The Annals of Statistics
1991, Vol. 19, No. 2, 531–556

RANK REGRESSION METHODS FOR LEFT-TRUNCATED AND RIGHT-CENSORED DATA[1]

BY TZE LEUNG LAI AND ZHILIANG YING

Stanford University and University of Illinois

A class of rank estimators is introduced for regression analysis in the presence of both left-truncation and right-censoring on the response variable. By making use of martingale theory and a tightness lemma for stochastic integrals of multiparameter empirical processes, the asymptotic normality of the estimators is established under certain assumptions. Adaptive choice of the score functions to give asymptotically efficient rank estimators is also discussed.

1. Introduction. Consider the problem of estimating the slope β in the regression model

$$y_i = \beta x_i + \varepsilon_i \qquad (i = 1, 2, \dots), \tag{1.1}$$

where ε_i are i.i.d. random variables with a continuous distribution function F, the x_i are either nonrandom or are independent random variables that are independent of $\{\varepsilon_n\}$ and the responses y_i are not completely observable due to left-truncation and right-censoring by the random variables t_i and c_i specified later. Throughout the sequel we shall assume that (t_i, c_i, x_i) are independent random vectors that are independent of the sequence $\{\varepsilon_n\}$ and such that $\infty > t_i \geq -\infty$ and $-\infty < c_i \leq \infty$. We shall also let $\tilde{y}_i = y_i \wedge c_i$ and $\Delta_i = I_{\{y_i \leq c_i\}}$, where we use $\wedge$ and $\vee$ to denote minimum and maximum, respectively. Thus, the responses y_i are right-censored by the censoring variables c_i. Moreover, we shall also assume left-truncation in the sense that $(\tilde{y}_i, \Delta_i, x_i)$ can be observed only when $\tilde{y}_i \geq t_i$. The data, therefore, consist of n observations $(\tilde{y}_i^0, t_i^0, \Delta_i^0, x_i^0)$ with $\tilde{y}_i^0 \geq t_i^0$, $i = 1, \dots, n$.

The special case $t_i \equiv -\infty$ corresponds to the censored regression model [cf. Kalbfleisch and Prentice (1980); Lawless (1982)], for which there is an extensive literature on hypothesis testing and estimation of β when the ε_i are assumed to belong to certain parametric families of distributions. The case $c_i \equiv \infty$ corresponds to the truncated regression model in the econometrics literature [cf. Tobin (1958); Goldberger (1981); Amemiya (1985)], which assumes the presence of truncation variables τ_i so that (x_i, y_i) can be observed only when $y_i \leq \tau_i$ (or equivalently, when $-y_i \geq -\tau_i = t_i$) and which uses the

Received March 1988; revised June 1990.
[1]Research supported by NSF.
AMS 1980 *subject classifications*. Primary 62J99, 62G20; secondary 60F05.
Key words and phrases. Censoring and truncation, regression, linear rank statistics, martingales, stochastic integrals, empirical processes, tightness, asymptotic normality, adaptive rank estimators.

method of maximum likelihood to estimate β in various parametric models for the ε_i.

Without assuming any parametric form for the distribution of the ε_i, we study herein rank regression methods for the estimation of β. In the case of complete data (for which $t_i \equiv -\infty$ and $c_i \equiv \infty$, so that $\tilde{y}_i^0 = y_i$ and $x_i^0 = x_i$), letting $e_i(b) = y_i - bx_i$ and $R_j(b)$ be the rank of $e_j(b)$ in the set of residuals $\{e_1(b), \dots, e_n(b)\}$, one can define the linear rank statistic

$$L_n(b) = \sum_{i=1}^{n} x_i a_n(R_i(b)), \tag{1.2a}$$

where the scores $a_n(j)$ are generated from a score function ϕ: $(0,1] \to (-\infty, \infty)$, that satisfies $\int_0^1 \phi(u)\,du = 0$ and $\int_0^1 \phi^2(u)\,du < \infty$, by

$$\begin{aligned} a_n(j) &= \phi(j/n), \\ a_n(j) &= \phi(j/(n+1)) \qquad \text{or} \\ a_n(j) &= E\phi(U_{(j)}^n), \end{aligned} \tag{1.2b}$$

$U_{(1)}^n \le \cdots \le U_{(n)}^n$ being the order statistics of a sample of size n from the uniform distribution on $(0,1)$. There is an extensive literature on the use of $L_n(0)$ as test statistics for testing the null hypothesis $\beta = 0$ and on the use of a zero-crossing of $L_n(b)$ as an estimate of β, cf. Hájek and Šidák (1967) and Jurečková (1969). We say that $\hat{b}$ is a zero-crossing of a step function $L(b)$ if the right- and left-hand limits $L(\hat{b}+)$ and $L(\hat{b}-)$ do not have the same sign, that is, if $L(\hat{b}+)L(\hat{b}-) \le 0$.

To extend these rank regression methods to the truncated and censored data $(\tilde{y}_i^0, t_i^0, \Delta_i^0, x_i^0)$, $i = 1, \dots, n$, define the residuals $e_i(b) = \tilde{y}_i^0 - bx_i^0$ and let $e_{(1)}(b) \le \cdots \le e_{(k)}(b)$ denote all the ordered uncensored residuals. For $i = 1, \dots, k$, let

$$\begin{aligned} J(i,b) &= \{j \le n : t_j^0 - bx_j^0 \le e_{(i)}(b) \le \tilde{y}_j^0 - bx_j^0\}, \\ n_i(b) &= \#J(i,b), \qquad \bar{x}(i,b) = \frac{\sum_{j \in J(i,b)} x_j^0}{n_i(b)}, \end{aligned} \tag{1.3}$$

where we use the notation $\#A$ to denote the number of elements of a set A. Let ψ be a twice continuously differentiable function on $(0,1)$. Let $0 < \varepsilon < 1$ and let p_n be a smooth function on $[0,1]$ such that $p_n(0) = 0$ and $p_n(x) = 1$ for $x \ge n^{-\varepsilon}$, as will be specified in Section 2. Define

$$1 - \hat{F}_{n,b}(u) = \prod_{i:\, e_{(i)}(b) < u,\, \Delta_{(i)}^0 = 1} \left\{1 - \frac{p_n(n^{-1}n_i(b))}{n_i(b)}\right\}, \tag{1.4}$$

$$S_n(b) = \sum_{i=1}^{k} \psi\big(\hat{F}_{n,b}(e_{(i)}(b))\big) p_n(n^{-1}n_i(b)) \{x_{(i)}^0 - \bar{x}(i,b)\}. \tag{1.5}$$

Recalling that for complete data a rank estimator of β is obtained as a zero-crossing of $L_n(b)$, we define in the present setting, in which the responses y_i are subject to left-truncation and right-censoring, a rank estimator $\hat{\beta}_n$ of β as a zero-crossing of $S_n(b)$.

For data subject only to right-censoring (i.e., $t_i \equiv -\infty$), Prentice (1978) and Cuzick (1985) have shown that the statistics (1.5) with $b = \beta_0$, $p_n \equiv 1$ and

$$\psi = \phi - \Phi, \quad \text{where } \Phi(u) = \frac{1}{1-u}\int_u^1 \phi(t)\,dt, \tag{1.6}$$

provide natural and analytically tractable extensions of the classical linear rank statistics (1.2) (with $b = \beta_0$) for testing the null hypothesis $\beta = \beta_0$. Recently, Tsiatis (1990), Ritov (1990) and Lai and Ying (1989) studied the use of such censored rank statistics for the estimation of β. In particular, Lai and Ying (1989) introduced the weight functions p_n to smooth out the jittery behavior of $\hat{F}_{n,b}(e_{(i)}(b))$ and $\bar{x}(i, b)$ as functions of b at those i for which the risk set size $n_i(b)$ is small compared to n. This enables them to apply the tightness theory for stochastic integrals of empirical-type processes, developed in Lai and Ying (1988) for right-censored data, to analyze the random function $S_n(b)$ in the censored case. For data subject only to left-truncation (i.e., $c_i \equiv \infty$), Lai and Ying (1989) have shown that the statistics (1.5) with ψ defined by (1.6) again provide natural and tractable extensions of the linear rank statistics (1.2).

In Section 2 we unify the results and methods for the censored regression model and the truncated regression model by generalizing them to data that are both right-censored and left-truncated, as previously described. Such data are often encountered in biostatistical applications. A basic tool in the development of this unified theory is a simple extension, given in Section 4, of the tightness results of Lai and Ying (1988) for stochastic integrals of certain empirical-type processes. Using this tightness lemma together with martingale theory, consistency and asymptotic normality of the rank estimators based on left-truncated and right-censored data are established under certain assumptions.

In Sections 3, we develop asymptotically efficient adaptive rank estimators in which the score function ψ in (1.5) is not predetermined in advance but is chosen on the basis of the observed data $(\tilde{y}_i^0, t_i^0, \Delta_i^0, x_i^0)$, $i = 1, \dots, n$. Such adaptive rank estimators can again be analyzed by making use of martingale theory and the tightness lemma of stochastic integrals of empirical-type processes. As will be shown in Section 3, these methods to construct and analyze rank estimators and adaptive rank estimators can also be extended to the case of $\nu \times 1$ vectors β and x_i, so that (1.1) with $\beta x_i = \beta^T x_i$ represents a multiple regression model.

2. Consistency and asymptotic normality of rank estimators. In this section we study the asymptotic properties of the estimate $\hat{\beta}_n$ of β, defined as a zero-crossing of the linear rank statistics $S_n(b)$ given by (1.5) for left-truncated and right-censored data. Throughout the sequel we shall restrict b to a bounded interval $[-\rho, \rho]$, assuming knowledge of an upper bound $\rho > |\beta|$. A basic step in the analysis of $S_n(b)$ is to represent it as a stochastic integral of empirical-type processes. First note that the sample $(\tilde{y}_i^0, t_i^0, \Delta_i^0, x_i^0)$, $i = 1, \dots, n$, of left-truncated and right-censored observations can be regarded as being generated by a larger sample of independent random vectors (y_i, t_i, c_i, x_i), $i = 1, \dots, m(n)$, where

$$m(n) = \inf\left\{m: \sum_{i=1}^{m} I_{\{t_i \le y_i \wedge c_i\}} = n\right\}. \tag{2.1}$$

