*(\mathbf{w}^{\mathrm{o}},\tau) = n^{-1/2}\sum_{i=1}^{n} I(\mathbf{W}_i^{\mathrm{o}} \leq \mathbf{w}^{\mathrm{o}})\mathcal{M}(\mathcal{O}_i, \widehat{\boldsymbol{\beta}}_n(\tau); h_n)G_i,$$

where $(G_1,\ldots,G_n)$ were generated independently from the standard normal distribution while fixing the data $\{(X_i,\Delta_i,\mathbf{W}_i), i=1,\ldots,n\}$ at their observed values. The supremum statistic $\sup_{\mathbf{w}^{\mathrm{o}},\tau}|\mathcal{T}(\mathbf{w}^{\mathrm{o}},\tau)|$ can be used to test the overall fit of the CQR model. We generated a large number of, say 1000, realizations from $\sup_{\mathbf{w}^{\mathrm{o}},\tau}|\mathcal{T}^*(\mathbf{w}^{\mathrm{o}},\tau)|$ and obtained its 95th percentile as 1.827. The observed value of $\sup_{\mathbf{w}^{\mathrm{o}},\tau}|\mathcal{T}(\mathbf{w}^{\mathrm{o}},\tau)|$ is 0.394, which indicates that the global linear CQR model fits the lung cancer data well.

5 Remarks

We have proposed a corrected estimating equation approach to CQR models for survival data when covariates are measured with errors. Using a smooth function to approximate the indicator function, the resultant estimating function is smoothed, and thus conventional iterative root-finding procedures, such as the Newton–Raphson algorithm, can be applied (Wang, Stefanski, and Zhu, 2012). We have established the asymptotic consistency and weak convergence of the proposed estimator through modern empirical process techniques. Numerical results show that the proposed method is promising in terms of correcting the bias arising from covariate measurement errors, whereas the naive method typically produces biased estimates.

For variance estimation, we also explored directly using the sandwich-type variance estimator based on the smoothed estimating equation (2.6), but the resulting coverage probability is found to be generally over the nominal level. A more interesting resampling method, known as the Markov chain marginal bootstrap, can be tailored for variance estimation in quantile regression (Kocherginsky, He, and Mu, 2005). When the covariates are of high dimension, particularly for those mismeasured ones, estimation could be difficult, whereas some regularization methods may potentially be incorporated into the estimation procedure to alleviate the instability caused by high dimensionality.

Identifiability is an inherent and subtle issue in CQR. Regression quantiles with τ close to 1 may not be identifiable due to the lack of event information in the upper tail. Theoretically, τ_U should satisfy the identifiability Assumption 4-(iii) in the Appendix. In practical implementation, we first set τ_U to be close to one minus the censoring rate, and then select the final τ_U in an adaptive manner as follows. If all the regression quantiles over $[\nu, \tau_U]$ can be estimated, we increase τ_U by some small step size, e.g., 0.05 or 0.1; otherwise, we decrease τ_U slightly. Through this trial-and-error process, we can push τ_U to the largest value at which the model parameters can be identified. Similarly, ν could also be chosen in such an adaptive way.

The proposed method requires the global linearity assumption as in Portnoy (2003), Peng and Huang (2008) and Wei and Carroll (2009); that is, in order to estimate the τth conditional regression quantile, it is necessary to assume that the conditional functionals at all the lower quantiles are also in the linear form. When the linearity assumption holds only at one specific quantile level τ of interest, research along the work of Wang and Wang (2009) is warranted.

Appendix: Assumptions

Let $\|\cdot\|$ denote the L_2-norm of the corresponding vector or matrix after vectorization. Define $r_0 = \inf\{j \geq 1\colon \int_{-\infty}^{\infty} x^j K^{(1)}(x)\mathrm{d}x \neq 0\}$. Let $S_C(t|\mathbf{z}) = P(C > t|\mathbf{z})$, $F_{X,\Delta=1}(t|\mathbf{z}) = P(X \leq t, \Delta = 1|\mathbf{z})$ and $F_X(t|\mathbf{z}) = P(X \leq t|\mathbf{z})$, and then $F_{X,\Delta=1}(t|\mathbf{z}) = \int_{-\infty}^{t} S_C(u|\mathbf{z})\mathrm{d}F_T(u|\mathbf{z})$ and $F_X(t|\mathbf{z}) = 1 - \{1 - F_T(t|\mathbf{z})\}S_C(t|\mathbf{z})$. Further denote $\mu_0(\mathbf{b}) = E\{\mathbf{Z}N(\mathbf{Z}^{\mathrm{T}}\mathbf{b})\}$, $\mu(\mathbf{b}; h_n) = E\{\Delta\bar{g}(\mathcal{O}, \mathbf{b}; h_n)\}$, $\widetilde{\mu}_0(\mathbf{b}) = E\{\mathbf{Z}I(X \geq \mathbf{Z}^{\mathrm{T}}\mathbf{b})\}$, and $\widetilde{\mu}(\mathbf{b}; h_n) = E\{g(\mathcal{O}, \mathbf{b}; h_n)\}$.

For $d > 0$, define $\mathcal{B}(d) = \{\mathbf{b} \in \mathbb{R}^p\colon \inf_{\tau \in (0,\tau_U]} \|\mu_0(\mathbf{b}) - \mu_0\{\boldsymbol{\beta}_0(\tau)\}\| \leq d\}$. Assume that there exists $d_0 > 0$ such that $\mathcal{B}(d_0)$ contains $\{\boldsymbol{\beta}_0(\tau)\colon \tau \in (0, \tau_U]\}$. Denote $f_{X,\Delta=1}(t|\mathbf{z})$ and $f_X(t|\mathbf{z})$ as the density functions corresponding to $F_{X,\Delta=1}(t|\mathbf{z})$ and $F_X(t|\mathbf{z})$, respectively. Let $\mathbf{B}_0(\mathbf{b}) = E\{\mathbf{Z}\mathbf{Z}^{\mathrm{T}} f_{X,\Delta=1}(\mathbf{Z}^{\mathrm{T}}\mathbf{b}|\mathbf{Z})\}$ and $\mathbf{J}_0(\mathbf{b}) = -E\{\mathbf{Z}\mathbf{Z}^{\mathrm{T}} f_X(\mathbf{Z}^{\mathrm{T}}\mathbf{b}|\mathbf{Z})\}$. Finally, denote $\dot{g}\{\mathcal{O}, \boldsymbol{\beta}(\tau); h_n\} = \partial g\{\mathcal{O}, \boldsymbol{\beta}(\tau); h_n\}/\partial\boldsymbol{\beta}(\tau)$.

