Simple resampling methods for censored regression quantiles

Yannis Bilias\textsuperscript{a,b}, Songnian Chen\textsuperscript{c,*}, Zhiliang Ying\textsuperscript{d}

\textsuperscript{a}Departments of Economics and Statistics, 266 Heady Hall, Iowa State University, Ames, IA 50011, USA
\textsuperscript{b}Department of Economics, University of Cyprus, P.O. Box 20537, 1678 Nicosia, Cyprus
\textsuperscript{c}Department of Economics, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong
\textsuperscript{d}Department of Statistics, Hill Center, Busch Campus, Rutgers University, Piscataway, NJ 08855, USA

Received 1 June 2000; accepted 25 June 2000

Abstract

Powell (Journal of Econometrics 25 (1984) 303–325; Journal of Econometrics 32 (1986) 143–155) considered censored regression quantile estimators. The asymptotic covariance matrices of his estimators depend on the error densities and are therefore difficult to estimate reliably. The difficulty may be avoided by applying the bootstrap method (Hahn, Econometric Theory 11 (1995) 105–121). Calculation of the estimators, however, requires solving a nonsmooth and nonconvex minimization problem, resulting in high computational costs in implementing the bootstrap. We propose in this paper computationally simple resampling methods by convexifying Powell’s approach in the resampling stage. A major advantage of the new methods is that they can be implemented by efficient linear programming. Simulation studies show that the methods are reliable even with moderate sample sizes. © 2000 Elsevier Science S.A. All rights reserved.

\textit{JEL classification:} C14; C24

\textit{Keywords:} Censored regression quantiles; Least absolute deviation; Linear programming; Resampling

*Corresponding author. Tel.: + 852-2358-7602; fax: + 852-2358-2084.
E-mail address: snchen@ust.hk (S. Chen).
1. Introduction

Censored quantile regression models proposed by Powell (1984, 1986) have attracted a great deal of interest in the recent literature, particularly due to their robustness to distributional misspecification of the error term and unknown conditional heteroscedasticity. For empirical applications, see Buchinsky (1994), Chamberlain (1994), Chay and Honore (1998), Horowitz and Neumann (1987), and Melenberg and van Soest (1996), among others. A survey of the subject may be found in Buchinsky (1998).

In order to carry out valid statistical inferences for censored quantile regression models, it is necessary to provide consistent estimators for the asymptotic variance–covariance matrices. However, the asymptotic variance–covariance matrices are difficult to estimate reliably since they involve conditional densities of error terms. For the uncensored counterpart, Parzen et al. (1994) (PWY) proposed a resampling method that allows approximation of the distribution of regression quantile estimators to avoid density estimation. Their algorithm requires only repeated executions of the same linear programming algorithm (see, e.g., Koenker and Bassett, 1978a) and is computationally simple and efficient. Hahn (1995), on the other hand, showed that the usual bootstrap approximation is also valid for both the uncensored and censored regression quantiles, and Buchinsky (1995) provided some simulation evidence. In principle, the bootstrap approximation can be used to approximate distribution of Powell's regression quantile estimators for censored model. In practice, however, implementation of the bootstrap can be extremely time consuming as each bootstrap replication involves solving a nonlinear and nonconvex minimization problem and accurate approximation with the resampling methods typically requires large number of such replications. This could lead to potentially prohibitive computational costs in empirical applications.

The main objective of this paper is to propose computationally efficient methods for approximating the distribution of Powell's estimator through certain resampling methods. This is accomplished via convexifying the objective function of Powell (1984) so that the standard linear programming algorithm can be used. Section 2 introduces our modified bootstrap method and a similar modification of the resampling method of Parzen et al. (1994). In Section 3, it is shown that both the resampling methods are asymptotically correct. Some simulation results are presented in Section 4. Section 5 contains some concluding remarks.

