Statistica Sinica **14**(2004), 333-347

LINEAR HYPOTHESIS TESTING IN CENSORED REGRESSION MODELS

L. C. Zhao

University of Science and Technology of China

Abstract: For testing a linear hypothesis in a censored regression (or censored "Tobit") model, three test criteria and four test statistics based on least absolute deviations estimates of parameters are proposed and their limiting chi-square distributions are established. Some consistent estimates of nuisance parameters are obtained for use in computing the test statistics. A simulation study for small sample performance of these test statistics is made by using iterative linear programming.

Key words and phrases: Censored regression model, least absolute deviations method, limiting distribution, linear hypothesis testing.

1. Introduction

Assume that in the regression model

$$Y_i = \mathbf{x}_i'\boldsymbol{\beta} + e_i, \qquad i = 1, \ldots, n, \tag{1.1}$$

only $Y_i^+ = Y_i I(Y_i \geq 0)$ and $\mathbf{x}_i$ are observable, where $I(\cdot)$ denotes the indicator function of a set, $\{\mathbf{x}_i\}$ is a sequence of known p-vectors, $\{e_i\}$ is a sequence of non-observable random errors, and β is the unknown p-vector of regression coefficients. In other words, we consider the following model with the non-negativity constraint on the dependent variables:

$$Y_i^+ = (\mathbf{x}_i'\boldsymbol{\beta} + e_i)^+, \qquad i = 1, \ldots, n. \tag{1.2}$$

Such a model is called the censored regression (or censored "Tobit") model. This is an important one among limited dependent variable (LDV) models, for which the range of the dependent variable is restricted to some subset of the real line. It is worthwhile to emphasize that many of the important advances in econometric theory are related to LDV models. For early literature, see Maddala (1983) or Powell (1984) for example. Recently, censored quantile regression models proposed by Powell (1984, 1986) have attracted a great deal of interest due to their robustness. For further developments see Pollard (1990), Rao and

Zhao (1993), Chen and Wu (1994), Fitzenberger (1997), Buchinsky and Hahn (1998) and Bilias, Chen and Ying (2000), among others. For related applications see, e.g., Buchinsky (1994), Chamberlain (1994), Chay and Honoré (1998), and a survey of the subject by Buchinsky (1998).

In the following we make two standard assumptions (see Pollard (1990) for example).

(A1) $e_1, e_2, \ldots$ are i.i.d. random variables such that the common distribution function F has zero median and positive derivative $f(0)$ at zero.

(A2) The parameter space B of $\boldsymbol{\beta}$ is a bounded open subset of $\mathcal{R}^p$ (with a closure $\overline{B}$).

Based on the fact that $\mathrm{med}(Y_i^+) = (\mathbf{x}_i'\boldsymbol{\beta})^+$, Powell (1984) introduced the least absolute deviations (LAD) estimate $\widehat{\boldsymbol{\beta}}_n$ of $\boldsymbol{\beta}$, which is a Borel-measurable solution of the minimization problem

$$\sum_{i=1}^{n} \Big| Y_i^+ - (\mathbf{x}_i'\widehat{\boldsymbol{\beta}}_n)^+ \Big| = \min \Big\{ \sum_{i=1}^{n} \Big| Y_i^+ - (\mathbf{x}_i'\boldsymbol{\beta})^+ \Big| : \boldsymbol{\beta} \in \overline{B} \Big\}. \tag{1.3}$$

The consistency and asymptotic normality of $\widehat{\boldsymbol{\beta}}_n$ have been studied by Powell (1984), Pollard (1990), Rao and Zhao (1993) and Chen and Wu (1994), among others.

In this paper we are interested in testing

$$H_0 : H'(\boldsymbol{\beta} - \mathbf{b}_0) = \mathbf{0} \quad \text{against} \quad H_1 : H'(\boldsymbol{\beta} - \mathbf{b}_0) \neq \mathbf{0}, \tag{1.4}$$

where H is a known $p \times q$ matrix of rank q, and $\mathbf{b}_0$ is a known p-vector ($0 < q \leq p$).

For testing (1.4) in the above semiparametric model, we present three criteria based on the LAD method. Under some mild conditions we establish the limiting chi-square distribution of these criteria under H_0. In addition, in order that these results can be used, we give some consistent estimates for nuisance parameters involved, and suggest four test statistics: $4\widehat{f}(0)M_n, 4\widehat{f}(0)^2\widehat{W}_n, \widehat{R}_n$ and $4M_n^2/\widehat{W}_n$. The first three of these tests resemble the likelihood-ratio, the Wald and the score test in the usual parametric models, respectively.

For solving the minimization problem (1.3), Buchinsky (1994) proposed an iterative linear programming algorithm. By using this method, we can calculate the relevant test statistics even a linear constraint is involved. A simple simulation study shows that the chi-square approximations for these test statistics are pretty good even with moderate sample size, and that the performances of the first two of them are better than the other two.

In Section 2, the tests and their limiting distributions are presented. Proofs of the main theorems are given in Section 3, along with the statements of some auxiliary lemmas (whose proofs are relegated to the Appendix). Some simulation results are presented in Section 4.

2. Methods and Main Results

To test (1.4), consider

$$M_n := \inf_{H'(\mathbf{b}-\mathbf{b}_0)=\mathbf{0}} \sum_{i=1}^{n} \left|(\mathbf{x}_i'\mathbf{b})^+ - Y_i^+\right| - \inf_{\mathbf{b}} \sum_{i=1}^{n} \left|(\mathbf{x}_i'\mathbf{b})^+ - Y_i^+\right|, \tag{2.1}$$

where infima are taken over all $\mathbf{b} \in \overline{B}$ with and without the constraint $H'(\mathbf{b} - \mathbf{b}_0) = \mathbf{0}$, and assumed to be attained at $\boldsymbol{\beta}_n^*$ and $\widehat{\boldsymbol{\beta}}_n$ respectively.

As mentioned above, we appeal to an iterative linear programming method, noting that the linear constraint in the first minimization does not affect its linear programming nature.

Let $\boldsymbol{\beta}$ be the true parameter, and take $\mu_i := \mathbf{x}_i'\boldsymbol{\beta}$ and $S_n := \sum_{i=1}^{n} I(\mu_i > 0)\mathbf{x}_i\mathbf{x}_i'$. Throughout the paper, it is always assumed that S_{n_0} is positive definite for some n_0 and that $n \geq n_0$, so S_n^{-1} exists. In addition to M_n, we also study the Wald-type test criterion $W_n := (\widehat{\boldsymbol{\beta}}_n - \mathbf{b}_0)'H(H'S_n^{-1}H)^{-1}H'(\widehat{\boldsymbol{\beta}}_n - \mathbf{b}_0)$ and Rao′s score-type test criterion $R_n := \boldsymbol{\xi}(\boldsymbol{\beta}_n^*)'S_n^{-1}\boldsymbol{\xi}(\boldsymbol{\beta}_n^*)$, with $\boldsymbol{\xi}(\mathbf{b}) := \sum_{i=1}^{n} I(\mathbf{x}_i'\mathbf{b} > 0)\ \mathrm{sgn}(\mathbf{x}_i'\mathbf{b} - Y_i^+)\mathbf{x}_i = \sum_{i=1}^{n} I(\mathbf{x}_i'\mathbf{b} > 0)\ \mathrm{sgn}(\mathbf{x}_i'\mathbf{b} - Y_i)\mathbf{x}_i$. Note that S_n depends on the unknown parameter $\boldsymbol{\beta}$, so that W_n and R_n are not test statistics.

To study the limiting distribution of M_n, W_n and R_n under H_0, we further assume the following.