Assumption 1. The smoothing function $K(\cdot)$ satisfies:

(i) $K^{(j)}(\cdot)$ is uniformly bounded for $j = 0, \ldots, 4$ in the Laplace measurement error model and for $j \geq 0$ in the normal measurement error model.

(ii) $r_0 \geq 2$ and for each integer j $(0 \leq j \leq r_0)$, $\int_{-\infty}^{\infty} |x^j K^{(1)}(x)|\mathrm{d}x < \infty$.

(iii) For each integer j $(0 \leq j \leq r_0)$, any $\eta > 0$, and any sequence h_n converging to 0, $\lim_{n\to\infty} h_n^{j-r_0} \int_{|h_n x|>\eta} |x^j K^{(1)}(x)|\mathrm{d}x = 0$ and $\lim_{n\to\infty} h_n^{-1} \int_{|h_n x|>\eta} |K^{(2)}(x)|\mathrm{d}x = 0$.

Assumption 2. For each integer j such that $1 \leq j \leq r_0$, $F_T^{(j)}(t|\mathbf{z})$ is a continuous function of $\mathbf{z}$ and uniformly bounded over t and $\mathbf{z}$, where $F_T^{(j)}(t|\mathbf{z})$ is the jth derivative of $F_T(t|\mathbf{z})$ with respect to t. So is $S_C^{(j)}(t|\mathbf{z})$.

Assumption 3. Each component of $\mu_0\{\boldsymbol{\beta}(\tau)\}$, $\mu\{\boldsymbol{\beta}(\tau); h_n\}$, $\widetilde{\mu}_0\{\boldsymbol{\beta}(\tau)\}$, and $\widetilde{\mu}\{\boldsymbol{\beta}(\tau); h_n\}$ as a function of τ is Lipschitz continuous.

Assumption 4. Boundedness conditions are imposed:

(i) Both $E(\|\mathbf{Z}\|^2)$ and $E(\|\mathbf{U}\|^2)$ are bounded, and $E(\mathbf{Z}\mathbf{Z}^{\mathrm{T}})$ is positive definite.

(ii) $f_{X,\Delta=1}(\mathbf{Z}^{\mathrm{T}}\mathbf{b}|\mathbf{Z})$ is bounded away from zero for all $\mathbf{b}$ in $\mathcal{B}(d_0)$.

(iii) $\inf_{\tau\in[\nu,\tau_U]} \mathrm{eigmin}[\mathbf{B}_0\{\boldsymbol{\beta}_0(\tau)\}] > 0$ for any $0 < \nu \leq \tau_U$, where eigmin$(\cdot)$ denotes the minimum eigenvalue of a matrix.

(iv) The norm of $\mathbf{J}_0(\mathbf{b})\mathbf{B}_0(\mathbf{b})^{-1}$ is uniformly bounded for all $\mathbf{b}$ in $\mathcal{B}(d_0)$.

Assumption 5. We assume that

(i) $\sup_{\tau\in[\nu,\tau_U]} \|E[\dot{g}\{\mathcal{O},\boldsymbol{\beta}_0(\tau);h_n\}]^2\|$ is bounded as $n\to\infty$ for any $\nu\in(0,\tau_U]$.

(ii) $E[\dot{g}\{\mathcal{O},\boldsymbol{\beta}(\tau);h_n\}]^2$ is component-wise continuous in sup-norm as a functional of $\boldsymbol{\beta}(\tau)$.

Assumption 1 holds if $K(\cdot)$ is the standard normal cumulative distribution function. Apparently, $F_{X,\Delta=1}(t|\mathbf{z})$ and $F_X(t|\mathbf{z})$ satisfy a similar boundedness condition as Assumption 2, which is a standard assumption in survival analysis. Assumptions 3 and 4 are commonly used in CQR models (Peng and Huang, 2009). Assumption 5 essentially imposes a higher convergence rate of the smoothing parameter h_n compared with Assumption 1-(iii). If we take $K(\cdot)$ to be the standard normal cumulative distribution function, Assumption 5 holds for both the Laplace and the normal measurement errors. This is because the exponential part can be factored out for any order of derivatives with respect to $K(\cdot)$ and the component-wise continuity follows directly.

References

Bang, H., and Tsiatis, A. A. (2002), "Median Regression With Censored Cost Data," *Biometrics*, 58, 643–649.

Brown, M. L. (1982), "Robust Line Estimation With Error in Both Variables," *Journal of the American Statistical Association*, 77, 71–79. Correction in 78, 1008.

Buchinsky, M., and Hahn, J. Y. (1998), "An Alternative Estimator for Censored Quantile Regression," *Econometrica*, 66, 653–671.

Buckley, J., and James, I. (1979), "Linear Regression With Censored Data," *Biometrika*, 66, 429–436.

Carroll, R. J., Ruppert, D., Stefanski, L. A., and Crainiceanu, C. (2006), *Measurement Error in Nonlinear Models: A Modern Perspective*, London: CRC Press.

Chernozhukov, V., and Hong, H. (2002), "Three-Step Censored Quantile Regression and Extramarital Affairs," *Journal of the American Statistical Association*, 97, 872–882.

Cox, D. R. (1972), "Regression Models and Life Tables" (with discussion), *Journal of Royal Statistical Society*, Ser. B, 34, 187–220.

Fitzenberger, B. (1997), "A Guide to Censored Quantile Regressions," in *Handbook of Statistics, Robust Inference*, eds. Maddala G. S., and Rao C. R., Amsterdam: North-Holland.

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

He, X., and Liang, H. (2000), "Quantile Regression Estimates for a Class of Linear and Partially Linear Errors-in-Variables Models," *Statistica Sinica*, 10, 129–140.

He, W., Yi, G. Y., and Xiong, J. (2007), "Accelerated Failure Time Models With Covariates Subject to Measurement Error," *Statistics in Medicine*, 26, 4817–4832.

Hong, H., and Tamer, E. (2003), "A Simple Estimator for Nonlinear Error in Variable Models," *Journal of Econometrics*, 117, 1–19.

Horowitz, J. (1992), "A Smoothed Maximum Score Estimator for the Binary Response Model," *Econometrica*, 60, 505–531.

Horowitz, J. (1998), "Bootstrap Methods for Median Regression Model," *Econometrica*, 66, 1327–1352.

Jin, Z., Lin, D. Y., Wei, L. J., and Ying, Z. (2003), "Rank-Based Inference for the Accelerated Failure Time Model," *Biometrika*, 90, 341–353.

