Journal of Statistical Research *ISSN 0256 - 422 X*
2013, Vol. 47, No. 1, pp. 51-61

QUANTILE REGRESSION MODELS WITH PARTIALLY FUNCTIONAL EFFECTS FOR RANDOMLY RIGHT CENSORED DATA: A SIMULATION STUDY

AMENA SULTANA
Centre for Equity and Health Systems (CEHS), ICDDR,B, Dhaka 1212, Bangladesh
Email: amena@icddrb.org

JAHIDA GULSHAN
Institute of Statistical Research and Training (ISRT)
University of Dhaka, Dhaka 1000, Bangladesh
Email: gulshan@isrt.ac.bd

SUMMARY

Quantile regression presents a flexible approach to the analysis of survival data, allowing for modeling quantile-specific covariate effect. Qian and Peng (2010) proposed profile estimating equations and a readily and stably implemented iterative algorithm for censored quantile regression tailored to the partially functional effect setting with a mixture of varying and constant effects and demonstrated improved efficiency of estimation over a naive two stage procedure. The aim of this study is to use the same algorithm on a quantile regression setting where some covariate effects follow general parametric pattern (e.g. normal, gamma or logistic distribution) rather than a constant function or value and to determine the strength of using the algorithm in such regression settings through simulation. Simulation studies demonstrate that the method works well, for moderately censored data, if the parametric pattern $g(.)$ is a known function with unknown parameter(s). A sensitivity analysis is performed to check the consequences of misspecification of such parametric pattern.

Keywords and phrases: Quantile regression; Censored data; Parametric pattern; Efficiency; Sensitivity analysis.

1 Introduction

Regression quantiles, a new class of statistics is a simple minimization problem yielding the ordinary sample quantiles in the location model (Koenker and Bassett, 1978). Quantile regression, first proposed by Koenker and Bassett (1978), has emerged as a significant extension of classic linear regression by seminally using the concept of conditional quantiles. Quantile regression has great flexibility and straightforward interpretation in assessing covariate effects on event times, resulting in growing interests in its applications in survival analysis. The quantile regression is more flexible because the effect of covariates is not restricted to be constant in contrast to the accelerated failure time model (Prentice, 1978; Buckley and James, 1979; Louis, 1981; Wei and Gail, 1983; Ritov, 1990; Tsiatis, 1990; Wei, Ying, and Lin, 1990). Quantile regression has become a useful complement to the use of the Cox model (Qian and Peng, 2010).

A quantile regression model for an event time of interest, Y (possibly transformed), and a covariate vector X including a constant term, can be expressed as,

$$Q_Y(r \mid X) = X^T\beta(r), \qquad r \in (0,1), \tag{1.1}$$

where $Q_Y(r \mid X) \equiv \inf\{y : pr(Y \leq y \mid X) \geq r\}$ denotes the rth conditional quantile of Y given X, and $\beta(r)$ represents the effect of X on $Q_Y(r \mid X)$.

In quantile regression analyses, location-shift effects may often be found adequate for a subset of covariates. For small samples, formulating all effects as functional (Peng and Huang, 2008) may offer unnecessary flexibility at the price of efficiency. For such, Qian and Peng (2010) adopted a quantile regression model with a mixture of varying and constant effects, referred to as partially functional effects that provides a simpler view of covariate effects, and may thus be preferred over its fully functional counterpart.

A few of the past studies on quantile regression with survival data centered around dealing with censoring. Approaches of Powell (1984, 1986), Ying et al. (1995), Fitzenberger (1997), Buchinsky and Hahn (1998), Yang (1999), Chernozhukov and Hong (2002), and Honore et al. (2002) assumed known censoring time or unconditional independence between T and C. Portnoy (2003) and Neocleous et al. (2006) studied the estimation of quantile regression model by adopting the self-consistency principle for the Kaplan-Meier estimate (Efron, 1967) under a weaker random censoring assumption that C is independent of T conditional on X. Later, Peng and Huang (2008) developed a new approach by utilizing the martingale feature associated with censored data, which greatly facilitates both large sample properties and inferential procedures. Qian and Peng (2010) used the same approach to represent censored quantile regression tailored to the partially functional effect setting with a mixture of varying and constant effects. They have shown that, such models can offer a simpler view regarding covariate-survival association with greater efficiency.