(A3) For any $\sigma > 0$, there exists a finite $\alpha > 0$ such that $\sum_{i=1}^{n} ||\mathbf{x}_i||^2 I(||\mathbf{x}_i|| > \alpha) < \sigma\lambda(S_n)$ for large n, where $\lambda(S_n)$ denotes the smallest eigenvalue of S_n.

(A4) For any $\sigma > 0$, there exists $\delta > 0$ such that $\sum_{i=1}^{n} ||\mathbf{x}_i||^2 I(|\mu_i| \leq \delta) < \sigma\lambda(S_n)$ for large n.

(A5) $\lambda(S_n)(\log n)^{-2} \to \infty$ as $n \to \infty$.

Note that the above conditions are weaker than those appearing in the literature, see, e.g., Pollard (1990).

Write $\mathbf{x}_{in} := S_n^{-1/2}\mathbf{x}_i$ and $H_n := S_n^{-1/2}H(H'S_n^{-1}H)^{-1/2}$. Then $\sum_{i=1}^{n} I(\mu_i > 0)\mathbf{x}_{in}\mathbf{x}_{in}' = I_p$ and $H_n'H_n = I_q$, where I_q is the identity matrix of order q.

Theorem 2.1. *Suppose that* (A1)$-$(A5) *are satisfied. If $\boldsymbol{\beta}$ is the true parameter and H_0 holds, then*

$$4f(0)M_n = ||\sum_{i=1}^{n} I(\mu_i > 0)\ \mathrm{sgn}(e_i)\ H_n'\mathbf{x}_{in}||^2 + o_p(1), \tag{2.2}$$

$$4f(0)^2 W_n = ||\sum_{i=1}^{n} I(\mu_i > 0)\ \mathrm{sgn}(e_i)\ H_n'\mathbf{x}_{in}||^2 + o_p(1), \tag{2.3}$$

$$R_n = ||\sum_{i=1}^{n} I(\mu_i > 0)\ \mathrm{sgn}(e_i)\ H_n'\mathbf{x}_{in}||^2 + o_p(1). \tag{2.4}$$

Consequently, $4f(0)M_n$, $4f(0)^2W_n$ *and* R_n *have the same limiting central chi-square distribution with* q *degrees of freedom.*

In order for the above results to be useful in testing the hypothesis H_0 against H_1, some consistent estimates of S_n and $f(0)$ (under H_0) should be obtained. We say that $\widehat{S}_n$ is a consistent estimate of the matrix S_n if $S_n^{-1/2}\widehat{S}_nS_n^{-1/2} \to I_p$ in probability as $n \to \infty$. Take $\widehat{S}_n := \sum_{i=1}^n I(\mathbf{x}_i'\widehat{\boldsymbol{\beta}}_n > 0)\mathbf{x}_i\mathbf{x}_i'$ as an estimate of S_n and, with $h = h_n > 0$, $h_n \to 0$, estimate $f(0)$ by $\widehat{f}_n(0) := \left(h\sum_{i=1}^n I(\mathbf{x}_i'\widehat{\boldsymbol{\beta}}_n > 0)\right)^{-1} \sum_{i=1}^n I(\mathbf{x}_i'\widehat{\boldsymbol{\beta}}_n > 0)\ I(0 < Y_i^+ - \mathbf{x}_i'\widehat{\boldsymbol{\beta}}_n \le h)$. Note that this is similar to that suggested by Powell (1984).

Write $d_n := \max_{1\le i\le n} ||\mathbf{x}_{in}||$.

Theorem 2.2. *Assume that* (A1)$-$(A5) *are satisfied and* $\boldsymbol{\beta}$ *is the true parameter. Then* $\widehat{S}_n$ *is a consistent estimate of* S_n*. Further, if*

$$h_n \to 0, \quad d_n/h_n \to 0, \tag{2.5}$$

$$h_n \sum_{i=1}^n I(\mathbf{x}_i'\boldsymbol{\beta} > 0) \to \infty \quad \text{as} \quad n \to \infty, \tag{2.6}$$

$$\lim_{n\to\infty} \sum_{i=1}^n I(|\mathbf{x}_i'\boldsymbol{\beta}| \le \delta_n) / \sum_{i=1}^n I(\mathbf{x}_i'\boldsymbol{\beta} > 0) = 0 \quad \text{for any} \quad \delta_n > 0, \tag{2.7}$$

then $\widehat{f}_n(0) \to f(0)$ *in probability as* $n \to \infty$.

Remark 2.1. Condition (2.6) involves the unknown parameter $\boldsymbol{\beta}$. In general, it cannot be used to choose h_n. However, it is easy to show that if all the conditions of Theorem 2.2 hold except (2.6), then Condition (2.6) is equivalent to

$$h_n \sum_{i=1}^n I(\mathbf{x}_i'\widehat{\boldsymbol{\beta}}_n > 0) \to \infty, \text{ in probability as } n \to \infty. \tag{2.8}$$

Moreover, Condition (2.8) can be tested computationally.

Now define $\widehat{W}_n := (\widehat{\boldsymbol{\beta}}_n - \mathbf{b}_0)'H(H'\widehat{S}_n^{-1}H)^{-1}H'(\widehat{\boldsymbol{\beta}}_n - \mathbf{b}_0)$ and $\widehat{R}_n := \boldsymbol{\xi}(\boldsymbol{\beta}_n^*)'\widehat{S}_n^{-1}\boldsymbol{\xi}(\boldsymbol{\beta}_n^*)$, and use $4\widehat{f}(0)M_n$, $4\widehat{f}(0)^2\widehat{W}_n$, $\widehat{R}_n$ and $4M_n^2/\widehat{W}_n$ as test statistics for testing H_0 against H_1.

As consequences of Theorems 2.1 and 2.2, we have the following.

Corollary 2.1. *Assume the conditions of Theorem* 2.2*. Under* H_0 *we have* $4\widehat{f}(0)M_n = 4\widehat{f}(0)^2\widehat{W}_n + o_p(1) = ||\sum_{i=1}^n I(\mu_i > 0)\ \mathrm{sgn}(e_i)H_n'\mathbf{x}_{in}||^2 + o_p(1) \xrightarrow{\mathcal{L}} \chi_q^2$, *as* $n \to \infty$.

Corollary 2.2. *Assume the conditions of Theorem* 2.1 *are satisfied. Under* H_0 *we have* $\widehat{R}_n = ||\sum_{i=1}^n I(\mu_i > 0)\ \mathrm{sgn}(e_i) H_n' \mathbf{x}_{in}||^2 + o_p(1) \xrightarrow{\mathcal{L}} \chi_q^2$, *as* $n \to \infty$, *and* $4M_n^2/\widehat{W}_n \xrightarrow{\mathcal{L}} \chi_q^2$, *as* $n \to \infty$.

3. Proof of Theorems

Suppose $0 < q < p$ and let K be a $p \times (p-q)$ matrix of rank $(p-q)$ such that $H'K = 0$. Write $K_n := S_n^{1/2} K (K' S_n K)^{-1/2}$, so $K_n' K_n = I_{p-q}$, $H_n' H_n = I_q$ and $H_n' K_n = 0$.

The following are direct corollaries of (A3), (A4) and (A5).

(A3′) For any $\sigma > 0$, there exists $\alpha > 0$ such that $\sum_{i=1}^n ||\mathbf{x}_{in}||^2 I(||\mathbf{x}_i|| > \alpha) < \sigma$ for large n.

(A4′) For any $\sigma > 0$, there exists $\delta > 0$ such that $\sum_{i=1}^n ||\mathbf{x}_{in}||^2 I(|\mu_i| \le \delta) < \sigma$ for large n.