Khan, S., and Powell, L. J. (2001), "Two-Step Estimation of Semiparametric Censored Regression Models," *Journal of Econometrics*, 103, 73–110.

Kocherginsky, M., He, X., and Mu, Y. (2005), "Practical Confidence Intervals for Regression Quantiles," *Journal of Computational and Graphical Statistics*, 14, 41–55.

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

Koenker, R., and Bassett, G. J. (1978), "Regression Quantiles," *Econometrica*, 46, 33–50.

Koenker, R., and Geling, O. (2001), "Reappraising Medfly Longevity: A Quantile Regression Survival Analysis," *Journal of the American Statistical Association*, 96, 458–468.

Kotz, S., Kozubowski, T. J., and Podgorski, K. (2001), *The Laplace Distribution and Generalizations*, New York: Springer.

Lai, T. L., and Ying, Z. (1991), "Rank Regression Methods for Left-Truncated and Right-Censored Data," *The Annals of Statistics*, 19, 531–556.

Li, Y., and Ryan, L. (2004), "Survival Analysis With Heterogeneous Covariate Measurement Error," *Journal of the American Statistical Association*, 99, 724–735.

Lindgren, A. (1997), "Quantile Regression With Censored Data Using Generalized L_1 Minimization," *Computational Statistics and Data Analysis*, 23, 509–524.

Ma, Y., and Yin, G. (2011), "Censored Quantile Regression With Covariate Measurement Errors," *Statistica Sinica*, 21, 949–971.

Nakamura, T. (1990), "Corrected Score Function for Errors-in-Variables Models: Methodology and Application to Generalized Linear Models," *Biometrika*, 77, 127–137.

Peng, L., and Huang, Y. (2008), "Survival Analysis With Quantile Regression Models," *Journal of the American Statistical Association*, 103, 637–649.

Portnoy, S. (2003), "Censored Regression Quantiles," *Journal of the American Statistical Association*, 98, 1001–1012.

Powell, J. L. (1984), "Least Absolute Deviations Estimation for the Censored Regression," *Journal of Econometrics*, 25, 303–325.

Prentice, R. L. (1978), "Linear Rank Tests With Right Censored Data," *Biometrika*, 65, 167–180.

Ritov, Y. (1990), "Estimation in a Linear Regression Model With Censored Data," *The Annals of Statistics*, 18, 303–328.

Stefanski, L. A. (1989), "Unbiased Estimation of a Nonlinear Function of a Normal-Mean With Application to Measurement Error Models," *Communications in Statistics-Theory and Methods*, 18, 4335–4358.

Tsiatis, A. A. (1990), "Estimating Regression Parameters Using Linear Rank Tests for Censored Data," *The Annals of Statistics*, 18, 354–372.

van der Vaart, A. W. (1998), *Asymptotic Statistics*, New York: Cambridge University Press.

Wang, H., and Wang, L. (2009), "Locally Weighted Censored Quantile Regression," *Journal of the American Statistical Association*, 104, 1117–1128.

Wang, H., Stefanski, L. A., and Zhu, Z. (2012), "Corrected-Loss Estimation for Quantile Regression With Covariate Measurement Errors," *Biometrika*, 99, 405–421.

Wei, L. J., Ying, Z., and Lin, D. Y. (1990), "Linear Regression Analysis of Censored Survival Data Based on Rank Tests," *Biometrika*, 77, 845–851.

Wei, Y., and Carroll, R. J. (2009), "Quantile Regression With Measurement Error," *Journal of the American Statistical Association*, 104, 1129–1143.

Yang, S. (1999), "Censored Median Regression Using Weighted Empirical Survival and Hazard Function," *Journal of the American Statistical Association*, 94, 137–145.

Ying, Z., Jung, S. H., and Wei, L. J. (1995), "Survival Analysis With Median Regression Models," *Journal of the American Statistical Association*, 90, 178–184.

Table 1: Simulation results for the log-transformed censored quantile regression with three different distributions of covariate measurement errors and heteroscedastic model errors for $\epsilon \sim N(0,1)$

	$\beta_0(\cdot)$				$\beta_1(\cdot)$			
τ	True	Est	SE/ESE	CP(%)	True	Est	SE/ESE	CP(%)
				$U \sim L_1(0, 0.5^2)$				
0.2	−1.342	−1.245	0.934	91.8	0.832	0.811	1.028	93.4
0.3	−1.024	−0.976	0.953	94.0	0.895	0.881	1.014	92.8
0.4	−0.753	−0.740	1.022	95.6	0.949	0.944	1.021	94.2
0.5	−0.500	−0.509	1.030	95.8	1.000	1.006	1.028	93.2
0.6	−0.247	−0.269	0.996	95.6	1.051	1.080	0.974	93.4
0.7	0.024	−0.001	1.040	95.0	1.105	1.117	1.038	93.2
				$U \sim N(0, 0.5^2)$				
0.2	−1.342	−1.265	0.967	92.8	0.832	0.830	0.993	94.0
0.3	−1.024	−0.975	0.943	92.0	0.895	0.885	0.932	93.6
0.4	−0.753	−0.715	0.928	93.8	0.949	0.936	0.925	93.8
0.5	−0.500	−0.465	0.879	93.0	1.000	0.993	0.935	94.2
0.6	−0.247	−0.222	0.886	94.2	1.051	1.053	0.972	93.6
0.7	0.024	0.044	0.913	94.6	1.105	1.099	0.942	92.8
				$U \sim \text{Unif}(-\sqrt{3}/2, \sqrt{3}/2)$				
0.2	−1.342	−1.260	0.987	92.6	0.832	0.817	0.891	94.8
0.3	−1.024	−0.971	0.969	93.6	0.895	0.878	0.894	94.8
0.4	−0.753	−0.720	0.951	93.8	0.949	0.934	0.962	92.8
0.5	−0.500	−0.476	0.942	94.2	1.000	0.990	0.986	92.6
0.6	−0.247	−0.222	0.969	93.8	1.051	1.050	1.042	93.6
0.7	0.024	0.060	1.004	94.2	1.105	1.102	1.052	92.6

Note: SE/ESE is the ratio of the sampling standard error and the estimated (bootstrap) standard error, and CP is the coverage probability.