To represent the linear rank statistics $S_n(b)$ given in (1.5) as stochastic integrals, define

$$t_i(b) = t_i - bx_i, \qquad c_i(b) = c_i - bx_i, \qquad y_i(b) = y_i - bx_i, \tag{2.2}$$

$$L_m(b, s) = \sum_{i=1}^{m} I_{\{t_i(b) \le y_i(b) \wedge c_i(b) \le s,\, y_i(b) \le c_i(b)\}}, \tag{2.3}$$

$$Y_m(b, s) = \sum_{i=1}^{m} x_i I_{\{t_i(b) \le y_i(b) \wedge c_i(b) \le s,\, y_i(b) \le c_i(b)\}}, \tag{2.4}$$

$$N_m(b, s) = \sum_{i=1}^{m} I_{\{t_i(b) \le s \le y_i(b) \wedge c_i(b)\}}, \tag{2.5}$$

$$X_m(b, s) = \sum_{i=1}^{m} x_i I_{\{t_i(b) \le s \le y_i(b) \wedge c_i(b)\}}, \tag{2.6}$$

$$\begin{aligned} &\log\left(1 - \hat{F}_{m,n,b}(y)\right) \\ &\quad = \int_{-\infty < s < y} \log\left\{1 - \frac{p_n\left(n^{-1}N_m(b, s)\right)}{N_m(b, s)}\right\} L_m(b, ds), \end{aligned} \tag{2.7}$$

$$\begin{aligned} V_{m,n}(b, s) &= \psi\left(\hat{F}_{m,n,b}(s)\right) p_n\left(n^{-1}N_m(b, s)\right), \\ \tilde{V}_{m,n} &= \frac{V_{m,n} X_m}{N_m}, \end{aligned} \tag{2.8}$$

$$T_{m,n}(b) = \int_{-\infty}^{\infty} V_{m,n}(b, s) Y_m(b, ds) - \int_{-\infty}^{\infty} \tilde{V}_{m,n}(b, s) L_m(b, ds). \tag{2.9}$$

From (1.4) and (1.5), it follows that

$$S_n(b) = T_{m(n),n}(b). \tag{2.10}$$

The following lemma approximates the stochastic integral $T_{m,n}(b)$ by replacing the empirical-type processes L_m, Y_m, N_m, X_m that appear in $T_{m,n}(b)$ by their expectations. As will be shown in Section 4, the lemma follows from a simple extension of the tightness results of Lai and Ying (1988), for stochastic integrals involving these empirical-type processes, under the following assumptions:

$$|x_i| \le B \quad \text{for all } i \text{ and some nonrandom constant } B, \tag{2.11}$$

F has a continuously differentiable density f such that

$$\int_{-\infty}^{\infty}\Big(\sup_{s\le t\le s+\eta}|f'(t)|\Big)\,ds < \infty \quad \text{for some } \eta > 0, \tag{2.12}$$

$$\begin{aligned}&\sup_{|b|\le\rho,\,-\infty<s<\infty}\sum_{1}^{m}P\{s\le t_i-bx_i\le s+h\}\\ &\qquad+\sup_{|b|\le\rho,\,-\infty<s<\infty}\sum_{1}^{m}P\{s\le c_i-bx_i\le s+h\}\\ &= O(mh) \text{ as } m\to\infty \text{ and } h\to 0 \text{ such that } mh\to\infty,\end{aligned} \tag{2.13}$$

$$\sup_i\Big\{E(c_i^-)^q + E\big(t_i^- I_{\{t_i>-\infty\}}\big)^q\Big\} < \infty \quad \text{for some } q > 0, \tag{2.14}$$

where $x^- = |x|I_{\{x\le 0\}}$. The role of these assumptions in the tightness results for the two-parameter process $L_m(b,s)$ or Y_m, N_m, X_m and for integrals with respect to the signed measure $Y_m(b,ds)$ or $L_m(b,ds)$ is discussed in the censored case (with $t_i \equiv -\infty$) by Lai and Ying [(1988), pages 342–343 and 347–348]. Note in particular that by (2.12),

$$\begin{aligned}EL_m(b,ds) &= \sum_{i=1}^{m}E\big[f(s+(b-\beta)x_i)I_{\{t_i-bx_i\le s\le c_i-bx_i\}}\big]\,ds,\\ EY_m(b,ds) &= \sum_{i=1}^{m}E\big[x_i f(s+(b-\beta)x_i)I_{\{t_i-bx_i\le s\le c_i-bx_i\}}\big]\,ds.\end{aligned} \tag{2.15}$$

LEMMA 1. *Let ψ be twice continuously differentiable on* $(0,1)$ *and such that*

$$\sup_{0<u<1}|\psi''(u)| < \infty. \tag{2.16}$$

Let p be a nondecreasing and twice continuously differentiable function on the real line such that

$$p(y) = 0 \ \text{for } y\le 0 \quad \text{and} \quad p(y) = 1 \ \text{for } y \ge 1. \tag{2.17}$$

For $m \ge n \ge 1$, define L_m, Y_m, N_m, X_m, $\hat F_{m,n,b}$, $V_{m,n}$, $\tilde V_{m,n}$ and $T_{m,n}$ by (2.3)–(2.9), *where the weight function p_n is of the form*

$$p_n(x) = p\big(n^{\lambda}(x - cn^{-\lambda})\big), \qquad 0\le x\le 1, \tag{2.18}$$

with $c > 0$ *and* $0 < \lambda < \frac{1}{18}$. *Define*

$$\Lambda_{m,n}(b, y) = \int_{-\infty < s < y} \left[\frac{p_n(n^{-1}EN_m(b, s))}{EN_m(b, s)} \right] EL_m(b, ds), \tag{2.19}$$

$$\begin{aligned} \phi_{m,n}(b) = \int_{-\infty}^{\infty} \psi(1 - e^{-\Lambda_{m,n}(b,s)}) p_n(n^{-1}EN_m(b, s)) \\ \times \left[EY_m(b, ds) - \frac{EX_m(b, s)}{EN_m(b, s)} EL_m(b, ds) \right]. \end{aligned} \tag{2.20}$$

Then under the assumptions (2.11)–(2.14), *for every* $D \ge 1$,

$$\max_{n \le m \le Dn} \sup_{|b| \le \rho} \frac{1}{n} |T_{m,n}(b) - \phi_{m,n}(b)| \to 0 \quad a.s., \tag{2.21}$$

$$\max_{n \le m \le Dn} \sup_{|b| \le \rho} \frac{|T_{m,n}(b) - T_{m,n}(\beta) - \phi_{m,n}(b) + \phi_{m,n}(\beta)|}{n^{1/2} \vee n|b - \beta|} \to 0 \quad a.s. \tag{2.22}$$

Lemma 1 provides a basic tool for the analysis of the random function $S_n(b)$ whose zero-crossing is the linear rank estimator $\hat{\beta}_n$ of β. Since $S_n(b)$ is not a smooth function of b, one cannot apply standard techniques (based on Taylor's expansion of the random function in a neighborhood of the true parameter) that are commonly used to prove asymptotic normality of maximum likelihood estimates, M-estimators, and so on. Moreover, $S_n(b)$ is not a monotone function of b [cf. Lai and Ying (1989)], so one cannot make use of the monotonicity and contiguity arguments [cf. Jurečková (1969)] that have been applied to prove asymptotic normality of rank estimators of β in the regression model (1.1) based on complete data (x_i, y_i). Lemma 1 enables us to approximate $S_n(b)$ by $S_n(\beta) + [\phi_{m(n),n}(b) - \phi_{m(n),n}(\beta)]$, which is much more tractable than $S_n(b)$.

To analyze the random variable $S_n(\beta)$ and the nonrandom function $\phi_{m(n),n}(b) - \phi_{m(n),n}(\beta)$, we begin by considering the case in which the (t_i, c_i, x_i) are i.i.d. and $t_i - \beta x_i$ is independent of $c_i - \beta x_i$. For $r = 0, 1, 2$, let

$$G_r(s) = E\left[x_1^r I_{\{t_1 - \beta x_1 \le s \le c_1 - \beta x_1\}} \right], \qquad \overline{G}(s) = P\{t_1 - \beta x_1 \le c_1 - \beta x_1 < s\}.$$

Define

$$\tau_0 = \inf\{s : G_0(s) > 0\}, \qquad \tau = \inf\{s > \tau_0 : (1 - F(s))G_0(s) = 0\}. \tag{2.23}$$

Suppose that $F(\tau_0) < 1$ (so $-\infty \le \tau_0 < \tau \le \infty$) and that

$$\int_{\tau_0 - \eta}^{\tau + \eta} \frac{\sup_{|t| \le \eta} f(s + t)}{1 - F(s)} dF(s) < \infty \quad \text{for some } \eta > 0. \tag{2.24}$$

Then for every $0 < \delta < 1$, it can be shown that uniformly in $m \ge n \ge \delta m$,

$$\phi_{m,n}(b) - \phi_{m,n}(\beta) = mA(b - \beta) + o(m^{1/2} \vee m|b - \beta|) \quad \text{as } m \to \infty \text{ and } b \to \infty, \tag{2.25}$$

where

$$A = \int_{\tau_0}^{\tau} \psi(\tilde{F}(s)) \left\{ \frac{f'(s)}{f(s)} + \frac{f(s)}{1 - F(s)} \right\} \left\{ G_2(s) - \frac{G_1^2(s)}{G_0(s)} \right\} dF(s), \tag{2.26}$$

$$\tilde{F}(s) = \frac{F(s) - F(\tau_0)}{1 - F(\tau_0)} = P\{\varepsilon_1 \leq s | \varepsilon_1 \geq \tau_0\}, \qquad s \geq \tau_0.$$

Moreover, it can be shown that $m(n)/n \to K$ a.s., where

$$\frac{1}{K} = \int_{-\infty}^{\infty} \{G_0(s) + \overline{G}(s)\} \, dF(s), \tag{2.27}$$

and that $n^{-1/2} S_n(\beta)$ has a limiting normal distribution with mean 0 and variance

$$v = K \int_{\tau_0}^{\tau} \psi^2(\tilde{F}(s)) \left\{ G_2(s) - \frac{G_1^2(s)}{G_0(s)} \right\} dF(s). \tag{2.28}$$

In fact, these results still hold when (t_i, c_i, x_i) are not identically distributed and $t_i - \beta x_i$ is not independent of $c_i - \beta x_i$, provided that the functions $\overline{G}$ and G_r $(r = 0, 1, 2)$ that appear in (2.23), (2.26) and (2.27) now take the form of (2.29). This is the content of:

THEOREM 1. *Suppose that in the regression model* (1.1), *the* ε_i *are i.i.d. random variables with a common distribution function* F *and* (t_i, c_i, x_i), $i = 1, 2, \ldots$, *are independent random vectors that are independent of* $\{\varepsilon_n\}$ *and such that* (2.11)–(2.14) *hold. Let* ψ *be a twice continuously differentiable function on* $(0, 1)$ *satisfying* (2.16) *and let* p *be a nondecreasing and continuously differentiable function on the real line satisfying* (2.17). *Define* $t_i(b)$, $c_i(b)$, $y_i(b)$ *by* (2.2), L_m, Y_m, N_m, X_m, $\hat{F}_{m,n,b}$, $V_{m,n}$, $\tilde{V}_{m,n}$ *and* $T_{m,n}$ *by* (2.3)–(2.9), *where the weight function* p_n *is of the form* (2.18) *with* $c > 0$ *and* $0 < \lambda < \frac{1}{18}$. *Assume furthermore that for* $r = 0, 1, 2$,

$$\lim_{m \to \infty} \frac{1}{m} \sum_{i=1}^{m} E\{x_i^r P[t_i - \beta x_i \leq s \leq c_i - \beta x_i | x_i]\} = G_r(s), \tag{2.29a}$$

$$\lim_{m \to \infty} \frac{1}{m} \sum_{i=1}^{m} P\{t_i - \beta x_i \leq c_i - \beta x_i < s\} = \overline{G}(s) \tag{2.29b}$$

exist for every s *with* $F(s) < 1$. *Defining* τ_0 *and* τ *by* (2.23), *assume that* $\tau_0 < \tau$, *that* F *satisfies* (2.24) *and that*

$$\lim_{m \to \infty} \frac{1}{m^{1-\lambda}} \sum_{i=1}^{m} \left[P\{t_i - \beta x_i < \tau_0 - \varepsilon\} I_{\{F(\tau_0) > 0\}} + P\{c_i - \beta x_i > \tau + \varepsilon\} I_{\{F(\tau) < 1\}} \right] = 0 \tag{2.30}$$

for every $\varepsilon > 0$,

where λ *is the same as in* (2.18). *Let* $S_n(b) = T_{m(n),n}(b)$, *where* $m(n)$ *is defined in* (2.1). *Then*

(i) $\sup_{|b| \le \rho} n^{-1}|S_n(b) - \phi_{m(n),n}(b)| \to 0$ *a.s., where* $\phi_{m(n),n}(b)$ *is given in* (2.20). *Moreover,* $m(n)/n \to K$ *a.s., where* K *is defined in* (2.27).