(A3″) For any $\sigma > 0$, there exists $\alpha > 0$ such that $\sum_{i=1}^n ||K'\mathbf{x}_i||^2 I(||K'\mathbf{x}_i|| > \alpha) < \sigma \lambda(K' S_n K)$ for large n.

(A4″) For any $\sigma > 0$, there exists $\delta > 0$ such that $\sum_{i=1}^n ||K'\mathbf{x}_i||^2 I(|\mu_i| \le \delta) < \sigma \lambda(K' S_n K)$ for large n.

(A5″) $\lambda(K' S_n K)(\log n)^{-2} \to \infty$ as $n \to \infty$.

(Refer to the following Remark 3.1.)

Write $\boldsymbol{\gamma} := \mathbf{b} - \boldsymbol{\beta}$,

$$\begin{aligned} G_n(\boldsymbol{\gamma}) &:= \sum_{i=1}^n (|(\mathbf{x}_i'\mathbf{b})^+ - Y_i^+| - |(\mathbf{x}_i'\boldsymbol{\beta})^+ - Y_i^+|) \\ &= \sum_{i=1}^n (|(\mu_i + \mathbf{x}_i'\boldsymbol{\gamma})^+ - Y_i| - |\mu_i^+ - Y_i|), \end{aligned} \tag{3.1}$$

and take $\widehat{\boldsymbol{\gamma}}_n := \widehat{\boldsymbol{\beta}}_n - \boldsymbol{\beta} = \arg\min G_n(\boldsymbol{\gamma})$, $\widehat{\boldsymbol{\zeta}}_n := S_n^{1/2} \widehat{\boldsymbol{\gamma}}_n$.

In light of (3.1) and Rao and Zhao (1993), we have the following.

Lemma 3.1. *Assume that* (A1)−(A5) *are satisfied and* $\boldsymbol{\beta}$ *is the true parameter. Then*

$$2f(0) S_n^{1/2} \widehat{\boldsymbol{\gamma}}_n = \sum_{i=1}^n I(\mu_i > 0)\ \mathrm{sgn}(e_i) \mathbf{x}_{in} + o_p(1), \tag{3.2}$$

$$\begin{aligned} G_n(\widehat{\boldsymbol{\gamma}}_n) &= -f(0) \widehat{\boldsymbol{\gamma}}_n' S_n \widehat{\boldsymbol{\gamma}}_n + o_p(1) \\ &= -(4f(0))^{-1} ||\sum_{i=1}^n I(\mu_i > 0)\ \mathrm{sgn}(e_i) \mathbf{x}_{in}||^2 + o_p(1). \end{aligned} \tag{3.3}$$

Suppose $H'(\mathbf{b}-\mathbf{b}_0)=\mathbf{0}$ and $H'(\boldsymbol{\beta}-\mathbf{b}_0)=\mathbf{0}$. It is easily seen that there exists a unique $\boldsymbol{\eta}\in\mathcal{R}^{p-q}$ such that $\mathbf{b}-\boldsymbol{\beta}=K\boldsymbol{\eta}$. Write

$$\begin{aligned}G_n^*(\boldsymbol{\eta}) &:= \sum_{i=1}^n \left(\left|(\mathbf{x}_i'\mathbf{b})^+ - Y_i^+\right| - \left|(\mathbf{x}_i'\boldsymbol{\beta})^+ - Y_i^+\right|\right)\\ &= \sum_{i=1}^n \left(\left|(\mu_i+\mathbf{x}_i'K\boldsymbol{\eta})^+ - Y_i\right| - \left|\mu_i^+ - Y_i\right|\right), \quad \boldsymbol{\eta}\in\mathcal{R}^{p-q}, \end{aligned}\tag{3.4}$$

and set $\boldsymbol{\beta}_n^*-\boldsymbol{\beta}:=K\widehat{\boldsymbol{\eta}}_n$, $\widehat{\boldsymbol{\eta}}_n=\arg\min G_n^*(\boldsymbol{\eta})$. Replacing $\mathbf{x}_i$ in Lemma 3.1 by $K'\mathbf{x}_i$, and noticing that (A3)−(A5) imply (A3″)−(A5″), we have the following.

Lemma 3.2. *Assume* (A1)−(A5), $0<q<p$, *and that* H_0 *holds. Then*

$$2f(0)(K'S_nK)^{1/2}\widehat{\boldsymbol{\eta}}_n = \sum_{i=1}^n I(\mu_i>0)\ \mathrm{sgn}(e_i)K_n'\mathbf{x}_{in} + o_p(1), \tag{3.5}$$

$$\begin{aligned}G_n^*(\widehat{\boldsymbol{\eta}}_n) &= -f(0)\widehat{\boldsymbol{\eta}}_n(K'S_nK)\widehat{\boldsymbol{\eta}}_n\\ &= -(4f(0))^{-1}||\sum_{i=1}^n I(\mu_i>0)\ \mathrm{sgn}(e_i)K_n'\mathbf{x}_{in}||^2 + o_p(1).\end{aligned}\tag{3.6}$$

Remark 3.1. At (3.1), $\widehat{\boldsymbol{\gamma}}_n$ minimizes $\sum_{i=1}^n(|(\mu_i+\mathbf{x}_i'\boldsymbol{\gamma})^+-\mu_i-e_i|)$ and leads to (3.2). At (3.4), $\widehat{\boldsymbol{\eta}}_n$ minimizes $\sum_{i=1}^n\left(\left|(\mu_i+\mathbf{x}_i'K\boldsymbol{\eta})^+-\mu_i-e_i\right|\right)$. To obtain the limiting distribution of $\widehat{\boldsymbol{\eta}}_n$, we need only replace $\mathbf{x}_i$ by $K'\mathbf{x}_i$ and keep μ_i unchanged in the conditions and (3.2), just as is done in Lemma 3.2. This is why (A3″) has $K'\mathbf{x}_i$ appearing in the indicator function while (A4″) has μ_i in the indicator function. Derivation of (A3″) and (A4″) are routine.

Now, for any given $p\times 1$ unit vector $\boldsymbol{\theta}$, we consider the functions

$$g_n(\boldsymbol{\zeta}) := \sum_{i=1}^n\{I(\mu_i+\mathbf{x}_{in}'\boldsymbol{\zeta}>0)\mathrm{sgn}(\mathbf{x}_{in}'\boldsymbol{\zeta}-e_i)+I(\mu_i>0)\ \mathrm{sgn}(e_i)\}\mathbf{x}_{in}'\boldsymbol{\theta},\ \boldsymbol{\zeta}\in\mathcal{R}^p, \tag{3.7}$$

$$t_n(\boldsymbol{\zeta}) := \sum_{i=1}^n t_{ni}(\boldsymbol{\zeta}) = \sum_{i=1}^n\ \mathrm{sgn}(\mathbf{x}_{in}'\boldsymbol{\zeta}-e_i)\{I(\mu_i+\mathbf{x}_{in}'\boldsymbol{\zeta}>0)-I(\mu_i>0)\}(\mathbf{x}_{in}'\boldsymbol{\theta}). \tag{3.8}$$

Lemma 3.3. *Assume* (A1)−(A5). *Then for any constant* $C>0$ *and any unit p-vector* $\boldsymbol{\theta}$, *we have* $\sup_{||\boldsymbol{\zeta}||\le C}|Et_n(\boldsymbol{\zeta})|\to 0$ *and* $\sup_{||\boldsymbol{\zeta}||\le C}|t_n(\boldsymbol{\zeta})|\to 0$ *in probability, as* $n\to\infty$.