In this paper, the estimation technique of Qian and Peng (2010) has been applied on quantile regression setting where some covariate effects follow a more general parametric pattern instead of a constant function. For this purpose, normal, extreme value and logistic distributions are used to present covariate effects. This generalization is straightforward and would be useful in practical settings where strong preliminary information on the evolving paths of quantile effects is available. Monte Carlo simulation is used to check the performance of the methodology under this scenario and a sensitivity analysis has been performed to examine the consequences of misspecification of parametric patterns.

2 Methods

Let T and C denote the survival time of interest and the censoring time, respectively. Define $V = \min(T,C)$ and $\delta = I(T \leq C)$,where $I(\cdot)$ is the indicator function. Without loss of generality, we let $Y = \log T$ and partition X as $(W^T, Z^T)^T$,where W is a $p \times 1$ vector of covariates with r-varying effects, and Z is a $q \times 1$ vector of covariates with constant effects. The observed data consist of n independent and identically distributed replicates of (V, δ, W, Z), denoted by $\{(V_i, \delta_i, W_i, Z_i), i = 1, \ldots, n\}$.

Then, a partially functional quantile regression model will take the form

$$Q_Y(r \mid W, Z) = W^T \beta(r) + Z^T \gamma, \qquad r \in (0, 1), \tag{2.1}$$

where $\beta(r)$ is the vector of functional regression coefficients for W and γ is the vector of constant coefficients for Z.

Consider, $F_T(t \mid W, Z) = pr(T \le t \mid W, Z)$ and $\Lambda_T(t \mid W, Z) = -\log\{1 - F_T(t \mid W, Z)\}$. Here $\Lambda_T(t \mid W, Z)$ is the cumulative hazard function of T conditional on W and Z. Let the counting process be $N(t) = I(V \le t, \delta = 1), Y(t) = I(V \ge t)$, the martingale process be $M(t) = N(t) - \Lambda_T(t \wedge V \mid W, Z)$ and $\{N_i(t), Y_i(t), M_i(t)\}(i = 1, \ldots, n)$ be the sample analogues of $\{N(t), Y(t), M(t)\}$. Also, $N_i = N_i(r)$, that is, $N_i = 1$ if the ith individual had the event at or prior to time r, and 0, otherwise.

Following the arguments in Peng and Huang (2008) and Qian and Peng (2010), the following equality holds for $r \in (0, 1)$:

$$E\Big\{\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \omega_i \Big(N_i \exp\big(W_i^T \beta_0(r) + Z_i^T \gamma_0\big) - \int_0^r I\big[V_i \ge \exp\big(W_i^T \beta_0(u) + Z_i^T \gamma_0\big)\big] dH(u)\Big)\Big\} = 0,$$

where ω_i represents either W_i (if covariate effects are varying) or Z_i (if covariate effects are constant) and $\beta_0(\cdot)$ and γ_0 are the true values of $\beta(\cdot)$ and γ in model (2.1) and $H(v) = -\log(1 - v)$ for $0 \le v \le 1$. Due to identifiability problem posed by censoring, let us restrict $\beta_0(\cdot)$ to $\{\beta_0(r) : r \in (0, r_U]\}$, where r_U is a constant in (0, 1) assumed to meet certain theoretical constraints (Qian and Peng, 2010).

To avoid potential efficiency loss by better utilization of the information on partially functional effect patterns Qian and Peng (2010) proposed a set of profile estimating equations. The concept was to estimate $\beta_0(r)$ keeping γ_0 as a constant nuisance while estimating γ_0 based on all martingale structured information in the estimation of $\beta_0(r)$. The estimating equations are,

$$n^{1/2} S_n(\beta, \gamma, r) = 0, \qquad 0 < r \le r_U \tag{2.2}$$

$$n^{1/2} U_n(\gamma) = 0, \tag{2.3}$$

where

$$S_n(\beta, \gamma, r) = n^{-1} \sum_{i=1}^{n} W_i \Big(N_i \exp\big(W_i^T \beta(r) + Z_i^T(\gamma)\big) - \int_0^r I\big[V_i \ge \exp\big(W_i^T \beta(u) + Z_i^T \gamma\big)\big] \mathrm{d}H(u)\Big)$$

and

$$U_n(\gamma) = n^{-1} \sum_{i=1}^{n} Z_i \Big(N_i \exp\big(W_i^T \hat\beta(r_U, \gamma) + Z_i^T \gamma\big) - \int_0^{r_U} I\big[V_i \ge \exp\big(W_i^T \hat\beta(u, \gamma) + Z_i^T \gamma\big)\big] \mathrm{d}H(u)\Big).$$