Table 2: Simulation results for the log-transformed censored quantile regression with three different distributions of covariate measurement errors and heteroscedastic model errors for ϵ from an extreme value distribution

	$\beta_0(\cdot)$				$\beta_1(\cdot)$			
τ	True	Est	SE/ESE	CP(%)	True	Est	SE/ESE	CP(%)
				$U \sim L_1(0, 0.5^2)$				
0.2	−2.000	−1.822	1.030	85.4	0.700	0.679	0.929	93.4
0.3	−1.531	−1.452	1.000	90.8	0.794	0.770	0.934	93.4
0.4	−1.172	−1.161	0.965	92.4	0.866	0.870	0.994	93.2
0.5	−0.867	−0.918	0.993	92.2	0.927	0.954	1.030	92.4
0.6	−0.587	−0.666	1.008	91.2	0.983	1.033	0.994	90.2
0.7	−0.314	−0.384	0.996	91.8	1.037	1.100	0.981	92.4
				$U \sim N(0, 0.5^2)$				
0.2	−2.000	−1.823	1.008	85.6	0.700	0.680	1.042	94.0
0.3	−1.531	−1.453	1.026	91.0	0.794	0.778	1.050	94.4
0.4	−1.172	−1.154	0.957	93.8	0.866	0.863	1.018	94.6
0.5	−0.867	−0.895	0.950	94.2	0.927	0.946	0.957	94.6
0.6	−0.587	−0.635	0.923	93.2	0.983	1.018	1.013	93.0
0.7	−0.314	−0.349	0.876	95.2	1.037	1.066	0.974	94.8
				$U \sim \mathrm{Unif}(-\sqrt{3}/2, \sqrt{3}/2)$				
0.2	−2.000	−1.771	1.008	85.4	0.700	0.657	0.933	94.4
0.3	−1.531	−1.418	1.023	90.4	0.794	0.749	0.911	94.8
0.4	−1.172	−1.119	1.018	91.6	0.866	0.847	0.959	93.8
0.5	−0.867	−0.854	1.073	93.2	0.927	0.926	1.031	92.8
0.6	−0.587	−0.601	1.061	92.2	0.983	1.012	1.051	91.8
0.7	−0.314	−0.314	1.080	91.8	1.037	1.073	1.032	91.8

Table 3: Simulation results for the log-transformed censored quantile regression with three different distributions of covariate measurement errors and heteroscedastic model errors for $\epsilon \sim t_2$

	$\beta_0(\cdot)$				$\beta_1(\cdot)$			
τ	True	Est	SE/ESE	CP(%)	True	Est	SE/ESE	CP(%)
				$U \sim L_1(0, 0.5^2)$				
0.2	−1.561	−1.453	0.974	90.6	0.788	0.788	0.923	96.6
0.3	−1.117	−1.101	0.973	95.8	0.877	0.881	0.960	95.8
0.4	−0.789	−0.797	0.978	96.0	0.942	0.947	0.916	96.4
0.5	−0.500	−0.512	0.960	96.0	1.000	1.020	0.898	95.2
0.6	−0.211	−0.223	0.955	94.2	1.058	1.085	0.872	93.8
0.7	0.117	0.116	0.991	95.0	1.123	1.122	0.920	93.8
				$U \sim N(0, 0.5^2)$				
0.2	−1.561	−1.452	0.975	92.0	0.788	0.810	0.908	95.0
0.3	−1.117	−1.082	0.973	95.0	0.877	0.885	0.935	94.6
0.4	−0.789	−0.763	1.026	95.2	0.942	0.947	0.981	95.8
0.5	−0.500	−0.477	1.041	95.0	1.000	1.006	1.025	94.6
0.6	−0.211	−0.193	1.036	95.0	1.058	1.073	1.030	94.8
0.7	0.117	0.161	0.967	94.8	1.123	1.114	0.967	93.6
				$U \sim \text{Unif}(-\sqrt{3}/2, \sqrt{3}/2)$				
0.2	−1.561	−1.490	0.937	90.6	0.788	0.810	0.920	94.2
0.3	−1.117	−1.109	1.050	93.2	0.877	0.889	0.918	93.2
0.4	−0.789	−0.789	1.026	93.0	0.942	0.943	0.962	93.8
0.5	−0.500	−0.478	0.985	93.4	1.000	1.001	0.981	93.0
0.6	−0.211	−0.179	1.004	93.4	1.058	1.059	1.018	92.6
0.7	0.117	0.156	0.946	92.0	1.123	1.131	1.033	91.4

Table 4: Simulation results for the log-transformed censored quantile regression with the heteroscedastic model error for $\epsilon \sim N(0,1)$ and covariate measurement error $U \sim N(0, 0.5^2)$, where the standard error of the measurement error is estimated based on five replications

	$\beta_0(\cdot)$				$\beta_1(\cdot)$			
τ	True	Est	SE/ESE	CP(%)	True	Est	SE/ESE	CP(%)
0.2	−1.342	−1.271	1.025	93.4	0.832	0.822	1.000	94.2
0.3	−1.024	−0.989	1.026	94.8	0.895	0.892	0.993	95.2
0.4	−0.753	−0.747	0.973	95.0	0.949	0.952	1.007	95.8
0.5	−0.500	−0.509	0.987	94.2	1.000	1.013	0.993	95.8
0.6	−0.247	−0.272	0.965	94.6	1.051	1.072	0.938	94.8
0.7	0.024	0.015	0.963	93.6	1.105	1.117	0.955	93.6

Table 5: Simulation results for the log-transformed censored quantile regression under a censoring rate of 50%, with the heteroscedastic model error for $\epsilon \sim N(0,1)$ and covariate measurement error $U \sim N(0, 0.5^2)$

	$\beta_0(\cdot)$				$\beta_1(\cdot)$			
τ	True	Est	SE/ESE	CP(%)	True	Est	SE/ESE	CP(%)
0.2	−1.342	−1.349	1.044	94.4	0.832	0.849	0.994	93.8
0.3	−1.024	−1.072	0.992	92.6	0.895	0.943	1.000	93.0
0.4	−0.753	−0.788	1.004	92.8	0.949	0.998	0.975	94.2
0.5	−0.500	−0.483	0.963	93.2	1.000	0.988	0.890	92.6

Table 6: Simulation results for the log-transformed censored quantile regression with the heteroscedastic model error for $\epsilon \sim N(0,1)$ and asymmetric covariate measurement error $U \sim \text{Exp}(1/\lambda) - \lambda$, where $\lambda = 0.5$