(ii) *Define* A *as in* (2.26). *Then with probability* 1,

$$S_n(b) - S_n(\beta) = nKA(b - \beta) + o\left(n^{1/2} \vee n|b - \beta|\right) \quad \text{as } n \to \infty \text{ and } b \to \beta. \tag{2.31}$$

(iii) *As* $n \to \infty$, $n^{-1/2}S_n(\beta)$ *has a limiting normal distribution with mean* 0 *and variance* v *given in* (2.28).

REMARK. The assumption (2.30) is clearly satisfied when (t_i, c_i, x_i) are i.i.d. and $t_i - \beta x_i$ is independent of $c_i - \beta x_i$. It is needed to give the basic limiting relations (2.37) and (2.41) that lead to the definitions of A and v in (2.26) and (2.28).

PROOF OF THEOREM 1(i)–(ii). Since $S_n(b) = T_{m(n),n}(b)$,

$$\sup_{|b| \le \rho} \frac{1}{n}\left|S_n(b) - \phi_{m(n),n}(b)\right| \to 0 \quad \text{a.s.}$$

by (2.21). By Kolmogorov's strong law of large numbers,

$$\frac{1}{m}\sum_{i=1}^{m}\left[I_{\{t_i \le y_i \wedge c_i\}} - P\{t_i \le y_i \wedge c_i\}\right] \to 0 \quad \text{a.s.} \tag{2.32}$$

By (1.1) and (2.29),

$$\begin{aligned}\frac{1}{m}\sum_1^m P\{t_i \le y_i \wedge c_i\} &= \frac{1}{m}\sum_1^m \left[P\{t_i - \beta x_i \le \varepsilon_i \le c_i - \beta x_i\}\right.\\ &\qquad \left.+P\{t_i - \beta x_i \le c_i - \beta x_i < \varepsilon_i\}\right]\\ &\to \int_{-\infty}^{\infty}\left[G_0(s) + \overline{G}(s)\right] dF(s) = \frac{1}{K}.\end{aligned} \tag{2.33}$$

From (2.1), (2.32) and (2.33), it follows that $m(n)/n \to K$ a.s.

To prove (2.31), it suffices in view of (2.22) to show that (2.25) holds, uniformly in $\delta m \le n \le m$ for every $0 < \delta < 1$, recalling that $S_n(b) = T_{m(n),n}(b)$ and that $m(n)/n \to K$ a.s. Let $\overline{F} = 1 - F$. By (2.5), (2.13) and (2.29),

$$\begin{aligned}EN_m(b,s) &= \sum_1^m E\left\{\overline{F}(s + (b - \beta)x_i)P\left[t_i - bx_i \le s \le c_i - bx_i \middle| x_i\right]\right\}\\ &\le \sum_1^m P\{t_i - bx_i \le s\},\end{aligned} \tag{2.34}$$

$$\frac{1}{m}EN_m(b,s) \to \overline{F}(s)G_0(s) \quad \text{as } m \to \infty \text{ and } b \to \beta. \tag{2.35}$$

In view of (2.18) and (2.19),

(2.36) $p_n(x) = 0 \quad \text{if } x \le cn^{-\lambda}, \qquad p_n(x) = 1 \quad \text{if } x \ge (c+1)n^{-\lambda}.$

From (2.13), (2.15), (2.30) and (2.34)–(2.36), it follows that for any $\tau^* \in (\tau_0, \tau)$,

$$\Lambda_{m,n}(b,y) = \int_{-\infty<s<y} \frac{p_n(n^{-1}EN_m(b,s))}{EN_m(b,s)} EL_m(b,ds)$$
$$\to \int_{\tau_0}^{y} \frac{f(s)}{\bar{F}(s)}\, ds$$

as $m \to \infty$ and $b \to \beta$, uniformly in $y \in [\tau_0, \tau^*]$ and $m \ge n \ge \delta m$ for every $0 < \delta < 1$. Therefore, by the definition of $\tilde{F}$ in (2.26),

(2.37) $1 - \exp(-\Lambda_{m,n}(b,y)) \to \tilde{F}(y)$ as $m \to \infty$ and $b \to \beta$, uniformly in $y \in [\tau_0, \tau^*]$ and $m \ge n \ge \delta m$.

To prove (2.25), let $H_i(x_i, s) = P[t_i - \beta x_i \le s \le c_i - \beta x_i | x_i]$ and assume without loss of generality that $\beta = 0$. By (2.34),

$$EN_m(b,s) = \sum_1^m E\{\bar{F}(s+bx_i)H_i(x_i, s+bx_i)\},$$
$$EX_m(b,s) = \sum_1^m E\{x_i\bar{F}(s+bx_i)H_i(x_i, s+bx_i)\},$$
$$EL_m(b,ds) = \sum_1^m E\{f(s+bx_i)H_i(x_i, s+bx_i)\}\,ds$$

by (2.15), since $\beta = 0$. Define

$$g_{m,n}(a,b) = \int_{-\infty}^{\infty} \psi(1-e^{-\Lambda_{m,n}(b,s)})p_n(n^{-1}EN_m(b,s)) \times \sum_1^m E\left\{\left[x_i - \frac{E\Sigma_1^m x_j\bar{F}(s+ax_j)H_j(x_j, s+bx_j)}{E\Sigma_1^m \bar{F}(s+ax_j)H_j(x_j, s+bx_j)}\right] \times f(s+ax_i)H_i(x_i, s+bx_i)\right\} ds.$$

Note that $\phi_{m,n}(b) = g_{m,n}(b,b)$ and that $g_{m,n}(0,b) = 0$. Therefore

$$\phi_{m,n}(b) - \phi_{m,n}(0) = \phi_{m,n}(b) = g_{m,n}(b,b) - g_{m,n}(0,b)$$
$$= Amb + o(m^{1/2} \vee mb) \quad \text{as } m \to \infty \text{ and } b \to 0,$$

uniformly in $m \ge n \ge m/(2K)$. The last relation follows by applying (2.24), (2.30), (2.36), (2.37) and the Taylor expansions

$$\bar{F}(s+bx_i) = \bar{F}(s) - bx_i f(s) + O(b^2) \text{ and}$$
$$f(s+bx_i) - f(s) = bx_i f'(s) + O(b^2)$$

to $g_{m,n}(b,b) - g_{m,n}(0,b)$, recalling that $\sup_i |x_i| \le B$ a.s. The technical details of the argument are similar to those given in the Appendix of Lai and Ying (1989) for the censored case. □

The proof of Theorem 1(iii) makes use of Rebolledo's (1980) martingale central limit theorem and the martingale structure of $S_n(\beta)$. This martingale structure is a corollary of the following lemma of Lai and Ying (1991), which is a generalization of the well-known martingale theory for right-censored data [cf. Gill (1980)] and of its extension to left-truncated data [cf. Keiding and Gill (1990)].

LEMMA 2. *Let $\varepsilon_1, \varepsilon_2, \dots$ be i.i.d. random variables with a continuous distribution function F and let $\Lambda = -\log(1 - F)$ denote the cumulative hazard function of F. Let (x_i, T_i, C_i), $i = 1, 2, \dots,$ be independent random vectors that are independent of $\{\varepsilon_n\}$. Let $\tilde{\varepsilon}_i = \varepsilon_i \wedge C_i$, $\Delta_i = I_{\{\varepsilon_i \le C_i\}}$ and let $\mathscr{F}(s)$ be the complete σ-field generated by*

$$x_i, T_i, I_{\{T_i \le \tilde{\varepsilon}_i\}}, \Delta_i I_{\{T_i \le \tilde{\varepsilon}_i \le s\}}, I_{\{T_i \le u \le \tilde{\varepsilon}_i\}}, I_{\{T_i \le \tilde{\varepsilon}_i \le u\}}, \qquad u \le s, \quad i = 1, 2, \dots.$$

Define

$$M_i(s) = I_{\{T_i \le \varepsilon_i \le s \wedge C_i\}} - \int_{-\infty}^{s} I_{\{\varepsilon_i \wedge C_i \ge u \ge T_i\}}\, d\Lambda(u). \tag{2.38}$$

Then $\{M_i(s), \mathscr{F}(s), -\infty < s < \infty\}$ is a martingale with predictable variation process

$$\langle M_i \rangle(s) = \int_{-\infty}^{s} I_{\{\varepsilon_i \wedge C_i \ge u \ge T_i\}}\, d\Lambda(u). \tag{2.39}$$

PROOF OF THEOREM 1(iii). Set $T_i = t_i - \beta x_i$, $C_i = c_i - \beta x_i$ in Lemma 2. By (2.9) and (2.10), $S_n(\beta) = S_n(\beta; \infty)$, where

$$\begin{aligned} S_n(\beta; t) = \sum_{i=1}^{m(n)} \int_{-\infty}^{t} & \psi\big(\hat{F}_{m(n), n, \beta}(s)\big) p_n\big(n^{-1} N_{m(n)}(\beta, s)\big) \\ & \times \left\{ x_i - \frac{X_{m(n)}(\beta, s)}{N_{m(n)}(\beta, s)} \right\} dM_i(s), \end{aligned} \tag{2.40}$$

noting that

$$N_{m(n)}(\beta, s) = \sum_{1}^{m(n)} I_{\{\varepsilon_i \wedge C_i \ge s \ge T_i\}} \quad \text{and} \quad X_{m(n)}(\beta, s) = \sum_{1}^{m(n)} x_i I_{\{\varepsilon_i \wedge C_i \ge s \ge T_i\}}.$$

Since $m(n) = \inf\{m: \sum_1^m I_{\{T_i \le \varepsilon_i \wedge C_i\}} = n\}$, it follows from the definition of $\mathscr{F}(s)$ in Lemma 2 that $m(n)$ is $\bigcap_{s=-\infty}^{\infty} \mathscr{F}(s)$-measurable. Furthermore, $\hat{F}_{m(n), n, \beta}(s)$ is left-continuous in s. It then follows that $\{S_n(\beta; t), \mathscr{F}(t), -\infty < t < \infty\}$ is a martingale, [cf. Lai and Ying (1991)]. Moreover, as shown in Theorem 3 of Lai and Ying (1991),

$$\sup_{t < \tau} \left| \hat{F}_{m(n), n, \beta}(t) - \tilde{F}(t) \right| \to 0 \quad \text{a.s.,} \tag{2.41}$$

under the assumption (2.30). The desired conclusion then follows from Rebolledo's martingale central limit theorem [cf. Gill (1980) and Section 2 of Lai and Ying (1989)]. □

COROLLARY 1. *With the same notation and assumptions as in Theorem* 1, *suppose that*

$$\liminf_{n\to\infty}\left\{\min_{(1-\delta)Kn\le m\le(1+\delta)Kn}\left[\inf_{\substack{|b|\le\rho\\|b-\beta|\ge\delta}} n^{-1}|\phi_{m,n}(b)|\right]\right\}>0 \tag{2.42}$$

for every $0<\delta<\delta_0$

for some sufficiently small $\delta_0>0$, *where* $\phi_{m,n}$ *is given in* (2.20). *Then* $\hat{\beta}_n\to\beta$ *a.s. If furthermore* $A\neq 0$, *where A is defined in* (2.26), *then* $n^{1/2}(\hat{\beta}_n-\beta)$ *has a limiting normal distribution with mean* 0 *and variance* $(KA)^{-2}v$ *as* $n\to\infty$, *where v is given in* (2.28).