Here $\hat{\beta}(r,\gamma)$ denotes the solution to estimating equation (2.2) given a fixed γ. The estimator of γ_0, denoted by $\hat{\gamma}$, is defined as the solution to estimating equation (2.3), and the estimator of $\beta_0(r)$ is defined as $\hat{\beta}(r) = \hat{\beta}(r,\hat{\gamma})$. By the definition of $Q_T(0 \mid W, Z)$, and from model (2.1), $\exp(W^T\beta_0(0)+Z^T\gamma_0) = 0$ is achieved. This along with the boundedness of γ_0, implies $\exp(W^T\beta_0(0)) = 0$. So, $\exp(W_i^T\hat{\beta}(0,\gamma)) = 0$ for $i = 1,\ldots,n$ could be set.

Let the effect of Z takes the form $g(r;\theta)$, where $g(.)$ is a known pdf and θ is an unknown parameter. Then the quantile regression model takes the form:

$$Q_Y(r \mid W, Z) = W^T\beta(r) + Z^T g(r;\theta), \qquad r \in (0,1), \tag{2.4}$$

The estimating equations are simple modifications of equations (2.2) and (2.3) and thus taking forms,

$$n^{1/2}S_n(\beta, g(r;\theta), r) = 0, \qquad 0 < r \leq r_U \tag{2.5}$$

and

$$n^{1/2}U_n(g(r;\theta)) = 0, \tag{2.6}$$

where

$$S_n(\beta, g(r;\theta), r) = n^{-1}\sum_{i=1}^{n} W_i\Big(N_i \exp\big(W_i^T\beta(r)+Z_i^T g(r;\theta)\big) - \int_0^r I\big[V_i \geq \exp\big(W_i^T\beta(u)+Z_i^T g(r;\theta)\big)\big]\,\mathrm{d}H(u)\Big)$$

and

$$U_n(g(r;\theta)) = n^{-1}\sum_{i=1}^{n} Z_i\Big(N_i \exp\big(W_i^T\hat{\beta}(r_U, g(r;\theta))+Z_i^T g(r;\theta)\big) - \int_0^{r_U} I\big[V_i \geq \exp\big(W_i^T\hat{\beta}(u, g(r;\theta))+Z_i^T g(r;\theta)\big)\big]\,\mathrm{d}H(u)\Big).$$

Based on equations (2.5) and (2.6) the parameters are estimated using the iterative algorithm proposed by Qian and Peng (2010).

3 Simulation Study

Two scenarios of Monte Carlo simulation schemes have been used in this paper (following Qian and Peng, 2010). In scenario (I) event times were generated from a log-linear model with independent and identically distributed errors,

$$\log T = Q_\epsilon(r) + g(r;\boldsymbol{\theta}_1)Z_1 + g(r;\boldsymbol{\theta}_2)Z_2 + \epsilon,$$

where the error ϵ followed the extreme value distribution. Under this scenario, model (2.1) held with $W = 1$, $Z = (Z_1, Z_2)^T$ and $\beta_0(r) = Q_\epsilon(r)$ and $\gamma_0 \equiv (\gamma_{01}, \gamma_{02})^T = \{g(r;\boldsymbol{\theta}_1), g(r;\boldsymbol{\theta}_2)\}^T$, where

$\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ are the vectors of parameters of the respective distributions. C is uniformly distributed in $C \sim Unif\{0.1I(Z_1 = 1), c_u\}$ and values of c_u under normal, extreme value and logistic distribution are 2.5, 7.0 and 5.6 respectively for 25% censoring and 4.6, 12.0 and 10.0 for 15% censoring respectively.

In scenario (II) event times were generated from a log-linear model with heteroscedastic errors,

$$\log T = Q_\epsilon(r) + g(r; \boldsymbol{\theta})Z_1 + Q_W(r)\xi W_1 + \epsilon,$$

where the error ϵ followed the standard normal distribution. The covariate ξ was standard extreme value distributed and was independent of ϵ. Under scenario (II) model (2.1) held with $W = (1, W_1)^T$ and $Z = Z_1$ and $\beta_0(r) = \{\beta_{01}(r), \beta_{02}(r)\}^T = \{Q_\epsilon(r), Q_W(r)\}^T$ and $\gamma_0 = g(r; \boldsymbol{\theta})$, where $\boldsymbol{\theta}$ is the vector of parameters of the corresponding distribution. C is uniformly distributed in $C \sim Unif\{0.3I(Z_1 = 1), c_u\}$. Values of c_u under normal, extreme value and logistic distribution are 4.25,10.0 and 8.75 for 25% censoring and 7.04, 15.5 and 17.5 for 15% censoring respectively.