2. Resampling methods for censored median regression

Consider the censored regression model

\[ y_i = \max \{ x_i' \beta + \varepsilon_i, 0 \} \quad i = 1, \ldots, n, \]  

(1)
where $e_i$ are error terms whose conditional medians given by $x_i$ are 0, $x_i$ are vectors of independent variables and $\beta_0$ is a conformable vector of true regression parameters. Powell (1984) recently proposed (CLAD) to estimate $\beta_0$ by

$$\hat{\beta} = \arg \min_{\beta} \sum_{i=1}^{n} |y_i - \max\{x_i'\beta, 0\}|.$$

(2)

Under certain regularity conditions, he showed that $\hat{\beta}$ is consistent and asymptotically normal in the sense that

$$\sqrt{n}(\hat{\beta} - \beta_0) \overset{d}{\to} N(0, \frac{1}{2}J_0^{-1}VJ_0^{-1}),$$

(3)

where $J_0 = E[f(0|x)xx'I(x'\beta_0 > 0)]$ and $V = E[xx'I(x'\beta_0 > 0)]$ with $f(\cdot|x)$ being the conditional density of $e$ given $x$.

Note that the objective function in (2) is a nonconvex and nonsmooth function and, consequently, its minimizer is difficult to calculate. A number of approaches have been developed for dealing with this problem. Womersley (1986) modified the reduced-gradient algorithm for the linear programming and his method is aimed at finding a local minimizer of the objective function. Buchinsky (1994) proposed an iterative linear programming algorithm, and Fitzenberger (1997) proposed an algorithm called BRECNS by adapting the standard Barrodale–Roberts Algorithm (Barrodale and Roberts, 1973). But, their methods do not guarantee convergence to global minimum. In addition, an interior point algorithm was developed by Koenker and Park (1996), but their method again does not guarantee convergence to the global minimizer. More recently, Fitzenberger and Winker (1999) proposed a new algorithm based on threshold accepting (TA), which requires a large number of iterations in its implementation. Even though their algorithm guarantees asymptotic convergence to the global optimum and a high probability of ending in a high-quality local optimum for a finite number of iterations, it is computationally expensive.

Note that the limiting variance–covariance matrix of $\hat{\beta}$ involves the conditional error density and thus can be difficult to estimate reliably. To avoid direct estimation of the error density, we can in principle apply the bootstrap method. Hahn (1995) provided theoretical justification for the bootstrap procedure. Buchinsky (1995) provided simulation evidence on the performance of the bootstrap in estimating the variance of quantile regression models. However, implementation of the bootstrap method requires the solving of nonconvex and nonsmooth minimization problem repeatedly, which could be prohibitively expensive computationally, especially when global maximization procedures, such as that of Fitzenberger and Winker (1999), are used.

We here proposed a modified bootstrap procedure by convexifying the objective function in the resampling stage and making use of the parameter estimator $\hat{\beta}$ already available during that stage. As a result, the bootstrap
For Eqs. (6) and (8), the left-hand sides could be $O_p(1)$ away from 0 due to its discontinuity in $\beta$ (see, e.g., Koenker and Bassett, 1978a, b; Powell, 1984).

Following the arguments in Powell (1984) and Pollard (1990), it is straightforward to show that Powell’s estimator is asymptotically equivalent to the solution of

$$
\min_{\beta} \sum_{i=1}^{n} |y_i - x'_i \beta| I(x'_i \beta_0 > 0). \tag{4}
$$

The minimization problem in (4) is a convex one and can be easily solved by the linear programming. However, this objective function is not realizable because $\beta_0$, or even a consistent estimator of it, is not available at that stage. Although the preceding convex modification of the objective function cannot be used to obtain an approximate solution to Powell’s estimator $\hat{\beta}$, it is suitable for obtaining the bootstrap estimators in the resampling stage. Specifically, let $(y^*_i, x^*_i)$, $i = 1, \ldots, n$, be a bootstrap sample from $\{(y_i, x_i), i = 1, \ldots, n\}$, define modified bootstrap estimator

$$
\hat{\beta}^* = \arg \min_{\beta} \sum_{i=1}^{n} |y^*_i - x^*_i \beta| I(x^*_i \hat{\beta} > 0). \tag{5}
$$

Note that in the resampling stage, $\hat{\beta}$ has already been computed. The minimization is over a convex function and can be easily executed via the linear programming. Because we need to have a large number of such $\hat{\beta}^*$ in the resampling stage, the computational savings could be substantial. The convexifying idea used here is similar to that of Buchinsky and Hahn (1998), Chernozhukov and Hong (1999), and Khan and Powell (1998), which modify Powell’s estimator.