PROOF. Write $n^{-1/2}S_n(\hat{\beta}_n)=n^{-1/2}S_n(\beta)+n^{-1/2}\{S_n(\hat{\beta}_n)-S_n(\beta)\}$ and apply (2.31) and Theorem 1(iii). The details of the argument are similar to those of Corollary 2 of Lai and Ying (1989). □

3. Asymptotically efficient adaptive rank estimators and extensions to multiple regression models.

By Corollary 1, the rank estimator $\hat{\beta}_n$, defined as a zero-crossing of the linear rank statistics (1.5), is asymptotically normal $N(\beta,v/(K^2A^2n))$ as $n\to\infty$. Letting $h=f/(1-F)$ denote the hazard function of F, we can express A given in (2.26) as

$$A=\int_{\tau_0}^{\tau}\psi(\tilde{F}(s))\left[\frac{h'(s)}{h(s)}\right]\left[G_2(s)-\frac{G_1^2(s)}{G_0(s)}\right]dF(s).$$

Since $1/K=\int_{-\infty}^{\infty}(G_0+\overline{G})\,dF$ and

$$v=K\int_{\tau_0}^{\tau}\psi^2(\tilde{F}(s))\left[G_2(s)-\frac{G_1^2(s)}{G_0(s)}\right]dF(s),$$

it then follows from the Schwarz inequality that

$$\frac{1}{(KA)^2}v\ge\left\{\int_{-\infty}^{\infty}(G_0+\overline{G})\,dF\right\}\left\{\int_{\tau_0}^{\tau}\left(\frac{h'}{h}\right)^2\left(G_2-\frac{G_1^2}{G_0}\right)dF\right\}^{-1} \tag{3.1}$$

and that equality holds in (3.1) in the case

$$\psi(\tilde{F}(s))=\frac{h'(s)}{h(s)}. \tag{3.2}$$

Since h is usually unknown in practice, we study in this section how to use the observed data $(\tilde{y}_i^0,t_i^0,\Delta_i^0,x_i^0)$, $i=1,\dots,n$, to estimate the asymptotically optimal score function (3.2) for the linear rank statistics (1.5) from which we obtain an asymptotically normal rank estimator $\hat{\beta}_n$ that attains the lower bound in (3.1).

Our basic idea is to divide the sample into two disjoint subsets, the first of which is $\{(\tilde{y}_i^0, t_i^0, \Delta_i^0, x_i^0): i \le n/2\}$. From the first subsample, define the residuals $e_i(b) = \tilde{y}_i^0 - bx_i^0$ $(i \le n/2)$ and order the uncensored ones among them as $e_{(1)}(b) \le \cdots \le e_{(k_1)}(b)$. Let $n_1 = [n/2]$, that is, the largest integer $\le n/2$, and define $J(i, b), n_i(b), \bar{x}(i, b)$ as in (1.3) but with n_1 replacing n (i.e., on the basis only of the first subsample). In analogy with (1.5), define

$$S_{n,1}(b) = \sum_{i=1}^{k_1} p_n(n^{-1}n_i(b))\big[x_{(i)}^0 - \bar{x}(i,b)\big]\psi_{n,2}(e_{(i)}(b)), \tag{3.3}$$

where p_n is defined by (2.17)–(2.18) and $\psi_{n,2}(s)$ is an estimate of $h'(s)/h(s)$ defined later from the second subsample of $n_2 = n - n_1$ observations $(\tilde{y}_r^0, t_r^0, \Delta_r^0, x_r^0)$, $n_1 < r \le n$. Likewise from the second subsample, define the residuals $e_i^*(b) = \tilde{y}_{n_1+i}^0 - bx_{n_1+i}^0$ $(i \le n_2)$ and order the uncensored ones among them as $e_{[1]}^*(b) \le \cdots \le e_{[k_2]}^*(b)$. As in (1.3), let $J^*(i, b) = \{n_1 < r \le n: t_r^0 - bx_r^0 \le e_{[i]}^*(b) \le \tilde{y}_r^0 - bx_r^0\}$, $n_i^*(b) = \#J^*(i, b)$, $\bar{x}^*(i, b) = (\sum_{r \in J^*(i,b)} x_r^0)/n_i^*(b)$. Define

$$S_{n,2}(b) = \sum_{i=1}^{k_2} p_n(n^{-1}n_i^*(b))\big[x_{[i]}^* - \bar{x}^*(i,b)\big]\psi_{n,1}(e_{[i]}^*(b)), \tag{3.4}$$

where $x_{[i]}^*$ denotes the covariate $x_{n_1+[i]}^0$ corresponding to $e_{[i]}^*(b)$ and $\psi_{n,1}(s)$ is an estimate of $h'(s)/h(s)$ defined later from the first subsample. Combining the two subsample statistics (3.3) and (3.4) gives the linear rank statistic

$$S_n^*(b) = S_{n,1}(b) + S_{n,2}(b). \tag{3.5}$$

An *adaptive rank estimator* β_n^* of β is defined as a zero-crossing of $S_n^*(b)$.

The analysis of the random function $S_n^*(b)$ and of its zero-crossing β_n^* uses the same basic ideas as those used in Section 2 for the analysis of $S_n(b)$ and $\hat{\beta}_n$. The first step is to extend the asymptotic linearity property (2.31) to $S_n^*(b) - S_n^*(\beta)$ and the second step is to establish the asymptotic normality of $S_n^*(\beta)$. Since the random function $\psi_{n,2}(s)$ is based entirely on the second subsample, $S_{n,1}(\beta)$ defined by (3.3) in terms of the first subsample values and the function $\psi_{n,2}(s)$ should still have the martingale structure used in the proof of Theorem 1(iii). Likewise, $S_{n,2}(\beta)$ should also have this martingale structure. In view of this martingale structure no matter how $\psi_{n,j}$ is constructed from the jth subsample $(j = 1, 2)$, there is considerable flexibility in choosing $\psi_{n,j}$. In particular, we shall choose $\psi_{n,j}$ to be sufficiently smooth so that Lemma 1 and the asymptotic linearity property (2.31) can be extended to $S_{n,j}(b)$. Moreover, (3.1) and (3.2) suggest that asymptotic efficiency of the corresponding adaptive rank estimator may be achieved by choosing $\psi_{n,j}$ such that

$$\sup_{s \in [s_1, s_2]:\, h(s) \ge d_n} \left| \psi_{n,j}(s) - \frac{h'(s)}{h(s)} \right| \to_P 0 \tag{3.6}$$

for all $s_1 > \tau_0$ and $s_1 < s_2 < \tau$, with $0 < d_n \to 0$ and

$$\int_{-\infty}^{s_1} \psi_{n,j}^2(s)\, dF(s) \to_P 0 \quad \text{and} \quad \int_{s_2}^{\infty} \psi_{n,j}^2(s)\, dF(s) \to_P 0 \tag{3.7}$$

as $n \to \infty$ and $s_1 \downarrow \tau_0$, $s_2 \uparrow \tau$.

Motivated by these considerations, we introduce the following data-dependent score function $\psi_{n,j}$ ($j = 1, 2$). First compute from the jth subsample a consistent estimator $b_{n,j}$ of β such that

$$b_{n,j} \to \beta \quad \text{a.s.} \quad \text{and} \quad b_{n,j} - \beta = O_P(n^{-d}) \quad \text{for some } d > 0. \tag{3.8}$$

For example, we can choose a smooth score function ψ satisfying (2.16) and use it to define the rank estimator $b_{n,j}$ as a zero-crossing of the rank statistics computed from the jth subsample by (1.4) and (1.5) (with $i \le k_1$ in the case $j = 1$ and with similar changes for $j = 2$). Then under the assumptions of Corollary 1, (3.8) holds with $d = \frac{1}{2}$. Let w be a twice continuously differentiable nonnegative function on the real line such that for some $\alpha > 0$,

$$w(t) = 0 \quad \text{for } t \notin (-\alpha, \alpha), \qquad \int_{-\alpha}^{\alpha} w(t)\, dt = 1. \tag{3.9}$$

Let δ_n be positive constants such that

$$\delta_n \to 0 \quad \text{and} \quad 1/\delta_n = o(n^{\varepsilon}) \quad \text{for every } \varepsilon > 0. \tag{3.10}$$

For example, take $\delta_n \sim (\log n)^{-1}$. Using w as the kernel and δ_n as the bandwidth, define kernel estimates $\hat{h}_{n,1}, \hat{h}_{n,2}$ of h based on the two subsamples separately by

$$\begin{aligned} \hat{h}_{n,1}(t) &= \delta_n^{-1} \sum_{i=1}^{k_1} w\frac{\big(\{t - e_{(i)}(b_{n,1})\}/\delta_n\big) p\big(\{n^{-1} n_i(b_{n,1}) - \delta_n\}/\delta_n\big)}{n_i(b_{n,1})}, \\ \hat{h}_{n,2}(t) &= \delta_n^{-1} \sum_{i=1}^{k_2} w\frac{\big(\{t - e^*_{[i]}(b_{n,2})\}/\delta_n\big) p\big(\{n^{-1} n^*_i(b_{n,2}) - \delta_n\}/\delta_n\big)}{n^*_i(b_{n,2})}, \end{aligned} \tag{3.11}$$

where p is the smooth function satisfying (2.17) that has been used in the definition of p_n in (3.3) and (3.4). Making use of (3.8), it will be shown in Lemma 6 that under certain assumptions on h,

$$\sup_s \left| \frac{\hat{h}'_{n,j}(s)}{\hat{h}_{n,j}(s)} - \frac{h'(s)}{h(s)} \right| I_{\{\hat{h}_{n,j}(s) \ge \delta_n^{1/4},\, n^{-1}\#_{n,j}(s) \ge \delta_n^{1/5}\}} \to_P 0, \qquad j = 1, 2, \tag{3.12}$$

where $\#_{n,1}(s) = \#\{i \le n_1\colon t_i^0 - b_{n,1}x_i^0 \le s \le \tilde{y}_i^0 - b_{n,1}x_i^0\}$, $\#_{n,2}(s) = \#\{n_1 < i \le n\colon t_i^0 - b_{n,2}x_i^0 \le s \le \tilde{y}_i^0 - b_{n,2}x_i^0\}$. Define for $j = 1, 2$,

$$\begin{aligned} \tilde{\psi}_{n,j}(s) &= \frac{\hat{h}'_{n,j}(s)}{\hat{h}_{n,j}(s)} I_{\{\hat{h}_{n,j}(s) \ge \delta_n^{1/4},\, n^{-1}\#_{n,j}(s) \ge \delta_n^{1/5}\}}, \\ \psi_{n,j}(s) &= \left\{\delta_n^{-1} \int_{-\infty}^{\infty} w\big((s-u)/\delta_n\big) \tilde{\psi}_{n,j}(u)\, du\right\} p\big(2 - \delta_n s^2\big). \end{aligned} \tag{3.13}$$

Thus, $\psi_{n,j}$ is a smoothed version of $\tilde{\psi}_{n,j}$. Making use of (3.12), we shall show that the $\psi_{n,j}$ thus constructed satisfies (3.6) and (3.7) in Lemma 6 that will be used to prove the following.