In each scenario, two different sample sizes 25 and 100 have been considered and 1000 datasets have been generated for each case. In (I) and (II), Z_1 was simulated from $uniform(0, 1)$, and Z_2, W_1 followed Bernoulli distributions with a success probability of 0.5.

Standard normal, standard extreme value and standard logistic distribution have been used to present the covariate effects to illustrate the situations where some covariate effects are believed to follow a specific distribution.

4 Sensitivity Analysis

One of the important advantages of quantile regression is its flexibility to have room for various effect patterns without assuming a distribution. A major concern under this scenario could be that the proposed estimator may produce bias if the parametric pattern is not correctly specified. Sensitivity analysis can be performed to examine the uncertainty in the output of a model that can be apportioned to different sources of uncertainty in its inputs (Salteli et.al., 2008). A large number of literature is available on sensitivity analysis, distinguished by the approaches or type of sensitivity measures based, for example, on variance decompositions, partial derivatives or elementary effects (Hossain and Muttlak, 2001; Jo, 2002; Cheung and Chappell, 2002; Majumdar et al., 2010).

In this study, the relative precision of the estimates under an assumed distribution over the correctly specified distribution is considered as a measure of sensitivity. Let $g(\cdot)$ has, e.g., a rectangular, normal, extreme value, logistic or gamma distribution. Under pattern misspecification one assumes $g(.)$ to follow some other distribution than the true one and obtains the estimates of the quantile regression model defined earlier. The precision of the estimates are calculated for the model with incorrectly specified parametric distribution and compared with the precision of the estimates under the correctly specified distributions.

The relative precision, ϕ, of the estimates are obtained as :

$$\phi = \frac{\big(1/mse(\text{estimates under wrongly specified distribution})\big)}{\big(1/mse(\text{estimates under correctly specified distribution})\big)}, \tag{4.1}$$

where mse is the mean squared error of the estimates which is a total of bias-squared and variance of the estimates. Values of relative precisions (for various estimators) for various assumed distributions relative to the correctly specified distributions under two scenarios (and for two sample sizes, which are 25 and 100) are presented in Tables 3 and 4 considering 15% of the data being censored.

5 Results

Table 1 reports simulation results under scenario (I) (i.e. model with independent and identically distributed errors) where some covariate effects follow standard normal distribution. Point estimates, $\hat{\beta}(r)(r = 0.1, 0.3, 0.5, 0.7), \hat{\theta}_1 \text{and} \hat{\theta}_2$ are approximately unbiased. The amount of bias tends to be smaller for larger sample size $(n = 100)$ and lower censoring rate.The resampling-based variance estimates, empirical variances and the difference between these two variances decrease with increasing size of sample. The resampling-based variance estimates agree with the empirical variances very well. The coverage probabilities are seen to be close to the nominal level for both sizes of the sample and for both censoring proportions. These probabilities increase as the sample size increases and censoring proportions decreases. The covariate effects were assumed to have several other distributions (e.g. extreme value, logistic) and similar results are obtained.

Table 1: Simulation results under scenario (I) for covariate effects having Normal distribution