Lemma 3.4. *Assume* (A1)−(A5). *Then for any constant* $C>0$ *and any given unit p-vector* θ, $\sup_{||\boldsymbol{\zeta}||\le C}|g_n(\boldsymbol{\zeta})-2f(0)\boldsymbol{\zeta}'\boldsymbol{\theta}|\to 0$ *in probability as* $n\to\infty$.

The following corollary may be useful elsewhere.

Corollary 3.1. *Assume that* (A1)−(A5) *are satisfied and* $\boldsymbol{\beta}$ *is the true parameter. Then* $\sum_{i=1}^{n} I(\mathbf{x}_i'\widehat{\boldsymbol{\beta}}_n > 0)\mathrm{sgn}(\mathbf{x}_i'\widehat{\boldsymbol{\beta}}_n - Y_i^+)S_n^{-1/2}\mathbf{x}_i \to 0$ *in probability as* $n \to \infty$.

Now we are in a position to prove the theorems.

Proof of Theorem 2.1. Since $\max_{1\le i\le n} ||\mathbf{x}_{in}|| \to 0$ (refer to (5.1)) and $H_n'H_n = I_q$, by the Lindeberg Theorem we have $\sum_{i=1}^{n} I(\mu_i > 0)\ \mathrm{sgn}(e_i)H_n'\mathbf{x}_{in} \xrightarrow{\mathcal{L}} N(0, I_q)$. By (2.1), Lemmas 3.1 and 3.2, noting that $H_nH_n' + K_nK_n' = I_p$, for $0 < q < p$ we have

$$\begin{aligned}
&4f(0)M_n = 4f(0)(G_n^*(\widehat{\boldsymbol{\eta}}_n) - G_n(\widehat{\boldsymbol{\gamma}}_n))\\
&= ||\sum_{i=1}^{n} I(\mu_i > 0)\ \mathrm{sgn}(e_i)\mathbf{x}_{in}||^2 - ||\sum_{i=1}^{n} I(\mu_i > 0)\ \mathrm{sgn}(e_i)K_n'\mathbf{x}_{in}||^2 + o_p(1)\\
&= ||\sum_{i=1}^{n} I(\mu_i > 0)\ \mathrm{sgn}(e_i)H_n'\mathbf{x}_{in}||^2 + o_p(1),
\end{aligned}$$

which proves (2.2) for $0 < q < p$. By Lemma 3.1, it is easy to show that (2.2) is still true with $H_n = I_p$ for $q = p$.

Since $H'(\boldsymbol{\beta} - \mathbf{b}_0) = \mathbf{0}$, $W_n = (\widehat{\boldsymbol{\beta}}_n - \boldsymbol{\beta})'H(H'S_n^{-1}H)^{-1}H'(\widehat{\boldsymbol{\beta}}_n - \boldsymbol{\beta})$. For $0 < q \le p$ we have $4f(0)^2W_n = 4f(0)^2\widehat{\boldsymbol{\gamma}}_n'S_n^{1/2}H_nH_n'S_n^{1/2}\widehat{\boldsymbol{\gamma}}_n = ||\sum_{i=1}^{n} I(\mu_i > 0)\ \mathrm{sgn}(e_i)H_n'\mathbf{x}_{in}||^2 + o_p(1)$, and (2.3) is obtained.

To prove (2.4), for $0 < q < p$, write $\widehat{\boldsymbol{\varphi}}_n := S_n^{1/2}K\widehat{\boldsymbol{\eta}}_n = K_n(K'S_nK)^{1/2}\widehat{\boldsymbol{\eta}}_n$. By (3.5), $2f(0)\widehat{\boldsymbol{\varphi}}_n = \sum_{i=1}^{n} I(\mu_i > 0)\mathrm{sgn}(e_i)K_nK_n'\mathbf{x}_{in} + o_p(1)$. Now $\mathbf{x}_i'\boldsymbol{\beta}_n^* = \mathbf{x}_i'\boldsymbol{\beta} + x_i'K\widehat{\boldsymbol{\eta}}_n = \mu_i + \mathbf{x}_{in}'\widehat{\boldsymbol{\varphi}}_n$, and $S_n^{-1/2}\boldsymbol{\xi}(\boldsymbol{\beta}_n^*) = \sum_{i=1}^{n} I(\mu_i + \mathbf{x}_{in}'\widehat{\boldsymbol{\varphi}}_n > 0)\ \mathrm{sgn}(\mathbf{x}_{in}'\widehat{\boldsymbol{\varphi}}_n - e_i)\mathbf{x}_{in}$. Evidently for any $\varepsilon > 0$, we can take a constant $C > 0$ such that $P(||\widehat{\boldsymbol{\varphi}}_n|| > C) < \varepsilon$ for $n \ge n_0$. Then, with Lemma 3.4, $I(||\widehat{\boldsymbol{\varphi}}_n|| \le C)||S_n^{-1/2}\boldsymbol{\xi}(\boldsymbol{\beta}_n^*) + \sum_{i=1}^{n} I(\mu_i > 0)\ \mathrm{sgn}(e_i)\mathbf{x}_{in} - 2f(0)\widehat{\boldsymbol{\varphi}}_n|| \to 0$ in probability, which implies that $S_n^{-1/2}\boldsymbol{\xi}(\boldsymbol{\beta}_n^*) = -\sum_{i=1}^{n} I(\mu_i > 0)\ \mathrm{sgn}(e_i)\mathbf{x}_{in} + 2f(0)\widehat{\boldsymbol{\varphi}}_n + o_p(1)$. Since $I_p - K_nK_n' = H_nH_n'$ and $H_n'H_n = I_q$, we have

$$\begin{aligned}
R_n &= \boldsymbol{\xi}'(\boldsymbol{\beta}_n^*)S_n^{-1}(\boldsymbol{\beta}_n^*)\\
&= ||(I_p - K_nK_n')\sum_{i=1}^{n} I(\mu_i > 0)\ \mathrm{sgn}(e_i)\mathbf{x}_{in}||^2 + o_p(1)\\
&= ||\sum_{i=1}^{n} I(\mu_i > 0)\ \mathrm{sgn}(e_i)H_n'\mathbf{x}_{in}||^2 + o_p(1),
\end{aligned}$$

which proves (2.4) for $0 < q < p$. It is easily seen that (2.4) is still true with $H_n = I_p$ when $q = p$. Theorem 2.1 is proved.

Proof of Theorem 2.2. Since $h_n/d_n \to \infty$ and $d_n \to 0$, there exist $c_n \to \infty$ such that $h_n/(c_nd_n) \to \infty$ and $c_nd_n \to 0$.

First we prove the consistency of $\widehat{S}_n$. Recall that $\widehat{\boldsymbol{\zeta}}_n = S_n^{1/2}\widehat{\boldsymbol{\gamma}}_n$. By Lemma 3.1,

$$2f(0)\widehat{\boldsymbol{\zeta}}_n \xrightarrow{\mathcal{L}} N(0, I_p), \quad \text{and} \quad P(||\widehat{\boldsymbol{\zeta}}_n|| \geq c_n) \to 0, \quad \text{as } n \to \infty. \tag{3.9}$$

We can write $S_n^{-1/2}\widehat{S}_n S_n^{-1/2} = \sum_{i=1}^n I(\mu_i + \mathbf{x}_{in}'\widehat{\boldsymbol{\zeta}}_n > 0)\mathbf{x}_{in}\mathbf{x}_{in}'$ and, for any unit p-vector $\boldsymbol{\theta}$, by (A4$'$) we have

$$\begin{aligned} &I(||\widehat{\boldsymbol{\zeta}}_n|| \leq c_n)\Big|\boldsymbol{\theta}'\left(S_n^{-1/2}\widehat{S}_n S_n^{-1/2} - I_p\right)\boldsymbol{\theta}\Big| \\ &\leq I(||\widehat{\boldsymbol{\zeta}}_n|| \leq c_n)\Big|\boldsymbol{\theta}'\sum_{i=1}^n (I(\mu_i + \mathbf{x}_{in}'\widehat{\boldsymbol{\zeta}}_n > 0) - I(\mu_i > 0))\mathbf{x}_{in}\mathbf{x}_{in}'\boldsymbol{\theta}\Big| \\ &\leq \sum_{i=1}^n I(|\mu_i| \leq c_n d_n)(\mathbf{x}_{in}'\boldsymbol{\theta})^2 \to 0, \quad \text{as } n \to \infty. \end{aligned} \tag{3.10}$$

Then consistency of $\widehat{S}_n$ follows from (3.9) and (3.10).