THEOREM 2. *With the same assumptions on* $(\varepsilon_i, t_i, c_i, x_i)$ *and on* p, p_n *as in Theorem* 1, *assume further that* f *is twice continuously differentiable with* $\sup_x |f''(x)| < \infty$ *and*

$$\int_{\tau_0-\eta}^{\tau+\eta} \sup_{|t|\le\eta} \left[\frac{f'(s+t)}{f(s)}\right]^2 dF(s) < \infty \quad \text{for some } \eta > 0. \tag{3.14}$$

Let w *be a twice continuously differentiable nonnegative function on the real line satisfying* (3.9) *and let* δ_n *be positive constants satisfying* (3.10). *Starting with consistent estimates* $b_{n,1}, b_{n,2}$ *of* β *satisfying* (3.8) *and based separately on the two subsamples as defined before, define* $\hat{h}_{n,1}, \hat{h}_{n,2}$ *by* (3.11) *and* $\psi_{n,1}, \psi_{n,2}$ *by* (3.13). *Define* $S_{n,1}(b)$, $S_{n,2}(b)$ *and* $S_n^*(b)$ *by* (3.3)–(3.5). *Let* $h = f/(1-F)$. *Then with probability* 1,

$$\begin{aligned} S_n^*(b) - S_n^*(\beta) &= nKA^*(b-\beta) + o(n^{1/2} \vee n|b-\beta|) \\ &\qquad \text{uniformly in } b \text{ with } |b-\beta| \le n^{-\varepsilon} \text{ as } n \to \infty \end{aligned} \tag{3.15}$$

for every $\varepsilon > 0$, *where* $A^* = \int_{\tau_0}^{\tau} (h'/h)^2 (G_2 - G_1^2/G_0)\, dF$ *and* K *is defined in* (2.27). *Moreover, as* $n \to \infty$, $n^{-1/2} S_n^*(\beta)$ *has a limiting normal distribution with mean* 0 *and variance* KA^*.

The proof of Theorem 2 will be given in Section 4 and uses the general tightness lemma for stochastic integrals of empirical-type processes presented there. From the asymptotic linearity property (3.15) of $S_n^*(b)$ in $[\beta - n^{-\varepsilon}, \beta + n^{-\varepsilon}]$ and the asymptotic normality of $n^{-1/2} S_n^*(\beta)$ established in Theorem 2, we obtain as in Corollary 1 the following result on the asymptotic normality of β_n^*.

COROLLARY 2. *With the same notation and assumptions as in Theorem* 2, *suppose that* $A^* > 0$ *and that* b_n *is an* (*auxiliary*) *estimator of* β *such that*

$$b_n - \beta = O_p(n^{-d}) \quad \text{for some } d > 0. \tag{3.16}$$

Then for every $0 < \varepsilon < \frac{1}{2}$, $P\{S_n^*(b)$ *has a zero-crossing in* $[\beta - n^{-\varepsilon}, \beta + n^{-\varepsilon}]$ *for all large* $n\} = 1$. *Consequently, if* β_n^* *is defined as the zero-crossing of* $S_n^*(b)$ *closest to* b_n, *then as* $n \to \infty$, $n^{1/2}(\beta_n^* - \beta)$ *has a limiting normal distribution with mean* 0 *and variance* $(KA^*)^{-1}$.

In view of (3.8), we can choose the b_n of Corollary 2 to be $(b_{n,1} + b_{n,2})/2$. Corollary 2 shows that the adaptive rank estimator β_n^* attains the lower bound (3.1) for the variances of the asymptotic normal distributions of the rank estimators $\hat{\beta}_n$ developed in Section 2. More generally, in the setting of Corollary 2, the variance $(KA^*)^{-1}$ of the limiting normal distribution of $n^{1/2}(\beta_n^* - \beta)$ is in fact an asymptotic lower bound for the variances of the limiting distributions of regular estimators [cf. Begun, Hall, Huang and Wellner (1983)] for the semiparametric problem of estimating β when the common distribution of the ε_i and the bivariate distributions of (t_i, c_i) are unknown. In the case of i.i.d. (t_i, c_i, x_i), the general theory of asymptotic lower bounds in semiparametric estimation developed by Begun, Hall, Huang and Wellner (1983) can be applied to the present problem. Extension of this theory to the setting of Corollary 2, in which the random vectors (t_i, c_i, x_i) need not be identically distributed, shows that for any regular estimator $\tilde{\beta}_n$, the limiting distribution of $n^{1/2}(\tilde{\beta}_n - \beta)$ can be represented as the convolution of two distributions one of which is $N(0, (KA^*)^{-1})$, establishing the asymptotic optimality of β_n^* within the class of regular estimators. The details will be presented elsewhere.

We conclude this section by considering extensions of the rank estimator $\hat{\beta}_n$ [which is a zero-crossing of (1.5)] and the adaptive rank estimator β_n^* to the case of multivariate covariates x_i. Suppose that the β and x_i in (1.1) are replaced by $\nu \times 1$ vectors $\beta = (\beta_1, \dots, \beta_\nu)^T$ and $x_i = (x_{i1}, \dots, x_{i\nu})^T$ and that by βx_i we mean $\beta^T x_i$, where β^T denotes the transpose of β. We also use $|b|$ to denote $(b^T b)^{1/2}$ for the vector b. As before, the y_i in (1.1) are not completely observable due to left-truncation and right-censoring. Defining $S_n(b)$ by (1.5), note that $S_n(b)$ is now a $\nu \times 1$ vector. Assuming that an upper bound $\rho > |\beta|$ is known, we define the multivariate rank estimator $\hat{\beta}_n$ as a minimizer of $|S_n(b)|$ with $|b| \le \rho$. As will be shown in Section 4, Lemma 1 can be easily extended to this multivariate setting. Therefore, the same arguments are those used in the proof of Theorem 1 and Corollary 1 can be used to prove the following.

THEOREM 3. *Suppose that in Theorem* 1, $x_i = (x_{i1}, \dots, x_{i\nu})^T$ *and* $\beta = (\beta_1, \dots, \beta_\nu)^T$ *as* $\nu \times 1$ *vectors and that the assumption* (2.29a) *is replaced by its multivariate version*:

For $j, k \in \{1, \dots, \nu\}$ *and for* $s < F^{-1}(1)$,

$$\lim_{m \to \infty} m^{-1} \sum_{i=1}^{m} P\{t_i - \beta^T x_i \le s \le c_i - \beta^T x_i\} = G_0(s),$$

$$\lim_{m \to \infty} m^{-1} \sum_{i=1}^{m} E\{x_{ik} P[t_i - \beta^T x_i \le s \le c_i - \beta^T x_i | x_i]\} = \Gamma_k(s), \tag{3.17}$$

$$\lim_{m \to \infty} m^{-1} \sum_{i=1}^{m} E\{x_{ij} x_{ik} P[t_i - \beta^T x_{i'} \le s \le c_i - \beta^T x_i | x_i]\} = \Gamma_{jk}(s).$$

Defining $\tilde{F}$ *as in* (2.26), *let* $A = (a_{ij})_{1 \le i, j \le \nu}$ *be defined by*

$$a_{ij} = \int_{\tau_0}^{\tau} \psi(\tilde{F}(s)) \left\{ \frac{f'(s)}{f(s)} + \frac{f(s)}{1 - F(s)} \right\} \left\{ \Gamma_{ij}(s) - \frac{\Gamma_i(s)\Gamma_j(s)}{G_0(s)} \right\} dF(s). \tag{3.18}$$

Then the conclusions of Theorem 1(i)–(ii) *still hold. Moreover, as* $n \to \infty$, $n^{-1/2} S_n(\beta)$ *has a limiting normal distribution with mean vector* 0 *and* *covariance matrix* $V = (v_{ij})_{1 \le i, j \le \nu}$, *where*

$$(3.19)\qquad v_{ij} = K\int_{\tau_0}^{\tau} \psi^2(\tilde{F}(s))\left\{\Gamma_{ij}(s) - \frac{\Gamma_i(s)\Gamma_j(s)}{G_0(s)}\right\} dF(s).$$

If furthermore (2.42) *holds and* A *is nonsingular, then* $\hat{\beta}_n \to \beta$ *a.s. and* $n^{1/2}(\hat{\beta}_n - \beta)$ *has a limiting normal distribution with mean vector* 0 *and covariance matrix* $K^{-2}A^{-1}VA^{-1}$.

Similarly, we can extend Theorem 2 and Corollary 2 to multivariate x_i and β. Letting bx_i denote $b^T x_i$ and defining $S_{n,1}(b)$, $S_{n,2}(b)$ and $S_n^*(b)$ by (3.3)–(3.5), the arguments used to prove Theorem 2 in Section 4 can also be used to prove the following.

THEOREM 4. *Suppose that in Theorem* 2, x_i *and* β *are* $\nu \times 1$ *vectors and that the assumption* (2.29a) *is replaced by its multivariate version* (3.17). *Then* (3.15) *still holds with* $A^* = (a_{ij}^*)_{1 \le i, j \le \nu}$, *where*

$$(3.20)\qquad a_{ij}^* = \int_{\tau_0}^{\tau} \left(\frac{h'(s)}{h(s)}\right)^2 \left\{\Gamma_{ij}(s) - \frac{\Gamma_i(s)\Gamma_j(s)}{G_0(s)}\right\} dF(s).$$

Moreover, as $n \to \infty$, $n^{-1/2}S_n^*(\beta)$ *has a limiting normal distribution with mean vector* 0 *and covariance matrix* KA^*.

As in Corollary 2, starting with an auxiliary consistent estimator b_n of β that satisfies (3.16), we define an adaptive rank estimator β_n^* in the multiple regression setting to be a minimizer of $|S_n^*(b)|$ in the ellipsoid $\{b: |b - b_n| \le n^{-\varepsilon}\}$ with $0 < \varepsilon < d \vee \frac{1}{2}$. Then under the assumptions of Theorem 4, if furthermore A^* is nonsingular, it follows from Theorem 4 that as $n \to \infty$, $n^{1/2}(\beta_n^* - \beta)$ has a limiting normal distribution with mean vector 0 and covariance matrix $K^{-1}(A^*)^{-1}$, which is asymptotically optimal as discussed before for the univariate case $\nu = 1$.