		Normal Covariate Effects							
		Bias[a]		EstVar[b]		EmpVar[c]		Cov95[d]	
Estimated Effects	Sample Size	25% cens[e]	15% cens[e]	25% cens[e]	15% cens[e]	25% cens[e]	15% cens[e]	25% cens[e]	15% cens[e]
$\hat{\beta}(0.1)$	25	0.07	0.04	0.07	0.05	0.07	0.05	93.0	94.9
$\hat{\beta}(0.3)$	25	0.08	0.04	0.06	0.04	0.05	0.04	92.0	93.9
$\hat{\beta}(0.5)$	25	-0.06	-0.04	0.05	0.04	0.05	0.04	93.0	94.8
$\hat{\beta}(0.7)$	25	-0.06	-0.06	0.05	0.04	0.05	0.04	91.9	92.7
$\hat{\theta}_1 = \binom{\hat{\theta}_{11}}{\hat{\theta}_{12}}$	25	$\binom{-0.06}{0.05}$	$\binom{-0.04}{0.05}$	$\binom{0.05}{0.05}$	$\binom{0.06}{0.06}$	$\binom{0.05}{0.05}$	$\binom{0.05}{0.05}$	$\binom{92.9}{92.2}$	$\binom{94.2}{93.7}$
$\hat{\theta}_2 = \binom{\hat{\theta}_{21}}{\hat{\theta}_{22}}$	25	$\binom{0.05}{0.05}$	$\binom{0.04}{0.04}$	$\binom{0.07}{0.06}$	$\binom{0.06}{0.06}$	$\binom{0.07}{0.06}$	$\binom{0.06}{0.05}$	$\binom{94.0}{95.0}$	$\binom{94.2}{96.7}$
$\hat{\beta}(0.1)$	100	0.03	0.02	0.04	0.03	0.04	0.03	94.6	95.2
$\hat{\beta}(0.3)$	100	0.03	0.01	0.03	0.02	0.03	0.02	94.7	96.0
$\hat{\beta}(0.5)$	100	-0.01	-0.01	0.03	0.02	0.02	0.02	95.5	96.0
$\hat{\beta}(0.7)$	100	-0.01	-0.01	0.02	0.02	0.02	0.02	95.7	96.2
$\hat{\theta}_1 = \binom{\hat{\theta}_{11}}{\hat{\theta}_{12}}$	100	$\binom{-0.02}{0.02}$	$\binom{-0.02}{0.02}$	$\binom{0.03}{0.03}$	$\binom{0.02}{0.02}$	$\binom{0.03}{0.03}$	$\binom{0.02}{0.03}$	$\binom{95.7}{95.5}$	$\binom{95.9}{96.0}$
$\hat{\theta}_2 = \binom{\hat{\theta}_{21}}{\hat{\theta}_{22}}$	100	$\binom{0.02}{0.02}$	$\binom{0.01}{0.02}$	$\binom{0.03}{0.03}$	$\binom{0.02}{0.02}$	$\binom{0.03}{0.03}$	$\binom{0.02}{0.02}$	$\binom{94.7}{94.8}$	$\binom{95.0}{95.2}$

a: empirical bias; b: average of resampling-based variance estimates; c: empirical variance;
d: empirical coverage probabilities of 95% confidence intervals (%); e: censoring.

Table 2 presents results from scenario (II), demonstrating that the estimators for effects following logistic distribution are still approximately unbiased. Similar to scenario I, the resampling-based variance estimates, empirical variances and the difference between these two variances are larger for smaller sizes of sample. The resampling-based variance estimates and the empirical variances are very similar. The coverage probabilities are also seen to be close to the nominal level for both sample sizes and for both censoring proportions. The covariate effects were tested under the assumption of several other distributions (e.g. normal, extreme value) and similar results are obtained.

Table 2: Simulation results under scenario (II) for covariate effects having logistic distribution