	$\beta_0(\cdot)$				$\beta_1(\cdot)$			
τ	True	Est	SE/ESE	CP(%)	True	Est	SE/ESE	CP(%)
0.2	−1.342	−1.197	0.962	88.2	0.832	0.793	1.021	92.8
0.3	−1.024	−0.946	0.938	92.6	0.895	0.872	0.979	95.4
0.4	−0.753	−0.719	1.013	94.0	0.949	0.953	0.972	94.6
0.5	−0.500	−0.508	1.000	94.8	1.000	1.038	0.926	93.2
0.6	−0.247	−0.304	0.969	93.4	1.051	1.124	0.961	91.4
0.7	0.024	−0.083	0.987	91.4	1.105	1.208	1.019	88.6

Table 7: Analysis results of the lung cancer data using the first-order bias correction method for the Cox proportional hazards model (Li and Ryan, 2004) and the simulation-extrapolation method for the accelerated failure time (AFT) model (He, Yi, and Xiong, 2007)

Model	Error	Covariate	Est	ESE	p-value
Cox	–	Histology	-0.503	0.219	0.022
		Age	0.433	0.118	< 0.001
		Sex	-0.081	0.209	0.699
		Biomarker	0.043	0.183	0.813
AFT	Normal	Intercept	1.275	0.614	0.038
		Histology	0.680	0.325	0.036
		Age	-0.500	0.159	0.002
		Sex	0.210	0.303	0.487
		Biomarker	0.178	0.300	0.552
	Extreme	Intercept	2.108	0.501	< 0.001
		Histology	0.569	0.262	0.030
		Age	-0.490	0.131	< 0.001
		Sex	0.100	0.238	0.673
		Biomarker	-0.070	0.246	0.776

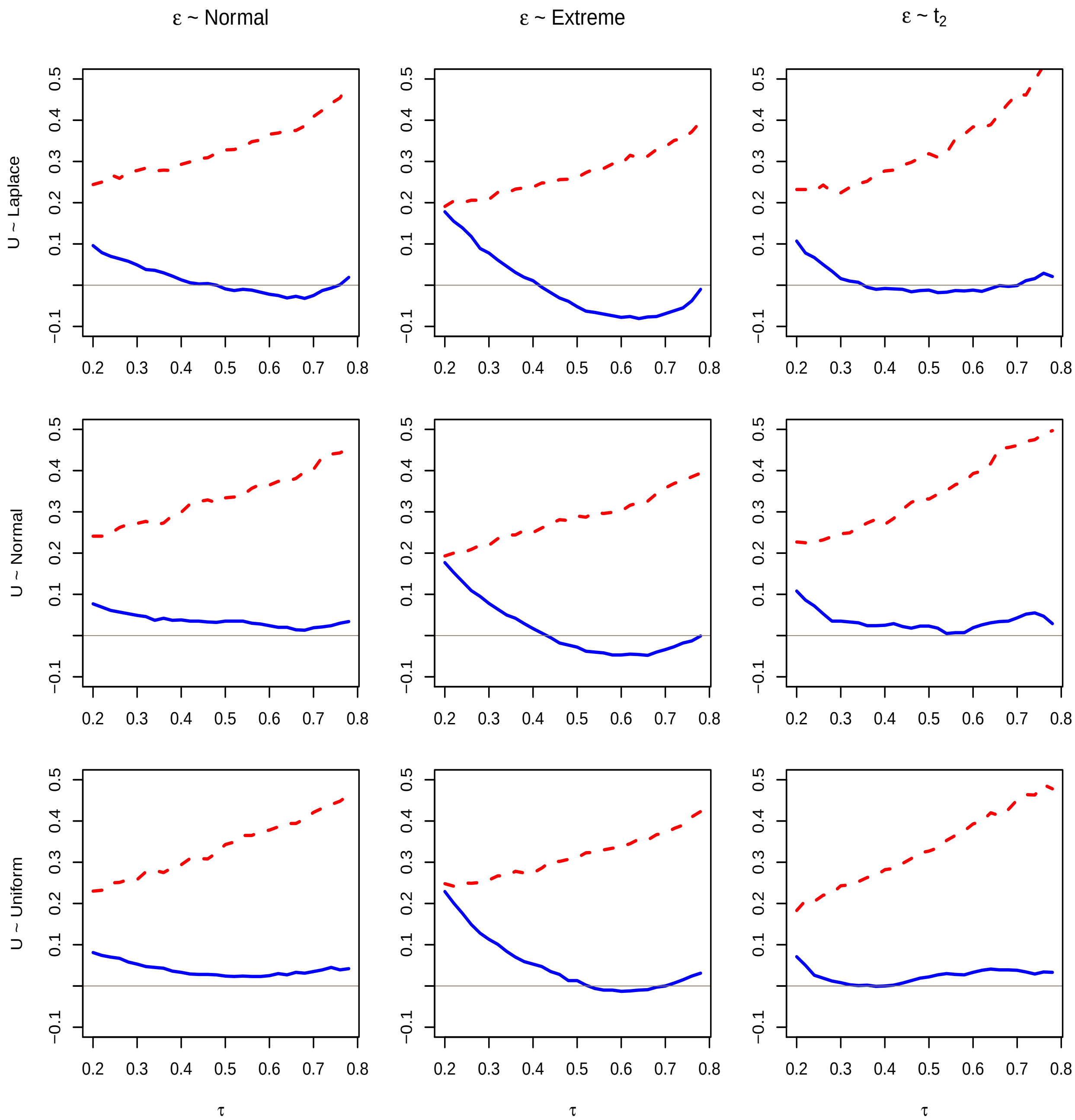

Figure 1: Biases of the estimated quantile regression intercept using the proposed method (solid lines) and the naive method (dashed lines) under three different model error distributions: normal, extreme value and t_2, and three different measurement error distributions: Laplace, normal and uniform, respectively.