Once a large number of $\hat{\beta}^*$ are computed, conditional distribution of $\hat{\beta}^* - \hat{\beta}$ can be approximated accurately. We will show in the next section that, under suitable regularity conditions, the conditional distribution scaled by $\sqrt{n}$ converges to the limit of the unconditional distribution of $\sqrt{n}(\hat{\beta} - \beta_0)$.

Our second resampling method is based on the observation that the estimating equation corresponding to (4) is pivotal, extending the resampling method by Parzen et al. (1994) for the uncensored median regression. More specifically, the estimating function corresponding to (4) takes the form

$$
\sum_{i=1}^{n} x_i \left[ I(y_i - x'_i \beta \leq 0) - \frac{1}{2} \right] I(x'_i \beta_0 > 0), \tag{6}
$$

---

1 For Eqs. (6) and (8), the left-hand sides could be $O_p(1)$ away from 0 due to its discontinuity in $\beta$ (see, e.g., Koenker and Bassett, 1978a, b; Powell, 1984).
which is conditionally pivotal at $\beta = \beta_0$ in the sense that its conditional distribution given the $x_i$ is the same as

$$U^* = \sum_{i=1}^{n} x_i \left[ B_i - \frac{1}{2} \right] I(x'_i \beta_0 > 0),$$

where $B_i$ are i.i.d. Bernoulli trials with success probability $1/2$. Our approach to extend the PWY method to censored data is to view Powell’s estimator $\hat{\beta}_K$ as an approximate solution to the estimating equation based on (6) and, in the resampling stage, to substitute the unknown $\beta_0$ by $\hat{\beta}_K$. Following Parzen et al. (1994), we generate $\{B_i, i = 1, \ldots, n\}$, a sequence of i.i.d. Bernoulli random variables with success probability $1/2$. Denote by $\hat{\beta}^{**}$ as the solution to

$$\sum_{i=1}^{n} x_i \left[ I(y_i - x'_i \hat{\beta} \leq 0) - \frac{1}{2} \right] I(x'_i \hat{\beta} > 0) + U^* = 0,$$

where $U^* = \sum_{i=1}^{n} x_i [B_i - \frac{1}{2}] I(x'_i \hat{\beta} > 0)$. It is worth pointing out that as in Parzen et al. (1994), this resampling method works for the median regression with stochastic as well as deterministic regressors.

The main advantage of $\hat{\beta}^{**}$ is computational. Following Parzen et al. (1994), we can compute $\hat{\beta}^{**}$ by minimizing a suitable $L_1$-objective function that can be solved by linear programming. Specifically, $\hat{\beta}^{**}$ is a minimizer of the following $L_1$-objective function

$$l(\beta) = \sum_{1 \leq i \leq n : x'_i \hat{\beta} > 0} |y_i - x'_i \beta| + |y_{n+1} - x'_{n+1} \beta|,$$

where $x_{n+1} = 2U^*$ and $y_{n+1}$ is chosen large enough so that $y_{n+1} > x'_{n+1} \beta$. See Parzen et al. (1994) for details.

To approximate the distribution of $\sqrt{n}(\hat{\beta}^{**} - \hat{\beta})$, one can simply simulate a large number of $U^*$’s, say $U^*_1, \ldots, U^*_M$, by generating $M$ sequences i.i.d. Bernoulli random variables. For each $U^*_j$, solve (9) to get $\hat{\beta}^{**}_j$. Then the empirical distribution of $\hat{\beta}^{**} - \hat{\beta}$ from the $\hat{\beta}^{**}_j$ should be close to the true distribution of $\hat{\beta} - \beta_0$, as will be shown in the next section.

3. Large sample properties

In this section, we show that the proposed resampling methods are asymptotically valid. We make the following assumptions.

**Assumption 1.** The true value of the parameter $\beta_0$ is in the interior of a known compact parameter space $B \subset \mathbb{R}^{n+1}$. 


Assumption 2. Conditional medians of the \( e_i \) given the \( x_i \) are uniquely equal to 0.