		Logistic Covariate Effects							
		Bias^a		EstVar^b		EmpVar^c		Cov95^d	
Estimated Effects	Sample Size	25% cens^e	15% cens^e	25% cens^e	15% cens^e	25% cens^e	15% cens^e	25% cens^e	15% cens^e
$\hat{\beta}_1(0.1)$	25	0.06	0.06	0.06	0.06	0.06	0.06	93.7	93.9
$\hat{\beta}_1(0.3)$	25	0.05	0.05	0.06	0.05	0.06	0.05	94.7	94.9
$\hat{\beta}_1(0.5)$	25	-0.06	-0.05	0.06	0.05	0.06	0.05	93.2	94.6
$\hat{\beta}_1(0.7)$	25	-0.08	-0.07	0.05	0.05	0.05	0.04	93.9	94.2
$\hat{\beta}_2(0.1)$	25	0.08	0.07	0.07	0.06	0.06	0.06	92.0	93.4
$\hat{\beta}_2(0.3)$	25	0.07	0.06	0.06	0.06	0.06	0.05	93.5	94.0
$\hat{\beta}_2(0.5)$	25	0.07	-0.05	0.06	0.05	0.06	0.05	94.6	95.2
$\hat{\beta}_2(0.7)$	25	-0.07	-0.05	0.06	0.04	0.05	0.04	95.0	95.2
$\hat{\theta} = \binom{\hat{\theta}_1}{\hat{\theta}_2}$	25	$\binom{0.05}{0.05}$	$\binom{0.04}{0.04}$	$\binom{0.06}{0.06}$	$\binom{0.06}{0.06}$	$\binom{0.06}{0.06}$	$\binom{0.06}{0.06}$	$\binom{94.2}{94.0}$	$\binom{94.4}{94.8}$
$\hat{\beta}_1(0.1)$	100	0.03	0.02	0.02	0.02	0.02	0.02	94.8	95.9
$\hat{\beta}_1(0.3)$	100	0.02	0.02	0.02	0.02	0.02	0.02	95.0	96.1
$\hat{\beta}_1(0.5)$	100	-0.02	-0.01	0.02	0.02	0.02	0.01	95.2	95.8
$\hat{\beta}_1(0.7)$	100	-0.03	-0.02	0.02	0.01	0.02	0.01	94.0	95.0
$\hat{\beta}_2(0.1)$	100	0.04	0.02	0.02	0.02	0.02	0.02	94.0	94.8
$\hat{\beta}_2(0.3)$	100	0.03	0.02	0.02	0.02	0.02	0.02	94.4	95.2
$\hat{\beta}_2(0.5)$	100	0.04	-0.03	0.02	0.02	0.02	0.02	93.0	94.7
$\hat{\beta}_2(0.7)$	100	-0.02	-0.02	0.02	0.01	0.02	0.01	94.9	96.0
$\hat{\theta} = \binom{\hat{\theta}_1}{\hat{\theta}_2}$	100	$\binom{0.02}{0.02}$	$\binom{0.01}{0.01}$	$\binom{0.03}{0.03}$	$\binom{0.02}{0.02}$	$\binom{0.03}{0.03}$	$\binom{0.02}{0.02}$	$\binom{94.7}{94.7}$	$\binom{96.0}{95.9}$

a: empirical bias; b: average of resampling-based variance estimates; c: empirical variance;
d: empirical coverage probabilities of 95% confidence intervals (%); e: censoring.

Tables 3 and 4 represents the sensitivity analysis for scenario (I) and scenario (II) respectively under 15% censoring with sample sizes 25 and 100. Both the tables report that the precision of the estimates are largely affected by the misspecification of the underlying distributions. Even though, the precision of the estimates increases for a relatively larger sample ($n = 100$), the cost of misspecification in terms of precision is very high for smaller samples (say $n = 25$).

Table 3: Relative Efficiency (ϕ) of the parameter estimates under model misspecification with 15% censoring for scenario I for the sample sizes ($n = 25, n = 200$)

Assumed Distributions	Coefficients	True distribution: Rectangular	Normal	Extreme Value	Logistic	Gamma(0.5)
Rectangular	$\hat{\beta}(0.2)$	(1.00,1.00)	(0.58,0.87)	(0.45,0.72)	(0.49,0.77)	(0.47,0.76)
	$\hat{\beta}(0.4)$	(1.00,1.00)	(0.60,0.89)	(0.49,0.75)	(0.52,0.78)	(0.51,0.77)
	$\hat{\beta}(0.6)$	(1.00,1.00)	(0.69,0.90)	(0.58,0.79)	(0.60,0.87)	(0.60,0.85)
Normal	$\hat{\beta}(0.2)$	(0.60,0.91)	(1.00,1.00)	(0.39,0.79)	(0.48,0.87)	(0.48,0.85)
	$\hat{\beta}(0.4)$	(0.63,0.93)	(1.00,1.00)	(0.41,0.80)	(0.50,0.88)	(0.52,0.87)
	$\hat{\beta}(0.6)$	(0.65,0.94)	(1.00,1.00)	(0.42,0.84)	(0.54,0.89)	(0.57,0.88)
Gamma(0.5)	$\hat{\beta}(0.2)$	(0.56,0.70)	(0.52,0.68)	(0.44,0.53)	(0.47,0.58)	(1.00,1.00)
	$\hat{\beta}(0.4)$	(0.59,0.72)	(0.55,0.69)	(0.45,0.55)	(0.49,0.63)	(1.00,1.00)
	$\hat{\beta}(0.6)$	(0.60,0.75)	(0.58,0.70)	(0.46,0.57)	(0.50,0.64)	(1.00,1.00)

6 Conclusion

In this study, the partially functional quantile regression methods for randomly right censored data are tailored to the regression settings where some covariate effects follow certain parametric patterns. The estimating equations for censored quantile regression by Qian and Peng (2010) are modified for estimating the quantiles and other parameter(s) of interest.