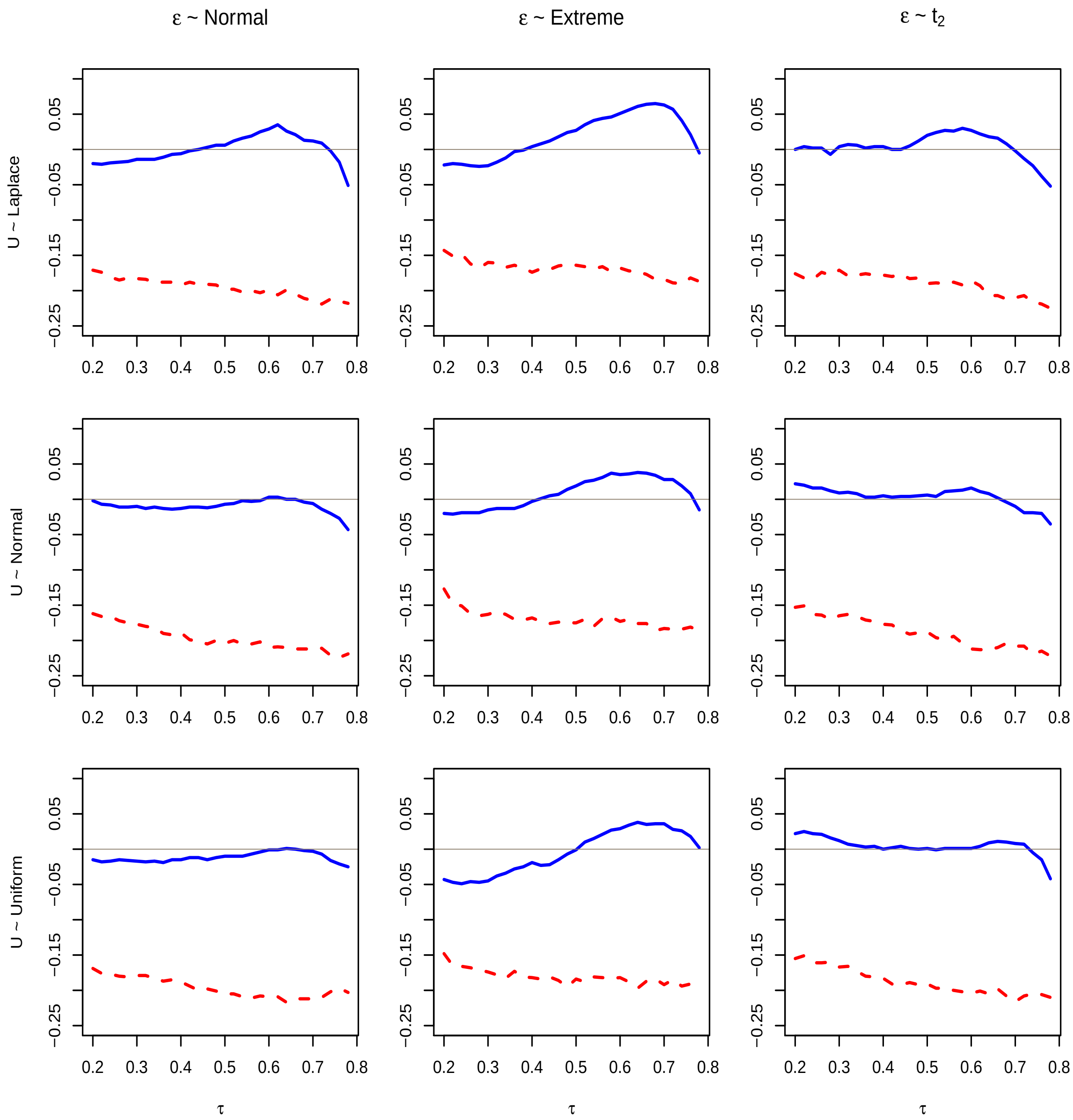

Figure 2: Biases of the estimated quantile regression slope using the proposed method (solid lines) and the naive method (dashed lines) under three different model error distributions: normal, extreme value and t_2, and three different measurement error distributions: Laplace, normal and uniform, respectively.

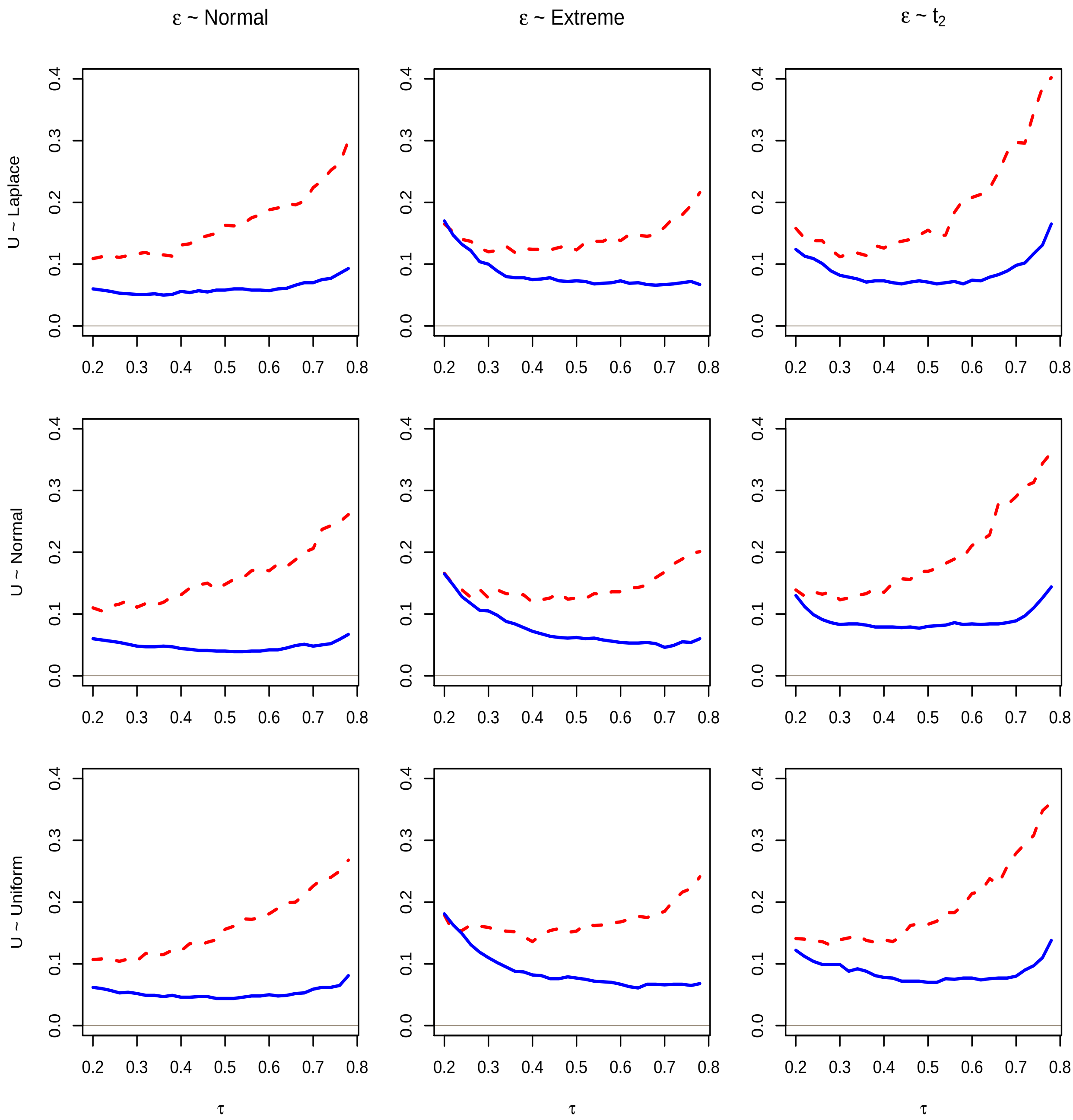

Figure 3: Mean squared errors of the estimated quantile regression intercept using the proposed method (solid lines) and the naive method (dashed lines) under three different model error distributions: normal, extreme value and t_2, and three different measurement error distributions: Laplace, normal and uniform, respectively.