4. A tightness lemma for stochastic integrals of empirical-type processes and the proofs of Lemma 1 and Theorem 2. Defining $L_m(b, s)$ and $Y_m(b, s)$ as in (2.3) and (2.4) but with b and x_i being $\nu \times 1$ vectors and $bx_i = b^T x_i$, consider stochastic integrals of the form

$$(4.1)\qquad \int_{s=-\infty}^{y} U_{m,n}(b, s)\, dL_m(b, s) \quad \text{or} \quad \int_{s=-\infty}^{y} U_{m,n}(b, s)\, dY_m(b, s),$$

where $\int_{s=-\infty}^{y}$ denotes either $\int_{-\infty < s < y}$ or $\int_{-\infty < s \le y}$. The tightness lemma is concerned with approximating $L_m(b, s)$ or $Y_m(b, s)$ in (4.1) by $EL_m(b, s)$ or $EY_m(b, s)$ and the random variables $U_{m,n}(b, s)$ by $u_{m,n}(b, s)$ which are independent of $\{(\varepsilon_i, t_i, c_i, x_i)\}$ and which satisfy the following assumptions for some

$D \geq \delta > 0$, $\xi \geq 0$ and $C_1, C_2, n_0, d_0 > 0$: For every $0 \leq \gamma < 1$ and $\varepsilon > 0$,

$$\max_{\delta n \leq m \leq Dn} \sup_{\substack{|b-a| \leq n^{-\gamma} \\ -\infty < s < \infty}} |U_{m,n}(b,s) - u_{m,n}(b,s) - U_{m,n}(a,s) + u_{m,n}(a,s)| = O(n^{-1/2-\gamma/2+\xi+\varepsilon}) \quad \text{a.s.,} \tag{4.2a}$$

$$\max_{\delta n \leq m \leq Dn} \sup_{\substack{|b| \leq \rho \\ -\infty < s < \infty}} |U_{m,n}(b,s) - u_{m,n}(b,s)| = O(n^{-1/2+\xi+\varepsilon}) \quad \text{a.s.,} \tag{4.2b}$$

$U_{m,n}(b,s)$ has bounded variation in s for every $b \in [-\rho, \rho]$ and

$$\max_{\delta n \leq m \leq Dn} \sup_{|b| \leq \rho} \int_{s=-\infty}^{\infty} |dU_{m,n}(b,s)| = O(n^{\xi}) \quad \text{a.s.,} \tag{4.2c}$$

$$\max_{\delta n \leq m \leq Dn} \sup_{\substack{|b| \leq \rho \\ -\infty < s < \infty}} |u_{m,n}(b,s)| \leq C_1 n^{\xi} \quad \text{a.s. for all } n \geq n_0, \tag{4.2d}$$

$$\max_{\delta n \leq m \leq Dn} \sup_{|b-b'|+|s-s'| \leq d} n^{-\xi} |u_{m,n}(b,s) - u_{m,n}(b',s')| \leq C_2 d \quad \text{a.s. for } 0 < d \leq d_0 \text{ and } nd \geq n_0. \tag{4.2e}$$

LEMMA 3. *Let $\varepsilon_1, \varepsilon_2, \ldots$ be i.i.d. random variables with a common distribution function F and let (t_i, c_i, x_i^T) be independent random vectors that are independent of $\{\varepsilon_i\}$. Suppose that* (2.11)–(2.14) *hold. Let $y_i = \beta^T x_i + \varepsilon_i$, with $|\beta| \leq \rho$, and define $L_m(b,s)$, $Y_m(b,s)$ by* (2.3) *and* (2.4). *Let $D \geq \delta > 0$ and $\xi \geq 0$ and let $U_{m,n}(b,s)$, $u_{m,n}(b,s)$ be random variables satisfying* (4.2a)–(4.2e) *for every $0 \leq \gamma < 1$ and $\varepsilon > 0$ and such that the family $\{u_{m,n}(b,s)\colon |b| \leq \rho, -\infty < s < \infty\}$ is independent of $\{(\varepsilon_i, t_i, c_i, x_i^T)\colon i = 1, \ldots, m\}$ for every n and $\delta n \leq m \leq Dn$. Then for every $0 \leq \gamma < 1$ and $\varepsilon > 0$,*

$$\max_{\delta n \leq m \leq Dn} \sup_{\substack{|b-a| \leq n^{-\gamma} \\ -\infty < y < \infty}} \left| \int_{-\infty}^{y} U_{m,n}(b,s) L_m(b,ds) - \int_{-\infty}^{y} u_{m,n}(b,s) EL_m(b,ds) - \int_{-\infty}^{y} U_{m,n}(a,s) L_m(a,ds) + \int_{-\infty}^{y} u_{m,n}(a,s) EL_m(a,ds) \right| = o(n^{(1-\gamma)/2+\xi+\varepsilon}) \quad a.s., \tag{4.3}$$

$$\max_{\delta n \leq m \leq Dn} \sup_{\substack{|b| \leq \rho \\ -\infty < y < \infty}} \left| \int_{-\infty}^{y} U_{m,n}(b,s) L_m(b,ds) - \int_{-\infty}^{y} u_{m,n}(b,s) EL_m(b,ds) \right| = o(n^{1/2+\xi+\varepsilon}) \quad a.s., \tag{4.4}$$

where $EL_m(b, ds)$ *is given by* (2.15). *Moreover*, (4.3) *and* (4.4) *still hold if* L_m *is replaced by* Y_m.

Lemma 3 is an extension of Theorem 2 of Lai and Ying (1988), which treats only the censored case (with $t_i \equiv -\infty$) and which only considers the case $m = n$ and nonrandom $u_{m,n}$. It can be proved by the same arguments as those used in the proof of Theorems 1 and 2 of Lai and Ying (1988). It says that under certain conditions we can approximate the stochastic integral $\int_{-\infty}^{y} U_{m,n}(b, s)L_m(b, ds)$ by $\int_{-\infty}^{y} u_{m,n}(b, s)EL_m(b, ds)$ and provides two kinds of error bounds for the approximation. We now apply Lemma 3 to prove Lemma 1.

PROOF OF LEMMA 1. Making use of Lemma 3 and arguments similar to those in the proof of Theorem 3 of Lai and Ying (1988), it can be shown that for every $D \geq 1$, $0 < \gamma < 1$ and $\varepsilon > 0$,

$$\max_{n \le m \le Dn} \sup_{|b-a| \le n^{-\gamma}} |T_{m,n}(b) - \phi_{m,n}(b) - T_{m,n}(a) + \phi_{m,n}(a)| \tag{4.5}$$
$$= O(n^{(1-\gamma)/2+3\lambda+\varepsilon}) \quad \text{a.s.},$$

$$\max_{n \le m \le Dn} \sup_{|b| \le \rho} |T_{m,n}(b) - \phi_{m,n}(b)| = O(n^{1/2+3\lambda+\varepsilon}) \quad \text{a.s.}, \tag{4.6}$$

where λ is given in (2.18). Since $\lambda < \frac{1}{18}$, (2.21) follows from (4.6). Moreover, by (4.5), with probability 1,

$$\max_{n \le m \le Dn} |T_{m,n}(b) - \phi_{m,n}(b) - T_{m,n}(\beta) + \phi_{m,n}(\beta)|$$
$$= o(n^{1/2}) \text{ uniformly in } b \text{ with } |b - \beta| \le n^{-1/3},$$
$$= o(n^{2/3}) \text{ uniformly in } |b| \le \rho \text{ with } |b - \beta| \ge n^{-1/3},$$

since $\lambda < \frac{1}{18}$. Noting that $n^{2/3} \le n|b - \beta|$ for $|b - \beta| \ge n^{-1/3}$, (2.22) follows. □

REMARK. Since Lemma 3 is stated for multivariate b and x_i, the preceding proof also extends Lemma 1 to the multiple regression setting, which is used in the proof of Theorem 3.

We now proceed to prove Theorem 2, noting that the same arguments can also be used to prove Theorem 4 on multivariate x_i and b. The first step is to develop a stochastic integral representation of $S_n^*(b)$. As in Section 2, we shall regard the first subsample of $[n/2]$ observations as being generated by a larger sample of independent random vectors (y_i, t_i, c_i, x_i), $i = 1, \dots, m_1(n)$, where

$$m_1(n) = \inf\left\{ m: \sum_{i=1}^{m} I_{\{t_i \le y_i \wedge c_i\}} = [n/2] \right\}. \tag{4.7}$$

Likewise, defining $m(n)$ $(> m_1(n))$ as in (2.1), we shall regard the second subsample as being generated by (y_i, t_i, c_i, x_i), $m_1(n) + 1 \le i \le m(n)$. Define

L_m, Y_m, N_m, X_m by (2.3)–(2.6). Moreover, for $M > m$, define also

(4.8a) $$L_{M,m}(b,s) = \sum_{i=m+1}^{M} I_{\{t_i(b) \le y_i(b) \wedge c_i(b) \le s,\, y_i(b) \le c_i(b)\}},$$

(4.8b) $$M_{M,m}(b,s) = \sum_{i=m+1}^{M} x_i I_{\{t_i(b) \le y_i(b) \wedge c_i(b) \le s,\, y_i(b) \le c_i(b)\}},$$

(4.8c) $$N_{M,m}(b,s) = \sum_{i=m+1}^{M} I_{\{t_i(b) \le s \le y_i(b) \wedge c_i(b)\}},$$

(4.8d) $$X_{M,m}(b,s) = \sum_{i=m+1}^{M} x_i I_{\{t_i(b) \le s \le y_i(b) \wedge c_i(b)\}}.$$

The estimate $\psi_{n,1}$ of h'/h defined from the first subsample by (3.13) can be written as $\psi_{n,1}(s) = u_{m_1(n),n}(s)$, where

(4.9) $$u_{m,n}(s) = \Psi_1(s, n; (y_i, t_i, c_i, x_i)_{1 \le i \le m})$$

for some Borel function Ψ_1. Likewise $\psi_{n,2}(s) = v_{m(n),m_1(n),n}(s)$, where

(4.10) $$v_{M,m,n}(s) = \Psi_2(s, n; (y_i, t_i, c_i, x_i)_{m+1 \le i \le M})$$

for some Borel function Ψ_2. Define for $M > m$,

(4.11) $$\begin{aligned} T^*_{M,m,n}(b) = &\int_{-\infty}^{\infty} p_n(n^{-1}N_m(b,s)) v_{M,m,n}(s) \\ &\quad\times \left[Y_m(b,ds) - \frac{X_m(b,s)}{N_m(b,s)} L_m(b,ds) \right] \\ &+ \int_{-\infty}^{\infty} p_n(n^{-1}N_{M,m}(b,s)) u_{m,n}(s) \\ &\quad\times \left[Y_{M,m}(b,ds) - \frac{X_{M,m}(b,s)}{N_{M,m}(b,s)} L_{M,m}(b,ds) \right]. \end{aligned}$$

Then analogous to (2.10), it follows from (3.3)–(3.5) that

(4.12) $$S_n^*(b) = T^*_{m(n),m_1(n),n}(b).$$

Note that $\{u_{m,n}(s), -\infty < s < \infty\}$ is independent of $\{(y_i, t_i, c_i, x_i): i \ge m+1\}$ and that $\{v_{M,m,n}(s), -\infty < s < \infty\}$ is independent of $\{(y_i, t_i, c_i, x_i): i \le m\}$ for every $M > m$ and n. Note also that the assumptions (4.2a) and (4.2b) are trivially satisfied by $U_{m,n}(b,s) = u_{m,n}(b,s) = u_{m,n}(s)$. Hence, applying Lemma 3 and a slight modification thereof (with an additional index M) to $U_{m,n}(b,s) = u_{m,n}(s)$ and to $U_{M,m,n}(b,s) = v_{M,m,n}(s)$, we have the following analog of Lemma 1.