To prove the consistency of $\widehat{f}_n(0)$, write $N_n = \sum_{i=1}^n I(\mu_i > 0)$, $\widehat{N}_n = \sum_{i=1}^n I(\mathbf{x}_i'\widehat{\boldsymbol{\beta}}_n > 0)$, $\widetilde{f}_n(0) := (hN_n)^{-1}\sum_{i=1}^n I(\mathbf{x}_i'\widehat{\boldsymbol{\beta}}_n > 0)I(0 < e_i - \mathbf{x}_{in}'\widehat{\boldsymbol{\zeta}}_n \leq h)$, $f_n^*(0) := (hN_n)^{-1}\sum_{i=1}^n I(\mathbf{x}_i'\widehat{\boldsymbol{\beta}}_n > 0)I(0 < e_i \leq h)$, and $f_n(0) := (hN_n)^{-1}\sum_{i=1}^n I(\mu_i > 0)I(0 < e_i \leq h)$. Since $\mathbf{x}_i'\widehat{\boldsymbol{\beta}}_n = \mu_i + \mathbf{x}_{in}'\widehat{\boldsymbol{\zeta}}_n$, by (2.7) we have

$$\begin{aligned} &I(||\widehat{\boldsymbol{\zeta}}_n|| \leq c_n)|\widehat{N}_n - N_n|/N_n \leq I(||\widehat{\boldsymbol{\zeta}}_n|| \leq c_n)\sum_{i=1}^n I(|\mu_i| \leq |\mathbf{x}_{in}'\widehat{\boldsymbol{\zeta}}_n|)/N_n \\ &\leq \sum_{i=1}^n I(|\mu_i| \leq c_n d_n)\Big(\sum_{i=1}^n I(\mu_i > 0)\Big)^{-1} \to 0. \end{aligned} \tag{3.11}$$

By (3.9), (3.11), (2.5) and (2.6), we have $\widetilde{f}_n(0)/\widehat{f}_n(0) = \widehat{N}_n/N_n \to 1$ in probability as $n \to \infty$.

Now

$$\begin{aligned} &I(||\widehat{\boldsymbol{\zeta}}_n|| \leq c_n)\left|\widetilde{f}_n(0) - f_n^*(0)\right| \\ &\leq I(||\widehat{\boldsymbol{\zeta}}_n|| \leq c_n)(hN_n)^{-1}\sum_{i=1}^n I(\mathbf{x}_i'\widehat{\boldsymbol{\beta}}_n > 0)\{I(|e_i| \leq c_n d_n) + I(h < e_i \leq h + c_n d_n)\} \\ &\leq (hN_n)^{-1}\sum_{i=1}^n \{I(\mu_i > 0) + I(|\mu_i| \leq c_n d_n)\}\{I(|e_i| \leq c_n d_n) \\ &\qquad\qquad + I(h < e_i \leq h + c_n d_n)\} := J_n. \end{aligned}$$

Since $h_n \to 0$, and $c_n d_n/h \to 0$, by (A1) and (2.7) we have

$$\begin{aligned} EJ_n = (N_n)^{-1}\{N_n + \sum_{i=1}^n I(|\mu_i| \leq c_n d_n)\} \\ \times\{P(|e_1| \leq c_n d_n)/h + P(h < e_1 \leq h + c_n d_n)/h\} \to 0, \end{aligned}$$

which, by (3.9), implies that $\widetilde{f}_n(0)-f_n^*(0) \to 0$ in probability as $n \to \infty$. Similarly, we have $f_n^*(0) - f_n(0) \to 0$ in probability as $n \to \infty$. By (A1) and (2.5), it is easily shown that $f_n(0) \to f(0)$ in probability as $n \to \infty$. The result follows.

4. A Simulation Study

We perform simulations to study properties of the proposed tests with practical sample sizes. For simplicity, p is fixed to be 2 at (1.2), and e_1 is taken to have the Laplace distribution with density function $(2c)^{-1}\exp(-|u|/c)$. The true value of the parameter vector $\boldsymbol{\beta}$ is set to be $(0,1)'$ and the linear hypothesis under consideration is $H_0 : H'\boldsymbol{\beta} = \mathbf{0}$ with $H' = (1,0)$. The test statistics $4\widehat{f}(0)M_n$, $4\widehat{f}^2(0)\widehat{W}_n$, $\widehat{R}_n$ and $4M_n^2/\widehat{W}_n$ are abbreviated to T_1, T_2, T_3 and T_4, respectively. To generate $\mathbf{x}_i = (x_i^{(1)}, x_i^{(2)})$, $i = 1, \ldots, n$, let $P(x^{(1)} = 1) = P(x^{(1)} = 0) = 1/2$, and take $x^{(2)}$ standard normal independent of $x^{(1)}$. Independent copies of $x^{(1)}$ and $x^{(2)}$ yield observations on $\mathbf{x} = (x^{(1)}, x^{(2)})$. In our simulations, sample size n is taken to be 50, 100 and 200, respectively, and $1,000$ repetitions are generated for each sample size.

To find minimization solutions $\widehat{\boldsymbol{\beta}}_n$ and $\boldsymbol{\beta}_n^*$ in (2.1), we use an iterative linear programming algorithm (see Buchinsky (1994)). Given a complete sample, to estimate a probability density the choice of window width h can be done by using the cross-validation method. However, this may be a difficult task for our situation. We will study these problems in the future. For this simple simulation, we only list a number of h's which are taken for estimating $f(0)$. Note that the choice of h depends on the sample size n and the scale of the distribution of random error e_i.

First we study distribution approximation of the four test statistics by simulation. To this end, we set $c = 0.1$ and take $\mathbf{x}_i$, $i = 1, \ldots, n$, to be fixed in all the $1,000$ samples. For all cases, we take $h = 0.04$. Under the true model, the functions $P(T_i \le \kappa(t))$ are plotted against t in Figure 1$-$3, where $\kappa(\cdot)$ is the quantile function of the χ_1^2 distribution. The figures show that the distributions of the four test statistics under the true model are close to the limiting χ_1^2 distribution even for moderate sample sizes, with the approximations for T_1 and T_2 better than the others.

Denote by $\chi_1^2(\alpha)$ the upper α-quantile of the limiting χ_1^2 distribution. We can define the T_i test by specifying its rejection region as $\{T_i > \chi_1^2(\alpha)\}$, where α is the approximate or nominal level of the T_i test, $i = 1, \ldots, 4$. To further study true levels (or type 1 errors) and power values under some alternative for the above four tests, we set $c = 1$ and take covariates to be random in all the samples, which is a more realistic case. In order to estimate $f(0)$ for construction of T_1 and T_2, we choose $h = 0.2, 0.3, \ldots, 0.7$, respectively, and list the relevant simulation results. In Table 1 we list $P(T_i > \chi_1^2(\alpha))$ under the true model vs.

α for four test statistics. The corresponding power values under the alternative $\beta = (1, 0)'$ are reported in Table 2. The relevant simulation results in Tables 1 and 2 have been rounded to two decimals.