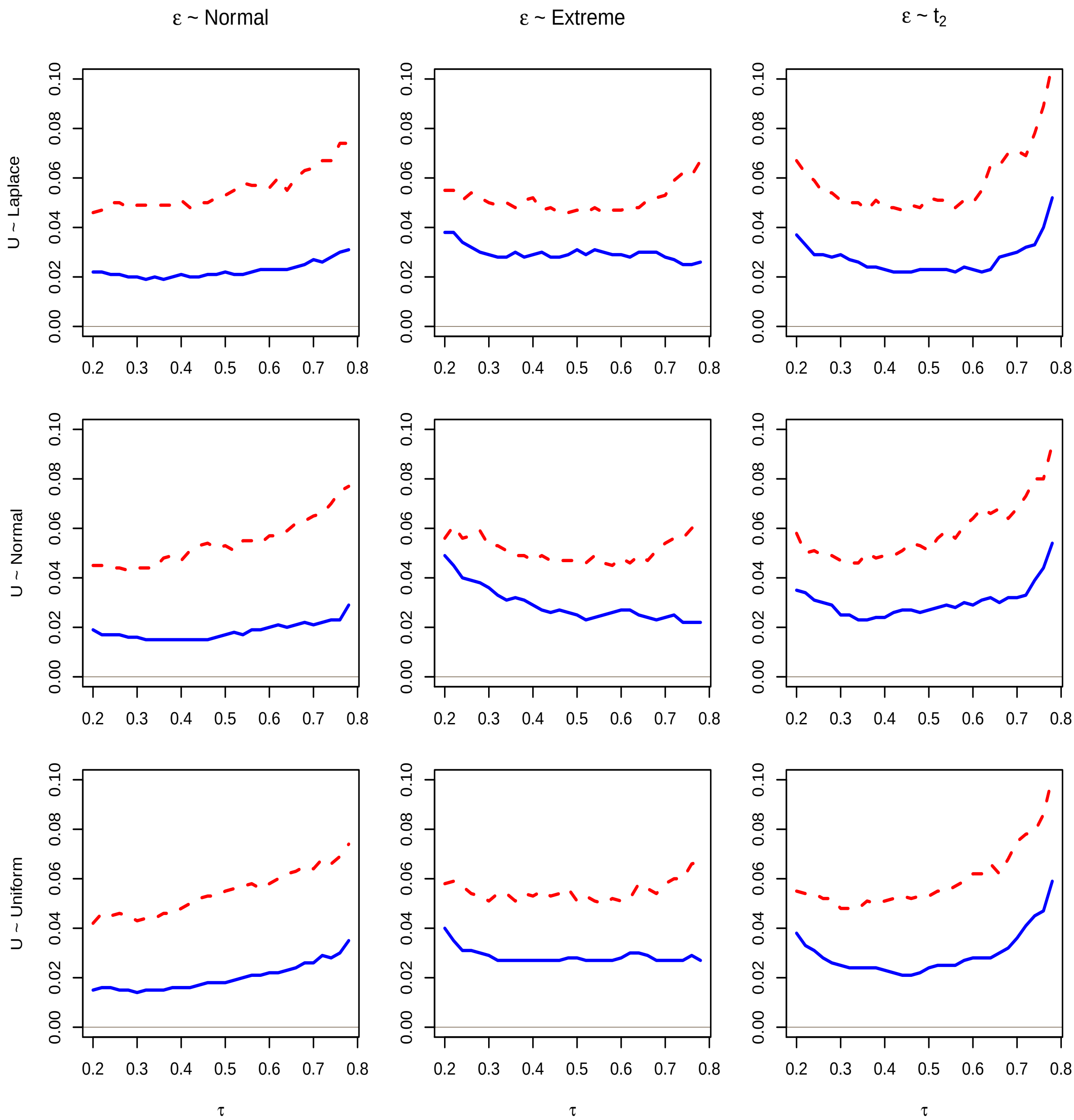

Figure 4: Mean squared errors of the estimated quantile regression slope using the proposed method (solid lines) and the naive method (dashed lines) under three different model error distributions: normal, extreme value and t_2, and three different measurement error distributions: Laplace, normal and uniform, respectively.