Assumption 3. \( P(x' \delta \neq 0 | x' \beta_0 > 0) > 0, \) for \( \delta \neq 0, P(x' \beta_0 = 0) = 0 \) and \( E[|x|^2] < \infty. \)

Assumption 4. \( E[f(0|x)x'I(x' \beta_0 > 0)] \) is nonsingular.

Assumption 5. The conditional density \( f(\cdot | x) \) is continuous and uniformly bounded.

The assumptions are standard for the desirable properties of Powell’s estimator to hold and are similar to those made in Powell (1984), Newey and McFadden (1994) and Pollard (1990). In particular, Assumption 1 is a standard assumption for an estimator defined through an optimization procedure with a nonconvex objective function. Assumption 2 is essential for \( \beta_0 \) to be uniquely defined; it is implied by the assumption that \( f(0|x) \) is nonzero and \( f(t|x) \) is continuous at \( t = 0. \) Assumptions 3 and 4 are crucial to ensure consistency and asymptotic normality of Powell’s estimator. Assumption 5 on the conditional density function is imposed to facilitate Taylor’s expansions used in the proofs.

Theorem 1. Under Assumptions 1–5 given above, the distribution for the modified bootstrap estimator \( \hat{\beta}^*, P(\sqrt{n}(\hat{\beta}^* - \hat{\beta}) \leq z|x_i, y_i, i = 1, \ldots, n), \) is consistent in the sense that

\[
P(\sqrt{n}(\hat{\beta}^* - \hat{\beta}) \leq z|x_i, y_i, i = 1, \ldots, n) - P(\sqrt{n}(\hat{\beta} - \beta_0) \leq z) \to 0
\]

in probability for every \( z. \)

Proof. Let \( F_n \) denote the empirical distribution formed from \( (x_1, y_1), \ldots, (x_n, y_n) \) and \( E_n^*(P_n^*) \) the corresponding expectation (probability measure) with respect to \( F_n. \) So \( (x_i^*, y_i^*) \) are i.i.d. with \( F_n \) as their common distribution. Define objective function

\[
G_n^*(b, \beta) = \frac{1}{n} \sum_{i=1}^{n} \left( |y_i^* - x_i^* \beta| - |y_i^* - x_i^* \beta|I(x_i^* \beta > 0) \right).
\]

Let \( G_n(b, \beta) = E_n^* G_n^*(b, \beta) \) and \( G(b, \beta) = E G_n(b, \beta). \) In view of Assumptions 3 and 4, the class of functions, \( \{|y - x' \beta| - |y - x' \beta|I(x' \beta > 0): b, \beta \in B\} \), is Euclidean with an integrable envelope in the sense of Pakes and Pollard (1989). Thus, by the uniform law of large numbers (Pollard, 1990),

\[
G_n(b, \beta) - G(b, \beta) = o(1) \quad \text{uniformly in } (b, \beta).
\]
Further, by the bootstrap uniform law of large numbers (Giné and Zinn, 1990, Theorem 2.6),

\[ G_n^*(b, \beta) - G_n(b, \beta) = o_B(1) \quad \text{uniformly in } (b, \beta), \]

where, for a random sequence \( v_n \) constructed from the bootstrap sample, \( v_n = o_B(1) \) if for every \( \eta > 0 \), \( P^*(\|v_n\| \geq \eta) \to 0 \) in probability (Hahn, 1995). Since \( \hat{\beta} \) is consistent and \( G \) is continuous, it follows that

\[ G_n^*(b, \hat{\beta}) - G(b, \beta_0) = o_B(1) \quad \text{uniformly in } b. \]

But \( G(b, \beta_0) \) is continuous in \( b \) and attains its unique minimum at \( b = \beta_0 \). Consequently, \( \hat{\beta}^* \), the minimizer of \( G_n^*(b, \hat{\beta}) \) over \( b \), is consistent, i.e., \( \hat{\beta}^* - \beta_0 = o_B(1) \) or, equivalently, \( \hat{\beta}^* - \beta = o_B(1) \).