In situations where the restricted model (under this study) is correctly specified, the estimators are more efficient than Qian and Peng (2010)'s estimator for model with partially functional effects. However, the sensitivity analysis shows that the consequences of misspecification of parametric forms could be severe, especially for small samples. So the practitioners should keep in mind that when the parametric pattern is not specified properly, the proposed estimator may cause bias. Therefore, this model is useful when there is a strong evidence or preliminary information on the pattern of the quantile effects. Nevertheless, while applying this technique, an important question arises as how to split the covariates in W and Z. We suggest that one start with fitting a partially functional model to the data and apply Qian and Peng (2010)'s approach to examine the form of the covariate effects first, and finally fit the proposed model as a restricted model if seems appropriate.

Table 4: Relative Efficiency (ϕ) of the parameter estimates under model misspecification with 15% censoring for scenario II for the sample sizes $(n = 25, n = 100)$

Assumed		True distribution				
Distributions	Coefficients	Rectangular	Normal	Extreme Value	Logistic	Gamma(0.5)
Rectangular	$\hat{\beta}_1(0.2)$	(1.00,1.00)	(0.49,0.82)	(0.36,0.71)	(0.45,0.77)	(0.44,0.75)
	$\hat{\beta}_2(0.2)$	(1.00,1.00)	(0.52,0.84)	(0.35,0.75)	(0.45,0.78)	(0.45,0.77)
	$\hat{\beta}_1(0.4)$	(1.00,1.00)	(0.53,0.87)	(0.39,0.77)	(0.48,0.80)	(0.47,0.78)
	$\hat{\beta}_2(0.4)$	(1.00,1.00)	(0.59,0.86)	(0.36,0.78)	(0.51,0.81)	(0.50,0.81)
	$\hat{\beta}_1(0.6)$	(1.00,1.00)	(0.54,0.88)	(0.44,0.79)	(0.50,0.82)	(0.48,0.81)
	$\hat{\beta}_2(0.6)$	(1.00,1.00)	(0.60,0.87)	(0.38,0.79)	(0.55,0.83)	(0.53,0.82)
Normal	$\hat{\beta}_1(0.2)$	(0.58,0.79)	(1.00,1.00)	(0.50,0.71)	(0.56,0.77)	(0.53,0.75)
	$\hat{\beta}_2(0.2)$	(0.55,0.77)	(1.00,1.00)	(0.48,0.70)	(0.50,0.76)	(0.50,0.75)
	$\hat{\beta}_1(0.4)$	(0.60,0.81)	(1.00,1.00)	(0.51,0.73)	(0.58,0.78)	(0.55,0.77)
	$\hat{\beta}_2(0.4)$	(0.57,0.79)	(1.00,1.00)	(0.51,0.70)	(0.54,0.78)	(0.53,0.77)
	$\hat{\beta}_1(0.6)$	(0.61,0.87)	(1.00,1.00)	(0.51,0.75)	(0.59,0.80)	(0.57,0.79)
	$\hat{\beta}_2(0.6)$	(0.60,0.82)	(1.00,1.00)	(0.52,0.72)	(0.56,0.79)	(0.54,0.78)
Gamma(0.5)	$\hat{\beta}_1(0.2)$	(0.59,0.78)	(0.57,0.77)	(0.50,0.67)	(0.54,0.72)	(1.00,1.00)
	$\hat{\beta}_2(0.2)$	(0.58,0.75)	(0.54,0.72)	(0.49,0.66)	(0.52,0.69)	(1.00,1.00)
	$\hat{\beta}_1(0.4)$	(0.61,0.80)	(0.58,0.78)	(0.52,0.70)	(0.54,0.75)	(1.00,1.00)
	$\hat{\beta}_2(0.4)$	(0.59,0.78)	(0.55,0.75)	(0.49,0.68)	(0.54,0.70)	(1.00,1.00)
	$\hat{\beta}_1(0.6)$	(0.64,0.83)	(0.62,0.80)	(0.54,0.71)	(0.57,0.77)	(1.00,1.00)
	$\hat{\beta}_2(0.6)$	(0.62,0.81)	(0.59,0.79)	(0.51,0.70)	(0.55,0.74)	(1.00,1.00)

References

[1] Barrodale, I. and Roberts, F.D.K. (1974). Solution of an Overdetermined System of Equations in the l_1 Norm. *Communications of the ACM,* **17***, 319-20.*

[2] Buchinsky, M. and Hahn, J. (1998). An Alternative Estimator for the Censored Quantile Regression Model. *Econometrica,* **66***, 653-71.*

[3] Buckley, J. and James, I. (1979). Linear Regression with Censored Data. *Biometrika,* **66***, 429-36.*

[4] Chernozhukov, V. and Hong, H. (2002). Three-Step Censored Quantile Regression and Extramarital Affairs. *J. Am. Statist. Assoc.,* **97***, 872-82*.