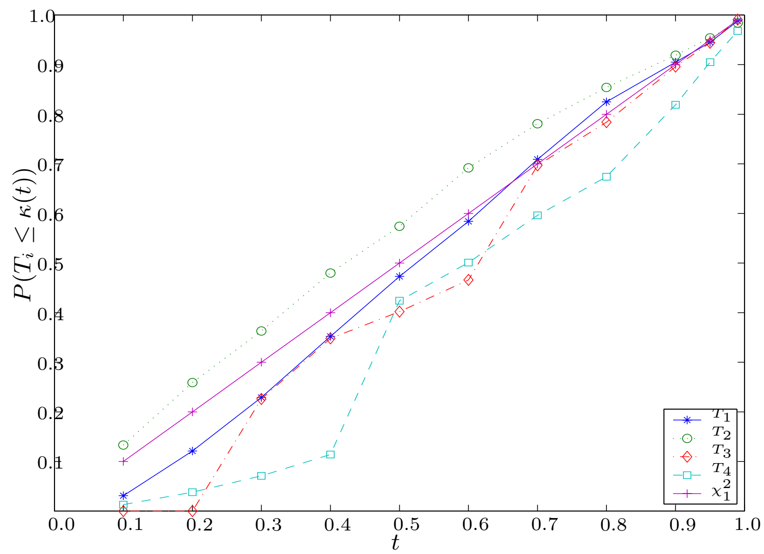

Figure 1. $P(T_i \le \kappa(t))$ plot for $n = 50$. $h = 0.04$ for T_1 and T_2; $c = 0.1$ and covariates are non-random; $\kappa(\cdot)$ is the quantile function of the χ_1^2 distribution.

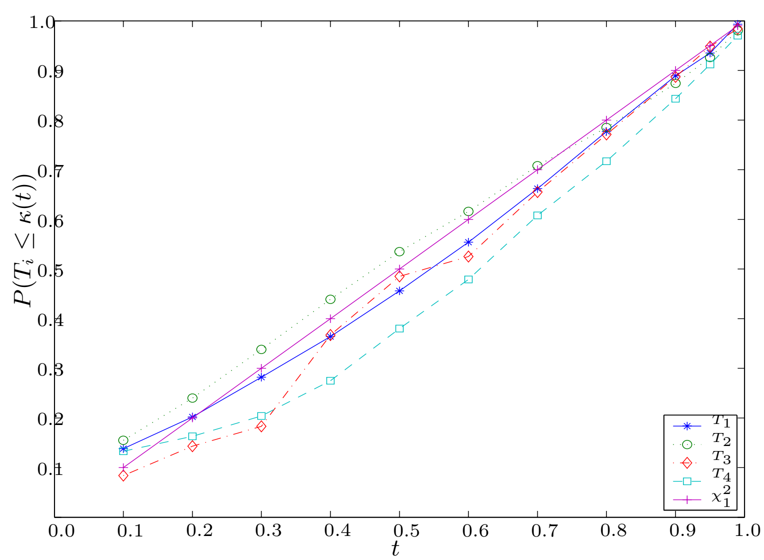

Figure 2. $P(T_i \le \kappa(t))$ plot for $n = 100$. $h = 0.04$ for T_1 and T_2; $c = 0.1$ and covariates are non-random; $\kappa(\cdot)$ is the quantile function of the χ_1^2 distribution.

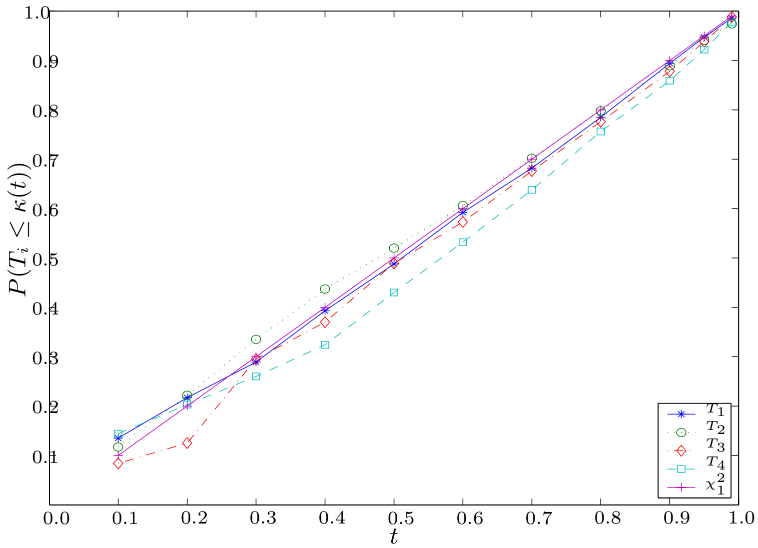

Figure 3. $P(T_i \leq \kappa(t))$ plot for $n = 200$. $h = 0.04$ for T_1 and T_2; $c = 0.1$ and covariates are non-random; $\kappa(\cdot)$ is the quantile function of the χ_1^2 distribution.

From the simulation results, it seems that convergence rates of the distribution for T_1 and T_2 are better than those of the other two in our simulations. The simulation also shows that the performance of estimate of $f(0)$ depends greatly on the choice of h.

Table 1. Empirical levels ($c = 1$).

n	α	T_1						T_2						T_3	T_4
		0.2 *	0.3 *	0.4 *	0.5 *	0.6 *	0.7 *	0.2 *	0.3 *	0.4 *	0.5 *	0.6 *	0.7 *		
50	0.10	0.14	0.11	0.10	0.09	0.08	0.07	0.18	0.14	0.12	0.09	0.08	0.07	0.11	0.21
	0.05	0.09	0.06	0.06	0.05	0.04	0.03	0.13	0.10	0.07	0.06	0.05	0.04	0.07	0.14
	0.01	0.03	0.02	0.01	0.01	0.01	0.01	0.07	0.05	0.03	0.02	0.02	0.01	0.03	0.06
100	0.10	0.12	0.10	0.09	0.08	0.07	0.06	0.13	0.11	0.09	0.08	0.07	0.06	0.12	0.19
	0.05	0.06	0.05	0.05	0.04	0.04	0.03	0.09	0.07	0.06	0.05	0.04	0.03	0.06	0.12
	0.01	0.02	0.01	0.01	0.01	0.01	0.00	0.05	0.03	0.02	0.01	0.01	0.01	0.02	0.04
200	0.10	0.11	0.09	0.08	0.07	0.06	0.06	0.11	0.09	0.07	0.06	0.05	0.03	0.11	0.17
	0.05	0.05	0.04	0.04	0.03	0.03	0.03	0.07	0.05	0.04	0.02	0.02	0.01	0.07	0.09
	0.01	0.01	0.01	0.01	0.00	0.00	0.00	0.02	0.01	0.01	0.01	0.01	0.00	0.02	0.03

* 0.2−0.7 are values of h, n is sample size, α is nominal level. The results are based on 1,000 repetitions. Covariates are random.