Next, we show that the bootstrap distribution is asymptotically correct. Let

\[ \Delta_i(b, \beta) = |y_i - x_i \beta| - |y_i - x_i \hat{\beta}| \quad \text{and} \quad \Delta_i^*(b, \beta) = |y_i^* - x_i^* \beta| - |y_i^* - x_i^* \hat{\beta}|. \]

So \( G_n(b, \beta) = n^{-1} \sum_{i=1}^n \Delta_i(b, \beta) I(x_i \hat{\beta} > 0) \) and \( G_n^*(b, \beta) = n^{-1} \sum_{i=1}^n \Delta_i^*(b, \beta) I(x_i^* \hat{\beta} > 0) \). Further, let

\[ D_i(\beta) = \text{sgn}(y_i^* - x_i^* \hat{\beta}) I(x_i^* \hat{\beta} > 0)x_i, \quad R_i(b, \beta) = [\Delta_i(b, \beta) - (b - \beta)D_i(\beta)] I(x_i \hat{\beta} > 0) \]

and \( W_n(\beta) = n^{-1/2} \sum_{i=1}^n (D_i(\beta) - E_nD_i(\beta)) \). Likewise, let \( D_i^*(\beta) = \text{sgn}(y_i^* - x_i^* \hat{\beta}) I(x_i^* \hat{\beta} > 0)x_i^*, \quad R_i^*(b, \beta) = [\Delta_i^*(b, \beta) - (b - \beta)D_i^*(\beta)] I(x_i^* \hat{\beta} > 0) \) and \( W_n^*(\beta) = n^{-1/2} \sum_{i=1}^n (D_i^*(\beta) - E_n^*D_i^*(\beta)) \) be their resampled counterparts. Then simple algebra results in

\[ G_n(b, \beta) = \frac{1}{\sqrt{n}} (b - \beta)' W_n(\beta) + G(b, \beta) + \frac{1}{n} \sum_{i=1}^n [R_i(b, \beta) - E_nR_i(b, \beta)] \]

and

\[ G_n^*(b, \beta) = \frac{1}{\sqrt{n}} (b - \beta)' W_n^*(\beta) + G_n(b, \beta) + \frac{1}{n} \sum_{i=1}^n [R_i^*(b, \beta) - E_n^*R_i^*(b, \beta)]. \]

The right-hand side of (11) involves (10), which will be dealt with first using the same argument given in Pollard (1990, p. 63). Specifically, it is easy to verify that

\[ |R_i(b, \beta)| \leq 2|x_i'(b - \beta)|I(|y_i - x_i \beta| \leq |x'(b - \beta)|), \]

which implies

\[
\sup_{|b - \beta_0| + |\beta - \beta_0| \leq r} \frac{|R_i(b, \beta)|}{|b - \beta|} \leq 2|x_i|I(|y_i - x_i \beta_0| \leq 2r|x_i|) + I(|x_i \beta_0| \leq 2r|x_i|) \leq 2|x_i|I(|e_i| \leq 2r|x_i|) + I(|x_i \beta_0| \leq 2r|x_i|).\]
So the maximal inequality of Pollard (1990, p. 38) can be applied to get, for some $C > 0$,

$$
E \left[ \sup_{|b - \beta_0| + |\beta - \beta_0| \leq r} \left| \frac{\sum_i (R_i(b, \beta) - ER_i(b, \beta))}{\sqrt{n}|b - \beta|} \right|^2 \right] \leq \frac{C}{n} \sum_{i=1}^{n} |x_i|^2 \left[ I(|\varepsilon_i| \leq 2r|x_i|) + I(|x'_i\beta_0| \leq 2r|x_i|) \right] = o(1)
$$
as $n \to \infty$ and $r \to 0$, where the last equality follows from Assumptions 3 and 5. Thus, uniformly in $b$ and $\beta$ over shrinking neighborhoods of $\beta_0$,

$$
\sum_{i=1}^{n} (R_i(b, \beta) - ER_i(b, \beta)) = o(|b - \beta|/\sqrt{n}). \tag{12}
$$