LEMMA 4. *Let ε_i, t_i, c_i, x_i and p_n be the same as in Lemma 1. Let $y_i = \beta x_i + \varepsilon_i$ with $|\beta| \le \rho$. For $M > m$, define L_m, Y_m, N_m, X_m, $L_{M,m}$, $Y_{M,m}$, $X_{M,m}$ by (2.3)–(2.6) and (4.8). Let $D \ge \delta > 0$. Defining $u_{m,n}(s)$ and $v_{M,m,n}(s)$*

as in (4.9) *and* (4.10) *for some Borel functions* Ψ_1 *and* Ψ_2, *assume that for every* $\xi > 0$, *there exist* n_0 *and* d_0 (*depending on* ξ) *such that*

$u_{m,n}(s)$ *and* $v_{M,m,n}(s)$ *have bounded variation in s and for all* $n \ge n_0$,

$$\max_{\delta n \le m < M \le Dn} \left\{ \int_{-\infty}^{\infty} |du_{m,n}(s)| + |u_{m,n}(0)| + \int_{-\infty}^{\infty} |dv_{M,m,n}(s)| + |v_{M,m,n}(0)| \right\} \le n^{\xi} \quad a.s., \tag{4.13}$$

$$\max_{\delta n \le m \le Dn} \sup_{|s-s'| \le d} \{ |u_{m,n}(s) - u_{m,n}(s')| + |v_{M,m,n}(s) - v_{M,m,n}(s')| \} \le n^{\xi} d \quad a.s. \text{ for all } 0 < d \le d_0 \text{ and } nd \ge n_0. \tag{4.14}$$

Letting $n_1 = [n/2]$ *and* $n_2 = n - n_1$, *define* $T^*_{M,m,n}(b)$ *by* (4.11) *and let*

$$\begin{aligned} \phi^*_{M,m,n}(b) = &\int_{-\infty}^{\infty} p_n(n^{-1}EN_m(b,s)) v_{M,m,n}(s) \\ &\times \left[EY_m(b,ds) - \frac{EX_m(b,s)}{EN_m(b,s)} EL_m(b,ds) \right] \\ &+ \int_{-\infty}^{\infty} p_n(n^{-1}EN_{M,m}(b,s)) u_{m,n}(s) \\ &\times \left[EY_{M,m}(b,ds) - \frac{EX_{M,m}(b,s)}{EN_{M,m}(b,s)} EL_{M,m}(b,ds) \right]. \end{aligned} \tag{4.15}$$

Then

$$\max_{\delta n \le m < M \le Dn} \sup_{|b| \le \rho} n^{-1} |T^*_{M,m,n}(b) - \phi^*_{M,m,n}(b)| \to 0 \quad a.s., \tag{4.16}$$

$$\max_{\delta n \le m < M \le Dn} \sup_{|b| \le \rho} \frac{|T^*_{M,m,n}(b) - \phi^*_{M,m,n}(b) - T^*_{M,m,n}(\beta) + \phi^*_{M,m,n}(\beta)|}{n^{1/2} \vee n|b - \beta|} \to 0 \quad a.s. \tag{4.17}$$

The following lemma shows that the $\psi_{n,j}$ defined in Theorem 2 by (3.13) satisfy the assumptions (3.6), (3.7), (4.13) and (4.14) for any $\xi > 0$. For notational simplicity we shall only consider the first subsample $j = 1$, as the case $j = 2$ can be treated by the same argument. In particular, part (ii) of the lemma shows that conditions (4.13) and (4.14) of Lemma 4 are satisfied by $u_{m,n}(s) = \Psi_{m,n}(b(y_1, t_1, c_1, x_1, \dots, y_m, t_m, c_m, x_m), s)$, in which $\Psi_{m,n}$ is defined by (4.21) and the function b represents the preliminary estimator of β based on the observable components of $\{(y_i, t_i, c_i, x_i): i \le m\}$, so that the case $m = m_1(n)$ gives the $b_{n,1}$ in (3.8) and (3.11).

LEMMA 5. *With the same notation and assumptions as in Theorem 2, define* L_m, Y_m, N_m, X_m *by* (2.3)–(2.6) *and* $m_1(n)$ *by* (4.7). *Let*

$$(4.18)\quad W_{m,n}(b,y) = \frac{1}{n\delta_n}\int_{-\infty}^{\infty} w\left(\frac{y-s}{\delta_n}\right) p\left(\frac{n^{-1}N_m(b,s)-\delta_n}{\delta_n}\right)\frac{L_m(b,ds)}{n^{-1}N_m(b,s)},$$

$$(4.19)\quad h_{m,n}(b,y) = \frac{1}{n\delta_n}\int_{-\infty}^{\infty} w\left(\frac{y-s}{\delta_n}\right) p\left(\frac{n^{-1}EN_m(b,s)-\delta_n}{\delta_n}\right)\frac{EL_m(b,ds)}{n^{-1}EN_m(b,s)},$$

$$(4.20)\quad \Psi^*_{m,n}(b,y) = \frac{[(\partial/\partial y)W_{m,n}(b,y)]\, I_{\{W_{m,n}(b,y)\ge \delta_n^{1/4},\, n^{-1}N_m(b,y)\ge \delta_n^{1/5}\}}}{W_{m,n}(b,y)},$$

$$(4.21)\quad \Psi_{m,n}(b,y) = \left\{\frac{1}{\delta_n}\int_{-\infty}^{\infty} w\left(\frac{y-u}{\delta_n}\right)\Psi^*_{m,n}(b,u)\,du\right\} p(2-\delta_n y^2).$$

Then

$$(4.22)\quad \hat{h}_{n,1}(y) = W_{m_1(n),n}(b_{n,1},y), \qquad \psi_{n,1}(y) = \Psi_{m_1(n),n}(b_{n_1},y).$$

(i) *For every* $D > \delta > 0$ *and* $\varepsilon > 0$,

$$\max_{\delta n\le m\le Dn}\ \sup_{\substack{|b|\le\rho\\ -\infty<y<\infty}} |W_{m,n}(b,y) - h_{m,n}(b,y)| = O\big((n\delta_n)^{-1}\delta_n^{-3}n^{(1/2)+\varepsilon}\big) \quad a.s.,$$

$$\max_{\delta n\le m\le Dn}\ \sup_{\substack{|b|\le\rho\\ -\infty<y<\infty}} \left|\frac{\partial}{\partial y}W_{m,n}(b,y) - \frac{\partial}{\partial y}h_{m,n}(b,y)\right| = O\big((n\delta_n)^{-1}\delta_n^{-4}n^{(1/2)+\varepsilon}\big) \quad a.s.$$

(ii) *For every* $b\in[-\rho,\rho]$, $\Psi_{m,n}(b,y)$ *is a twice continuously differentiable function of* y *and*

$$\sup_{\substack{|b|\le\rho\\ -\infty<y<\infty}} |\Psi_{m,n}(b,y)| \le \left(\frac{m}{n}\right)\delta_n^{-13/4}\|w'\|_\infty,$$

$$\sup_{\substack{|b|\le\rho\\ -\infty<y<\infty}} \left|\frac{\partial}{\partial y}\Psi_{m,n}(b,y)\right| \le \left(\frac{m}{n}\right)\delta_n^{-17/4}\|w'\|_\infty\|w'\|_1 + 4\delta_n^{1/2}\|p'\|_\infty \sup_{b,y}|\Psi_{m,n}(b,y)|,$$

$$\sup_{|b|\le\rho}\int_{-\infty}^{\infty}\left|\frac{\partial}{\partial y}\Psi_{m,n}(b,y)\right| dy \le 8\left(\frac{m}{n}\right)\delta_n^{-13/4}\|w'\|_\infty\{\|p'\|_\infty + \delta_n^{-3/2}\|w'\|_1\},$$

where $\|g\|_1 = \int_{-\infty}^{\infty}|g(t)|\,dt$ *and* $\|g\|_\infty = \sup_t|g(t)|$.

PROOF. First note that by (4.18),

$$\frac{\partial}{\partial y} W_{m,n}(b, y) = \frac{1}{n\delta_n^2} \int_{-\infty}^{\infty} w'\left(\frac{y-s}{\delta_n}\right) p\left(\frac{n^{-1}N_m(b,s) - \delta_n}{\delta_n}\right) \frac{L_m(b, ds)}{n^{-1}N_m(b,s)}. \tag{4.23}$$

Moreover, $0 \le p(y) \le 1$ and by (2.17),

$$\begin{aligned} &p((x-\delta_n)/\delta_n) = 0 \quad \text{if } x \le \delta_n \quad \text{and} \\ &p((x-\delta_n)/\delta_n) = 1 \quad \text{if } x \ge 2\delta_n. \end{aligned} \tag{4.24}$$

The proof of (i) is similar to that of Theorems 2 and 3 of Lai and Ying (1988). Since $\int_{-\infty}^{\infty} L_m(b, ds) \le m$, it follows from (4.23) and (4.24) that

$$\sup_{\substack{|b| \le \rho \\ -\infty < y < \infty}} \left| \frac{\partial}{\partial y} W_{m,n}(b, y) \right| \le \left(\frac{m}{n}\right) \delta_n^{-3} \|w'\|_\infty.$$

Therefore by (4.20), $\sup_{\|b\| \le \rho,\, -\infty < y < \infty} |\Psi^*_{m,n}(b,y)| \le (m/n)\delta_n^{-13/4}\|w'\|_\infty$. Noting that $p(2 - \delta_n y^2) = 0$ if $|y| \ge (2/\delta_n)^{1/2}$ by (2.17) and making use of (3.9), we obtain (ii). □

PROOF OF THEOREM 2. We shall use the stochastic integral representation (4.11) of $S_n^*(b) = T^*_{m(n), m_1(n), n}(b)$. As shown in the proof of Theorem 1, $m(n)/n \to K$ a.s. and the same argument can be used to show that $m_1(n)/n \to K/2$ a.s. It will be shown in Lemma 6 that (3.7) and (3.6) are satisfied with $d_n = 3\delta_n^{1/4} (\to 0)$. Defining $\phi^*_{M,m,n}(b)$ as in (4.15) and making use of (3.6), (3.7), (3.14) and (2.24), it can be shown by arguments similar to those in the Appendix of Lai and Ying (1989) that as $n \to \infty$,

$$\begin{aligned} &\phi^*_{m(n), m_1(n), n}(b) - \phi^*_{m(n), m_1(n), n}(\beta) \\ &\quad = m(n)(b-\beta) \int_{\tau_0}^{\tau} \frac{h'(s)}{h(s)} \left\{ \frac{f'(s)}{f(s)} + \frac{f(s)}{1-F(s)} \right\} \\ &\qquad \times \left\{ G_2(s) - \frac{G_1^2(s)}{G_0(s)} \right\} f(s)\, ds + o(n^{1/2} \vee n|b-\beta|), \\ &\qquad + o\left(n^{1/2} \vee n|b-\beta|\right), \end{aligned} \tag{4.25}$$

uniformly in $b \in [\beta - n^{-\varepsilon}, \beta + n^{-\varepsilon}]$ for every $\varepsilon > 0$. Since $h'/h = f'/f + f/(1-F)$, (3.15) follows from (4.25) and Lemmas 4 and 5. Note in this connection that (3.14) implies that

$$\begin{aligned} &\int_{\tau_0 - \eta}^{\tau + \eta} \left(\frac{h'}{h}\right)^2 dF < \infty \quad \text{for some } \eta > 0 \text{ and} \\ &f(s) = O\left((1 - F(s))^{1/2}\right) \text{ as } s \uparrow \tau, \end{aligned} \tag{4.26}$$

as shown in Lemma 2 of Lai and Ying (1989).