Table 2. Power values under alternative $\boldsymbol{\beta} = (1,0)'$ ($c = 1$).

n	α	T_1						T_2						T_3	T_4
		0.2 *	0.3 *	0.4 *	0.5 *	.6 *	0.7 *	0.2 *	0.3 *	0.4 *	.5 *	0.6 *	0.7 *		
50	0.10	0.92	0.93	0.93	0.94	0.93	0.94	0.90	0.91	0.92	0.93	0.93	0.93	0.22	0.90
	0.05	0.90	0.90	0.90	0.90	0.90	0.90	0.86	0.88	0.89	0.89	0.90	0.89	0.09	0.84
	0.01	0.79	0.80	0.79	0.79	0.78	0.77	0.79	0.81	0.80	0.80	0.78	0.78	0.01	0.71
100	0.10	0.99	1.00	1.00	1.00	1.00	1.00	0.99	1.00	1.00	1.00	1.00	1.00	0.54	0.99
	0.05	0.99	0.99	1.00	0.99	0.99	0.99	0.99	0.99	1.00	1.00	1.00	1.00	0.32	0.99
	0.01	0.99	0.99	1.00	0.99	0.99	0.99	0.96	0.98	0.98	0.99	0.98	0.98	0.06	0.96
200	0.10	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.90	1.00
	0.05	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.76	1.00
	0.01	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.36	1.00

Rejection region is $\{T_i > \chi_1^2(\alpha)\}$ with $\chi_1^2(\alpha)$ the upper α-quantile of χ_1^2 distribution.
* 0.2−0.7 are values of h, n is sample size, α is nominal level. The results are based on 1,000 repetitions. Covariates are random.

5. Appendix

Proof of Lemma 3.3. Since $\lambda(S_n) \to \infty$ and (A3′) holds, it is easily shown that

$$d_n := \max_{1\le i\le n} ||\mathbf{x}_{in}|| \to 0 \text{ as } n \to \infty. \tag{5.1}$$

By (3.8), $t_{ni}(\boldsymbol{\zeta})$ can be written as a sum of those terms of the form $\pm t_{ni}^{j\pm}(\boldsymbol{\zeta})$, $j = 1,\ldots,4$, where $x^- = |x|I(x<0)$ and $t_{ni}^{1\pm}(\boldsymbol{\zeta}) := I(x_{in}'\boldsymbol{\zeta} > e_i, -\mathbf{x}_{in}'\boldsymbol{\zeta} < \mu_i \le 0)(\mathbf{x}_{in}'\boldsymbol{\theta})^{\pm}$, $t_{ni}^{2\pm}(\boldsymbol{\zeta}) := I(\mathbf{x}_{in}'\boldsymbol{\zeta} < e_i, -\mathbf{x}_{in}'\boldsymbol{\zeta} < \mu_i \le 0)(\mathbf{x}_{in}'\boldsymbol{\theta})^{\pm}$, $t_{ni}^{3\pm}(\boldsymbol{\zeta}) := I(\mathbf{x}_{in}'\boldsymbol{\zeta} > e_i, 0 < \mu_i \le -\mathbf{x}_{in}'\boldsymbol{\zeta})(\mathbf{x}_{in}'\boldsymbol{\theta})^{\pm}$, $t_{ni}^{4\pm}(\boldsymbol{\zeta}) := I(\mathbf{x}_{in}'\boldsymbol{\zeta} < e_i, 0 < \mu_i \le -\mathbf{x}_{in}'\boldsymbol{\zeta})(\mathbf{x}_{in}'\boldsymbol{\theta})^{\pm}$.

Write $\mathcal{T}_n^{j\pm} := \{(t_{ni}^{j\pm}(\boldsymbol{\zeta}),\ 1 \le i \le n),\ \boldsymbol{\zeta} \in \mathcal{R}^p\}$, $j = 1,2,3,4$. It is easily shown that, if $h_{ni}^{\pm}(\boldsymbol{\zeta}) = I(\mathbf{x}_{in}'\boldsymbol{\zeta} < a_i)(\mathbf{x}_{in}'\boldsymbol{\theta})^{\pm}$ with a_i being a random variable or real number, and $\mathcal{H}_n^{\pm} := \left\{\left(h_{ni}^{\pm}(\boldsymbol{\zeta}), 1 \le i \le n\right), \boldsymbol{\zeta} \in \mathcal{R}^p\right\}$, then $\mathcal{H}_n^{\pm}$ has pseudo-dimension at most p according to the sense of Pollard (1990, pp.14-22). By Lemma 5.1 in Pollard (1990), if both $\mathcal{F}$ and $\mathcal{G}$ have pseudo-dimension at most V, then $\mathcal{F} \wedge \mathcal{G} := \{(f \wedge g) : f \in \mathcal{F},\ g \in \mathcal{G}\}$ has pseudo-dimension less than $10V$, where $f \wedge g = \min(f,g)$. From this, all $\mathcal{T}_n^{j\pm}$ have bounded pseudo-dimensions. By the maximal inequality for manageable processes (Pollard (1990, 7.10), we have

$$E \sup_{||\boldsymbol{\zeta}||\le C} |t_n(\boldsymbol{\zeta}) - Et_n(\boldsymbol{\zeta})|^2 \le C_1 \sum_{i=1}^n (\mathbf{x}_{in}'\boldsymbol{\theta})^2 I(|\mu_i| \le Cd_n) \to 0. \tag{5.2}$$

Here (5.1) and (A4′) are used, and $C_1 > 0$ is a constant.

By (A1), (5.1) and (A4′), there exist constants C_2 and C_3 such that

$$\sup_{||\boldsymbol{\zeta}||\leq C} |Et_n(\boldsymbol{\zeta})| = 2 \sup_{||\boldsymbol{\zeta}||\leq C} \Big| \sum_{i=1}^{n} \Big(F(\mathbf{x}_{in}'\boldsymbol{\zeta}) - \frac{1}{2} \Big) \{ I(\mu_i + \mathbf{x}_{in}'\boldsymbol{\zeta} > 0) - I(\mu_i > 0) \} \mathbf{x}_{in}'\boldsymbol{\theta} \Big|$$

$$\leq C_2 \sup_{||\boldsymbol{\zeta}||\leq C} \sum_{i=1}^{n} I(|\mu_i| \leq C d_n) |(\mathbf{x}_{in}'\boldsymbol{\zeta})(\mathbf{x}_{in}'\boldsymbol{\theta})|$$

$$\leq C_3 \sum_{i=1}^{n} ||\mathbf{x}_{in}||^2 I(|\mu_i| \leq C d_n) \to 0 \quad \text{as } n \to \infty, \tag{5.3}$$

and Lemma 3.3 follows from (5.2) and (5.3).

To prove Lemma 3.4, we need the following.

Lemma A.1. *Let E be an open convex subset of $\mathcal{R}^p$ and let $g_1, g_2, \ldots$ be a sequence of random convex functions on E such that for any $\mathbf{x} \in E$, $g_n(\mathbf{x}) \to g(\mathbf{x})$* a.s. (or in probability) *as $n \to \infty$, where g is some real function on E. Then g is also convex. Furthermore, assume that $\partial g(\mathbf{x}), \partial g_1(\mathbf{x}), \partial g_2(\mathbf{x}), \ldots$ are sub-gradients of $g, g_1, g_2, \ldots$ at $\mathbf{x}$, and $\partial g(\mathbf{x})$ is continuous on E. Then for all compact $D \subset E$, $\lim_{n\to\infty} \sup_{\mathbf{x}\in D} ||\partial g_n(\mathbf{x}) - \partial g(\mathbf{x})|| = 0$ a.s. (or in probability).*

Note that we call $\partial g(\mathbf{x})$ a sub-gradient of g at $\mathbf{x}$ if for all $\mathbf{z} \in E$, $g(\mathbf{z}) - g(\mathbf{x}) \geq (\partial g(\mathbf{x}))'(\mathbf{z} - \mathbf{x})$. For a proof, refer to Lemmas 4.2 and 4.3 in Heiler and Willers (1988). For the case of convergence in probability, a diagonal technique should be used.