By the Taylor expansion, $G(b, \beta) = (b - \beta)'J_0(b - \beta) + o(|b - \beta|^2)$, where $J_0 = E[f(0|x)xx'I(x'\beta_0 > 0)]$. This and (12) can now be applied to (10) to get

$$
G_n(b, \beta) = \frac{1}{n} (b - \beta)' \sum_{i=1}^{n} x_i \text{sgn}(y_i - x'_i\beta)I(x'_i\beta > 0)
$$

$$
+ (b - \beta)'J_0(b - \beta) + o_p \left( \frac{|b - \beta|}{\sqrt{n}} + |b - \beta|^2 \right)
$$

uniformly in $b$ and $\beta$ over shrinking neighborhoods of $\beta_0$. But $\hat{\beta}$, Powell’s estimator, satisfies $\sum x_i \text{sgn}(y_i - x'_i\hat{\beta})I(x'_i\hat{\beta} > 0) = o_p(\sqrt{n})$. Hence,

$$
G_n(b, \hat{\beta}) = (b - \hat{\beta})'J_0(b - \hat{\beta}) + o_p \left( \frac{|b - \hat{\beta}|}{\sqrt{n}} + |b - \hat{\beta}|^2 \right).
$$

Exactly the same arguments as those for (12) lead to

$$
\sum_{i=1}^{n} (R_i^*(b, \beta) - ER_i^*(b, \beta)) = o_p(|b - \beta|/\sqrt{n}).
$$

Consequently, we obtain

$$
G_n^*(b, \hat{\beta}) = \frac{1}{\sqrt{n}} \left( (b - \hat{\beta})'W_n^*(\hat{\beta}) + (b - \hat{\beta})'J_0(b - \hat{\beta}) \right)
$$

$$
+ o_p \left( \frac{|b - \hat{\beta}|}{\sqrt{n}} + |b - \hat{\beta}|^2 \right).
$$

Since, conditional on $\{(y_i, x_i), i = 1, \ldots, n\}$, $W_n^*(\hat{\beta})$ is a sum of i.i.d. random variables with mean 0 and variance–covariance matrix converging to $V = E[xx'I(x'\beta_0 > 0)]$, we know that the conditional distribution of $W_n^*(\hat{\beta})$
converges to \( N(0, V) \) in probability. Thus, an application of Lemma A.1 below with \( \theta = b - \hat{\beta} \) and \( \theta_n = \hat{\beta}^* - \hat{\beta} \) gives \( \hat{\beta}^* - \hat{\beta} = O_P(1/\sqrt{n}) \). A further application of Lemma A.2 results in

\[
\sqrt{n}(\hat{\beta}^* - \hat{\beta}) = -\frac{1}{2} J_0^{-1} W_n^*(\hat{\beta}) + o_P(1),
\]

whose conditional distribution clearly converges in probability to \( N(0, J_0^{-1} V J_0^{-1}) \). Hence Theorem 1 holds.

**Theorem 2.** Under Assumptions 1–5,

\[
P(\sqrt{n}(\hat{\beta}^* - \hat{\beta}) \leq z|x_i, y_i, i = 1, \ldots, n) = P(\sqrt{n}(\hat{\beta} - \beta_0) \leq z) \to 0
\]

in probability for every \( z \in \mathbb{R}^{p+1} \).

**Proof.** The derivation given of Parzen et al. (1994, Appendix) can be very much carried over to derive our result. So we will only sketch a few key steps in our proof. By Assumptions 1–4 and in view of (8), we can derive an asymptotic linearity result similar to (A1.1) in Parzen et al. (1994), i.e.,

\[
-U^* = \sum_{i=1}^{n} f(0|x_i) x_i \mathbb{I}(x_i^* \hat{\beta} > 0)(\hat{\beta}^* - \hat{\beta}) + o(n|\hat{\beta}^* - \hat{\beta}| + \sqrt{n})
\]

\[
= J_0 n(\hat{\beta}^* - \hat{\beta}) + o(n|\hat{\beta}^* - \hat{\beta}| + \sqrt{n}).
\]

Therefore, we have the following representation:

\[
\sqrt{n}(\hat{\beta}^* - \hat{\beta}) = -\frac{1}{\sqrt{n}} J_0^{-1} U^* + o(1),
\]

which clearly converges to \( N(0, J_0^{-1} V J_0^{-1}/4) \), the same limiting distribution as that of \( \sqrt{n}(\hat{\beta} - \beta_0) \). Hence the theorem holds.