[5] Cheung, Y.K. and Chappell, R. (2002). A Simple Technique to Evaluate Model Sensitivity in the Continual Reassessment Method. *Biometrics.* **58***, 671-74.*

[6] Efron, B. (1967). The Two-Sample Problem with Censored Data. In Proc. *Fifth Berkeley Symp. Math. Statist. Prob.*, vol. IV, Ed. L. LeCam and J. Neyman. New York: Prentice-Hall.

[7] Fitzenberger, B. (1999). It Wages and Employment Across Skill Groups. Heidelberg: Physica-Verlag.

[8] Honoré, B., Khan, S. and Powell, J. L. (2002). Quantile Regression Under Random Censoring. *J. Economet,* **109***, 67-105.*

[9] Hossain, S. S. and Muttlak, J. H. A. (2001). Selected Ranked Set Sampling. *Aust. N. Z. J. Stat.,* **43***, 311-25.*

[10] Jo, B. (2002). Model Misspecification Sensitivity Analysis in Estimating Causal Effects of Interventions with Non-compliance. *Stat Med. 2002* **21***, 3161-81.*

[11] Koenker, R. and Bassett, G. (1978). Regression Quantiles. *Econometrica,* **46***, 33-50.*

[12] Louis, T. A. (1981). Nonparametric Analysis of an Accelerated Failure Time Model. *Biometrika,* **68***, 381-90.*

[13] Neocleous, T., Branden, K. V., and Portnoy, S. (2006). Correction to Censored Regression Quantiles by S. Portnoy, 98 (2003), 1001-1012. *J. Am. Statist. Assoc.,* **101***, 860-1.*

[14] Majumdar, A. Paul, D. and Kaye J. (2010). Sensitivity Analysis and Model Selection for a Generalized Convolution Model for Spatial Processes. *Bayesian Analysis.* **5***, 3, 493-518.*

[15] Peng, L. and Huang, Y. (2008). Survival Analysis with Quantile Regression Models. *J. Am. Statist. Assoc,* **103***, 637-49.*

[16] Portnoy, S. (2003). Censored Regression Quantiles. *J. Am. Statist. Assoc,* **98***, 1001-12.*

[17] Powell, J. (1984). Least Absolute Deviations Estimation for the Censored Regression Model. *J. Economet,* **25***, 303-25.*

[18] Powell, J. (1986). Censored Regression Quantiles. *J. Economet,* **32***, 143-55.*

[19] Prentice, R. L. (1978). Linear Rank Tests with Right Censored Data (Corr: vol. 70, p. 304). *Biometrika,* **65***, 167-80.*

[20] Qian, J. and Peng, L. (2010). Censored Quantile Regression with Partially Functional Effects. *Biometrika,* **97***, 839-50.*

[21] Ritov, Y. (1990). Estimation in a Linear Regression Model with Censored Data. *The Annals of Statistics,* **18***, 303-28.*

[22] Salteli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gateli, D. Saisana, M. and Tarantola, S. (2008). *Global Sensitivity Analysis*. The Primer, John Wiley & Sons.

[23] Tsiatis, A. A. (1990). Estimating Regression Parameters Using Linear Rank Tests for Censored Data. *Ann. Statist,* **18***, 354-72.*

[24] Wei, L. J. and Gail, M. H. (1983). Nonparametric Estimation of a Scale Change with Censored Observations. *J. Am. Statist. Assoc,* **78***, 382-88.*

[25] Wei, L. J., Ying, Z. and Lin, D.Y. (1990).Linear Regression Analysis of Censored Survival Data Based on Rank Tests. *Biometrika,* **77***, 845-51.*

[26] Yang, S. (1999). Censored Median Regression Using Weighted Empirical Survival and Hazard Functions. *J. Am. Statist. Assoc,* **94***, 137-45.*

[27] Ying, Z., Jung, S. H. and Wei, L. J. (1995). Survival Analysis with Median Regression Models. *J. Am. Statist. Assoc,* **90***, 178-84.*