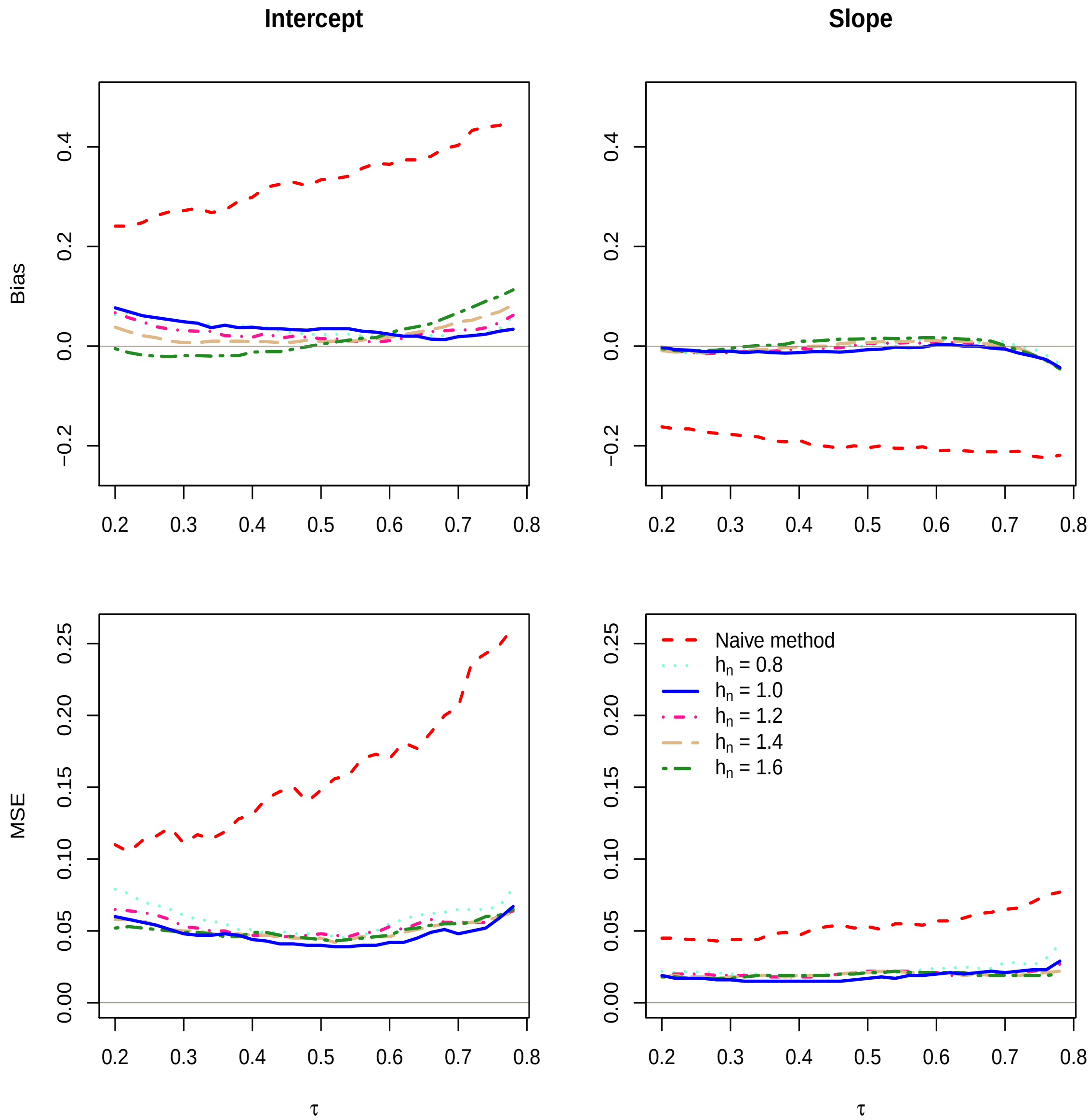

Figure 5: Comparison of biases and mean squared errors (MSEs) of the estimated quantile regression intercept (left panel) and slope (right panel) for the proposed method with different h_n and the naive method under the normal measurement error and normal model error.

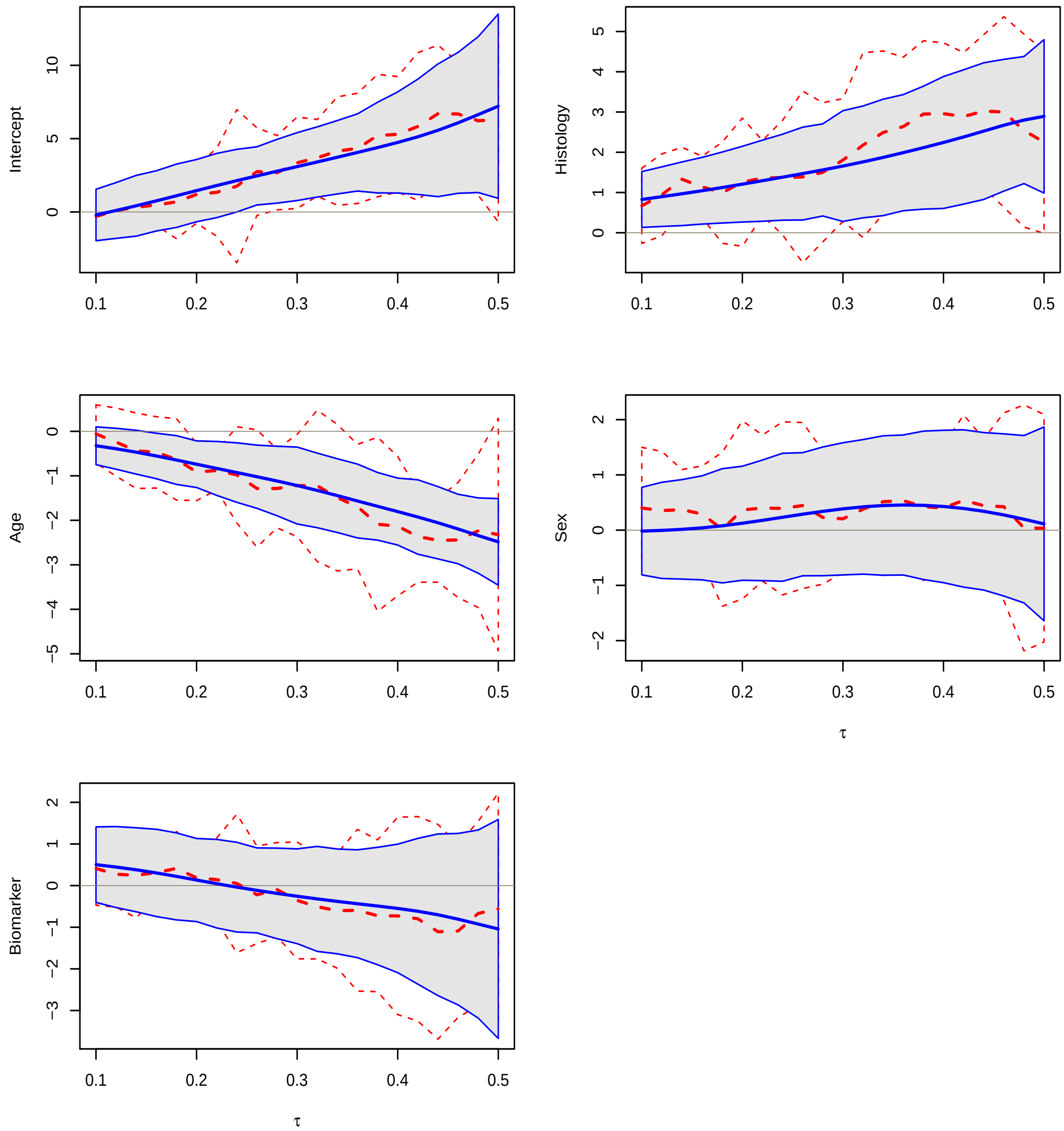

Figure 6: Estimated covariate effects and the corresponding 95% pointwise confidence intervals under the proposed measurement error quantile regression (solid lines) and the naive method (dashed lines) for the lung cancer data.