4. A simulation study

Monte Carlo experiments were carried out to assess the performance of our methods. For the censored regression model, data are generated as follows in all of the designs:

\[
y = \max\{1 + x_1 \beta_1 + x_2 \beta_2 + \epsilon, 0\}, \tag{13}
\]

where \( x_1 \) is a Bernoulli random variable centered at zero with a success probability \( \frac{1}{2} \), \( x_2 \) a standard normal random variable, and \((\beta_1, \beta_2) = (1,1)\). Different designs are constructed by varying the distributions of the error term. Specifically, the three distributions for \( \epsilon \) are the standard normal distribution
N(0, 1) (Standard Normal), a heteroscedastic normal \((1 + x_2) \ast N(0, 1)\) (Heteroscedastic Normal) and a normal mixture \(0.75 \ast N(0, 1) + 0.25 \ast N(0, 4)\) (Normal Mixture). As a result, the censoring level is approximately 30% for all the designs.

Here we investigate the finite sample performance of three resampling methods – the direct heteroscedastic bootstrap method (Bootstrap), our modified bootstrap (M-Bootstrap) and the resampling method (Resampling) based on Eq. (8), where Powell’s CLAD estimator is chosen as the first-step estimator \(\hat{\beta}\). For that purpose, 1000 random samples are drawn with sample size equal to 100 according to Eq. (13). For each random sample \(\{(y_i, x_i), i = 1, \ldots, 100\}\), 1000 bootstrap samples are drawn and, accordingly, 1000 bootstrap estimates and 1000 modified bootstrap estimates are obtained. In addition, 1000 replicates of \(U^*\), the weighted sum of independent and centered Bernoulli variables in (8), are drawn and their corresponding resampling estimates are obtained. The standard (S) and percentile (P) methods (Efron and Tibshirani, 1993) are then used to construct confidence intervals of the regression coefficient of the continuous regressor \(x_2\) using the three methods. The empirical coverage probabilities are summarized in Table 1. All the three methods appear to give reasonable coverage probabilities and are comparable.

All the resampling methods are implemented in Gauss code, and each of our methods is about eight times faster than the direct bootstrap. Powell’s CLAD was implemented through iterative linear programming (Buchinsky, 1994), which behaves reasonably well for the designs under consideration here. It is worth pointing out that, in general, there is no guarantee that the iterative linear programming or other locally based search method would converge, or converge to the global minimum. In cases where global minimization methods such as that of Fitzenberger and Winker (1999) are called for and/or a large number of bootstrap replicates are needed for accurate statistical inference (Efron and Tibshirani, 1993), the computational advantage of our resampling methods will be even more significant.

5. Concluding remarks

In this paper we have proposed two simple resampling methods for censored median regression estimator introduced by Powell (1984). Unlike the uncensored case (Koenker and Bassett, 1978a), where the linear programming may be used, computation of Powell’s estimator involves solving nonconvex and non-smooth objective function. Such computational complexity can be of serious concern when the bootstrap and other resampling methods are called for to avoid estimating the limiting variance–covariance matrix that involves the error density. Two resampling methods are proposed for approximating the distribution of Powell’s estimator. One is a simple modification of the bootstrap method.
Table 1
Empirical coverage probabilities (ECP) for confidence intervals

<table>
<thead>
<tr>
<th>Confidence level</th>
<th>Bootstrap ECP</th>
<th>Resample ECP</th>
<th>M-Bootstrap ECP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>S</td>
<td>P</td>
<td>S</td>
</tr>
<tr>
<td>0.95</td>
<td>0.956</td>
<td>0.929</td>
<td>0.912</td>
</tr>
<tr>
<td>0.90</td>
<td>0.974</td>
<td>0.952</td>
<td>0.943</td>
</tr>
<tr>
<td>0.85</td>
<td>0.909</td>
<td>0.878</td>
<td>0.863</td>
</tr>
<tr>
<td></td>
<td>0.941</td>
<td>0.900</td>
<td>0.886</td>
</tr>
<tr>
<td></td>
<td>0.868</td>
<td>0.830</td>
<td>0.812</td>
</tr>
<tr>
<td></td>
<td>0.906</td>
<td>0.846</td>
<td>0.833</td>
</tr>
</tbody>
</table>