Proof of Lemma 3.4. Write $v_n(\boldsymbol{\zeta}) = \sum_{i=1}^n I(\mu_i > 0)\{ \operatorname{sgn}(\mathbf{x}_{in}'\boldsymbol{\zeta} - e_i) + \operatorname{sgn}(e_i)\}\mathbf{x}_{in}$. By (3.7), (3.8) and Lemma 3.3, to prove Lemma 3.4 we need only prove

$$\sup_{||\boldsymbol{\zeta}||\leq C} ||v_n(\boldsymbol{\zeta}) - 2f(0)\boldsymbol{\zeta}|| \to 0 \quad \text{in probability,} \quad \text{as } n \to \infty. \tag{5.4}$$

Define $V_n(\boldsymbol{\zeta}) = \sum_{i=1}^n I(\mu_i > 0) \int_0^{\mathbf{x}_{in}'\boldsymbol{\zeta}} (\operatorname{sgn}(v - e_i) + \operatorname{sgn}(e_i)) dv$. $V_n(\boldsymbol{\zeta})$ is convex in $\boldsymbol{\zeta}$ and has a sub-gradient $v_n(\boldsymbol{\zeta})$ at $\boldsymbol{\zeta}$. By (A1) and (5.1),

$$EV_n(\boldsymbol{\zeta}) = 2 \sum_{i=1}^{n} I(\mu_i > 0) \int_0^{\mathbf{x}_{in}'\boldsymbol{\zeta}} (F(v) - 1/2) dv$$

$$= \sum_{i=1}^{n} I(\mu_i > 0) \int_0^{\mathbf{x}_{in}'\boldsymbol{\zeta}} 2f(0)v(1 + o(1)) dv \to f(0)\boldsymbol{\zeta}'\boldsymbol{\zeta}. \tag{5.5}$$

By the Schwarz inequality, (A1) and (5.1),

$$\operatorname{Var} V_n(\boldsymbol{\zeta}) \leq \sum_{i=1}^{n} I(\mu_i > 0) E\Big[\int_0^{\mathbf{x}_{in}'\boldsymbol{\zeta}} (\operatorname{sgn}(v - e_i) + \operatorname{sgn}(e_i)) dv \Big]^2$$

$$\leq \sum_{i=1}^{n} I(\mu_i > 0)|\mathbf{x}_{in}'\boldsymbol{\zeta}| \cdot \Big|4\int_0^{\mathbf{x}_{in}'\boldsymbol{\zeta}} E(I(e_i > 0) - I(e_i > v))^2 dv\Big|$$
$$\leq 4\sum_{i=1}^{n} I(\mu_i > 0)|\mathbf{x}_{in}'\boldsymbol{\zeta}| \cdot \Big|\int_0^{\mathbf{x}_{in}'\boldsymbol{\zeta}} P(|e_i| \leq |v|) dv\Big|$$
$$\leq C_4\sum_{i=1}^{n} I(\mu_i > 0)|\mathbf{x}_{in}'\boldsymbol{\zeta}|^3 \leq C_5 d_n \to 0, \tag{5.6}$$

where C_4 and C_5 are constants. By (5.5) and (5.6), for each $\boldsymbol{\zeta}$ we have $V_n(\boldsymbol{\zeta}) \to f(0)\boldsymbol{\zeta}'\boldsymbol{\zeta}$ in probability, as $n \to \infty$, which, by the convexity of $V_n(\boldsymbol{\zeta})$ and Lemma 5.1, implies (5.4). The lemma is proved.

Proof of Corollary 3.1. With Lemma 3.1 $\mathbf{x}_i'\widehat{\boldsymbol{\beta}}_n = \mu_i + \mathbf{x}_{in}'\widehat{\boldsymbol{\zeta}}_n$ and $||\widehat{\boldsymbol{\zeta}}_n|| = O_p(1)$ as $n \to \infty$. From (5.4), $\sum_{i=1}^n I(\mu_i > 0)\{\ \mathrm{sgn}(\mathbf{x}_{in}'\widehat{\boldsymbol{\zeta}}_n - e_i) +\ \mathrm{sgn}(e_i)\}\mathbf{x}_{in} - 2f(0)\widehat{\boldsymbol{\zeta}}_n \to 0$ in probability which, in view of Lemma 3.1, implies that $\sum_{i=1}^n I(\mu_i > 0)\ \mathrm{sgn}(\mathbf{x}_{in}'\widehat{\boldsymbol{\zeta}}_n - e_i)\mathbf{x}_{in} \to 0$ in probability as $n \to \infty$. By Lemma 3.3, $\sum_{i=1}^n \{I(\mathbf{x}_i'\widehat{\boldsymbol{\beta}}_n > 0) - I(\mu_i > 0)\}\ \mathrm{sgn}(\mathbf{x}_{in}'\widehat{\boldsymbol{\zeta}}_n - e_i)\mathbf{x}_{in} \to 0$ in probability. It then follows that $\sum_{i=1}^n I(\mathbf{x}_i'\widehat{\boldsymbol{\beta}}_n > 0)\ \mathrm{sgn}(\mathbf{x}_i'\widehat{\boldsymbol{\beta}}_n - Y_i)\mathbf{x}_{in} \to 0$ in probability as $n \to \infty$, which implies the result.

Acknowledgement

The research was partially supported by the National Natural Science Foundation of China (10171094), Ph.D. Program Foundation of Ministry of Education of China and Special Foundations of the Chinese Academy of Sciences and USTC.

The author would like to express his thanks to the referee, an associate editor and the Chair Editor for their valuable comments, which led to this much improved version of the paper. Thanks are also given to Mr. H. Zhang, who carried out the simulations.

References

Bilias, Y., Chen, S. N. and Ying, Z. (2000). Simple resampling methods for censored regression quantiles. *J. Econometrics* **99**, 373-386.

Buchinsky, M.(1994). Changes in the U. S. wage structure 1963-1987: application of quantile regression. *Econometrica* **62**, 405-458.

Buchinsky, M. (1998). Recent advances in quantile regression models. *The Journal of Human Resources* **27**, 88-126.

Buchinsky, M. and Hahn, J. (1998). An alternative estimator for the censored quantile regression model. *Econometrica* **66**, 653-671.

Chamberlain, G. (1994). Quantile regression, censoring, and the structure of wages. In: Sims, C. (Ed.), *Advances in Econometrics.* Sixth World Congress, **1**, 171-209. Canbridge University Press, New York.

Chay, K. and Honoré, B. (1998). Estimation of semiparametric censored regression models: an application to changes in black-white earnings inequality during the 1960s. *J. Human Resources* **27**, 4-38.

Chen, X. R. and Wu, Y. (1994). Consistency of L_1 estimates in censored linear regression models. *Comm. Statist. Theory Method* **23**, 1847-1858.

Fitzenberger, B. (1997). A guide to censored quantile regressions. In *Handbook of Statistics, Robust Inference* **15** (Edited by G. S. Maddala and C. R. Rao), 405-437. North-Holland, Amsterdam.

Heiler, S. and Willers, R. (1988). Asymptotic normality of R-estimates in the linear model. *Statistics* **19**, 173-184.

Maddala, G. S. (1983). *Limited Dependent and Qualitative Variables in Econometrics.* Cambridge University Press, Cambridge.

Pollard, D. (1990). *Empirical Processes: Theory and Applications.* NSF-CBMS Regional Conference Series in Probability and Statistics, Volume 2.

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

Powell, J. L. (1986). Censored regression quantiles. *J. Econometrics* **32**, 143-145.

Rao, C. R. and Zhao, L. C. (1993). Asymptotic normality of LAD estimators in censored regression models. *Math. Methods Statist.* **2**, 228-239.

Department of Statistics and Finance, University of Science and Technology of China, Hefei, Anhui, 230026, China.

E-mail: lczhao@ustc.edu.cn

(Received March 2002; accepted May 2003)