(a) Standard normal

(b) Normal mixture

<table>
<thead>
<tr>
<th>Confidence level</th>
<th>Bootstrap ECP</th>
<th>Resample ECP</th>
<th>M-Bootstrap ECP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>S</td>
<td>P</td>
<td>S</td>
</tr>
<tr>
<td>0.95</td>
<td>0.957</td>
<td>0.936</td>
<td>0.926</td>
</tr>
<tr>
<td>0.90</td>
<td>0.975</td>
<td>0.938</td>
<td>0.935</td>
</tr>
<tr>
<td>0.85</td>
<td>0.923</td>
<td>0.892</td>
<td>0.875</td>
</tr>
<tr>
<td></td>
<td>0.941</td>
<td>0.878</td>
<td>0.872</td>
</tr>
<tr>
<td></td>
<td>0.879</td>
<td>0.843</td>
<td>0.824</td>
</tr>
<tr>
<td></td>
<td>0.901</td>
<td>0.829</td>
<td>0.822</td>
</tr>
</tbody>
</table>

(c) Heteroscedastic normal

(Hahn, 1995), which convexifies the nonconvex objective function so that the efficient linear programming algorithm can be used to compute bootstrap estimates. The other is a similar modification of a resampling method proposed by Parzen et al. (1994). Again the computation of resampling estimates can be achieved via linear programming.

Our results could be extended in several directions. Koenker and Bassett (1978a) and Powell (1986) extended uncensored and censored median regressions to general regression quantiles, and the results are useful for testing heteroscedasticity for the error distribution (Koenker and Bassett, 1982; Powell, 1986). Our resampling results can also be extended for general censored regression quantiles to test for heteroscedasticity. Recently, Fitzenberger (1998) considered the moving block bootstrap in the presence of serial dependence. It might be possible to extend the moving block bootstrap and our convexifying idea to the censored quantile regression model with dependent data. It is a topic for future research.
Acknowledgements

We would like to thank two anonymous referees and an associate editor, as well as Bo Honoré, Shakeeb Khan, Roger Koenker, Lungfei Lee and Jim Powell for their helpful comments. All remaining errors are ours.

Appendix

The following lemmas are bootstrap analogues of Theorems 1 and 2 of Sherman (1993). For a weakly consistent bootstrap estimator $\theta_n = o_B(1)$, defined by a minimization procedure, Lemma A.1 provides conditions under which $\sqrt{n}\theta_n = o_B(1)$, while Lemma A.2 strengthens this $\sqrt{n}$-consistency result to weak convergence provided there exists a very good quadratic approximation to the objective function within $o_B(n^{-1/2})$ neighborhoods of 0. The proofs of the theorems are similar to Sherman (1993), thus omitted here.

Lemma A.1. Let $\theta_n$ be a minimizer of $\Gamma_n^*(\theta)$. Suppose that $\theta_n = o_B(1)$, and that uniformly over $o(1)$ neighborhoods of 0,

$$
\Gamma_n^*(\theta) = \theta'V\theta + O_B\left(\frac{|\theta|}{\sqrt{n}}\right) + o_B(|\theta|^2) + O_B\left(\frac{1}{n}\right),
$$

where $V$ is a positive definite matrix. Then

$$\theta_n = O_B(1/\sqrt{n}).$$

Lemma A.2. Let $\Theta$ be a subset of $\mathbb{R}^d$. Suppose $\theta_n$ minimizes $\Gamma_n^*(\theta)$ over $\Theta$ and that $\sqrt{n}\theta_n = o_B(1)$. Suppose 0 is an interior point of $\Theta$ and that uniformly over $O_B(n^{-1/2})$ neighborhoods of 0,

$$
\Gamma_n^*(\theta) = \frac{1}{2} \theta'V\theta + \frac{1}{\sqrt{n}} Z_n^*\theta + o_B\left(\frac{1}{n}\right),
$$

where $V$ is a positive definite matrix, and $Z_n^*$ converges in distribution to a $\mathcal{N}(0, \Sigma)$ random vector in probability. Then

$$\sqrt{n}\theta_n = -V^{-1}Z_n^* + o_B(1).$$

References

Chernozhukov, V., Hong, H., 1999. Censored quantile regression with the extreme quantile estimate of the unknown censoring points, unpublished manuscript.