To prove the asymptotic normality of $S_n^*(\beta)$, let $T_i = t_i - \beta x_i$, $C_i = c_i - \beta x_i$, and define $M_i(s)$ by (2.38) and $\mathscr{F}(s)$ as in Lemma 2. Then by Lemma 2 $\{M_i(s), \mathscr{F}(s), -\infty < s < \infty\}$ is a martingale. Analogous to (2.40), we now have $S_n^*(\beta) = S_{n,1}^*(\beta;\infty) + S_{n,2}^*(\beta;\infty)$, where

$$S_{n,1}^*(\beta;t) = \sum_{i=1}^{m_1(n)} \int_{-\infty}^{t} \psi_{n,2}(s) p_n\big(n^{-1}N_{m_1(n)}(\beta,s)\big) \times \left\{x_i - \frac{X_{m_1(n)}(\beta,s)}{N_{m_1(n)}(\beta,s)}\right\} dM_i(s),$$

$$S_{n,2}^*(\beta;t) = \sum_{i=m_1(n)+1}^{m(n)} \int_{-\infty}^{t} \psi_{n,1}(s) p_n\big(n^{-1}N_{m(n),m_1(n)}(\beta,s)\big) \times \left\{x_i - \frac{X_{m(n),m_1(n)}(\beta,s)}{N_{m(n),m_1(n)}(\beta,s)}\right\} dM_i(s).$$

As noted in the proof of Theorem 1(iii), the random variables $m(n)$ and $m_1(n)$ are $\bigcap_{s=-\infty}^{\infty} \mathscr{F}(s)$-measurable. Define

$$S_{n,1}(\beta;t) = \sum_{i=1}^{m_1(n)} \int_{-\infty}^{t} \frac{h'(s)}{h(s)} p_n\big(n^{-1}N_{m_1(n)}(\beta,s)\big) \times \left\{x_i - \frac{X_{m_1(n)}(\beta,s)}{N_{m_1(n)}(\beta,s)}\right\} dM_i(s),$$

$$S_{n,2}(\beta;t) = \sum_{i=m_1(n)+1}^{m(n)} \int_{-\infty}^{t} \frac{h'(s)}{h(s)} p_n\big(n^{-1}N_{m(n),m_1(n)}(\beta,s)\big) \times \left\{x_i - \frac{X_{m(n),m_1(n)}(\beta,s)}{N_{m(n),m_1(n)}(\beta,s)}\right\} dM_i(s).$$

Since h'/h is nonrandom, the same argument as that in the proof of Theorem 1(iii) shows that $\{S_{n,j}(\beta;t), \mathscr{F}(t), -\infty < t < \infty\}$ is a martingale for $j = 1,2$, and that $n^{-1/2}\{S_{n,1}(\beta;\infty) + S_{n,2}(\beta;\infty)\}$ has a limiting normal distribution with mean 0 and variance K^*A.

Let $\mathscr{B}^*$ denote the σ-field generated by $\{(\varepsilon_i, t_i, c_i, x_i): i \geq m_1(n)+1\}$ and let $\mathscr{F}^*(s)$ be the σ-field generated by $\mathscr{F}(s) \cup \mathscr{B}^*$. Then $\psi_{n,2}(u)$ is measurable with respect to $\mathscr{B}^*$ for every u and $\{n^{-1/2}(S_{n,1}^*(\beta;t) - S_{n,1}(\beta;t)), \mathscr{F}^*(t), -\infty < t < \infty\}$ is a martingale with predictable variation process

$$\begin{aligned}\big\langle n^{-1/2}\{S_{n,1}^*(\beta;\cdot) - S_{n,1}(\beta;\cdot)\}\big\rangle(t) &\leq (2B)^2 n^{-1} \sum_{i=1}^{m_1(n)} \int_{-\infty}^{t} \left\{\psi_{n,2}(s) - \frac{h'(s)}{h(s)}\right\}^2 \\ &\qquad \times I_{\{c_i - \beta x_i \geq s \geq t_i - \beta x_i\}} I_{\{\varepsilon_i \geq s\}}\, d\Lambda(s),\end{aligned}$$

by Lemma 2, recalling that $|x_i| \le B$ and that $0 \le p_n \le 1$. From (3.6), (3.7) and (4.26), it then follows that

$$\sup_t \langle n^{-1/2}\{S^*_{n,1}(\beta;\cdot) - S_{n,1}(\beta;\cdot)\}\rangle(t) \to_P 0.$$

Hence by Lenglart's inequality [cf. Gill (1980), pages 18–19],

$$n^{-1/2}\{S^*_{n,1}(\beta;\infty) - S_{n,1}(\beta;\infty)\} \to_P 0.$$

By a similar argument, $n^{-1/2}\{S^*_{n,2}(\beta;\infty) - S_{n,2}(\beta;\infty)\} \to_P 0$. □

LEMMA 6. *With the same notation and assumptions as in Theorem 2, $\hat{h}_{n,j}$ satisfies (3.12) and $\psi_{n,j}$ satisfies (3.7) and (3.6) with $d_n = 3\delta_n^{1/4}$ for $j = 1, 2$.*

PROOF. We only consider the first subsample $j = 1$, as the case $j = 2$ can be proved similarly. We shall make use of Lemma 5(i), which approximates $W_{m,n}(b, y)$ by the nonrandom function $h_{m,n}(b, y)$ defined in (4.19). We first show that for every $D > \delta > 0$ and $\varepsilon > 0$,

$$\max_{\delta n \le m \le Dn} \sup_{|b-\beta| \le n^{-\varepsilon}} \sup_{n^{-1}EN_m(b,y) \ge (1/2)\delta_n^{1/5}} \{|h_{m,n}(b,y) - h(y)|\} = O(\delta_n^{4/5}), \tag{4.27}$$

$$\max_{\delta n \le m \le Dn} \sup_{|b-\beta| \le n^{-\varepsilon}} \sup_{n^{-1}EN_m(b,y) \ge (1/2)\delta_n^{1/5}} \{|(\partial/\partial y)h_{m,n}(b,y) - h'(y)|\} = O(\delta_n^{7/10}). \tag{4.28}$$

Let $\bar{F} = 1 - F$. By the assumption on f'' and (2.12), $\sup(f(s) + |f'(s)| + |f''(s)|) < \infty$. Hence by (2.11), (2.13), (2.34) and (3.10), we have for large n,

$$\begin{aligned}
&|b - \beta| \le n^{-\varepsilon},\ \delta n \le m \le Dn \text{ and } n^{-1}EN_m(b,y) \ge \delta_n^{1/5}/2 \\
&\Rightarrow n^{-1}EN_m(b, y + \delta_n t) \wedge n^{-1}EN_m(\beta, y + \delta_n t) \ge \delta_n^{1/5}/3 \\
&\qquad\qquad \text{and } \bar{F}(y + \delta_n t) \ge \delta_n^{1/5}/(3D) \quad \text{for all } |t| \le \alpha \\
&\Rightarrow (\partial/\partial y)h_{m,n}(\beta,y) = \delta_n^{-1}\int_{-\alpha}^{\alpha} w'(t)h(y - \delta_n t)\,dt \\
&\qquad = \int_{-\alpha}^{\alpha} w(t)h'(y - \delta_n t)\,dt \\
&\qquad = h'(y) + O\left(\sup_{|x|\le\alpha}|h''(y - \delta_n x)|\int_{-\alpha}^{\alpha}\delta_n|t|w(t)\,dt\right) \\
&\qquad = h'(y) + O\left((\delta_n^{-1/5})^{3/2}\delta_n\right),
\end{aligned} \tag{4.29}$$

in view of (3.9), since $h' = f'/\bar{F} + h^2$ and $h'' = f''/\bar{F} + hf'/\bar{F} + 2hh'$, and since $f(s)/\bar{F}^{1/2}(s) = O(1)$ as $s \uparrow \tau$ in the case $\bar{F}(\tau) = 0$ by (4.26). Moreover, using (2.13), (2.15), (2.34), (3.9), (4.24) and (4.29), it can be shown that for every $\varepsilon > 0$,

$$\begin{aligned}
\max_{\delta n \le m \le Dn} \sup_{|b-\beta| \le n^{-\varepsilon}} \sup_{n^{-1}EN_m(b,y) \ge (1/2)\delta_n^{1/5}} &\{|h_{m,n}(b,y) - h_{m,n}(\beta,y)| \\
&+ |(\partial/\partial y)h_{m,n}(b,y) - (\partial/\partial y)h_{m,n}(\beta,y)|\} = O(n^{-\varepsilon}\delta_n^{-2/5}).
\end{aligned} \tag{4.30}$$

Combining (4.29) and (4.30) gives (4.28), and a similar argument can be used to prove (4.27).

By a straightforward modification of Theorem 1 of Lai and Ying (1988), it can be shown that for every $\varepsilon > 0$,

$$\sup_{|b| \le \rho,\, -\infty < y < \infty} |N_m(b,y) - EN_m(b,y)| = O(m^{1/2+\varepsilon}) \quad \text{a.s.} \tag{4.31}$$

From (3.10) and (4.31) together with Lemma 5(i) and (4.27), it follows that with probability 1, for all large n,

$$\begin{aligned} &\delta n \le m \le Dn,\ |b-\beta| \le n^{-\varepsilon},\ n^{-1}N_m(b,y) \ge \delta_n^{1/5} \text{ and } W_{m,n}(b,y) \ge \delta_n^{1/4} \\ &\quad \Rightarrow n^{-1}EN_m(b,y) \ge \delta_n^{1/5}/2,\ h(y) \ge \delta_n^{1/4}/2 \text{ and } \bar{F}(y) \ge \delta_n^{1/5}/(3D). \end{aligned} \tag{4.32}$$

Since $h' = (f' + f^2/\bar{F})/\bar{F} = O(1/\bar{F})$ by (4.26) and since $b_{n,1} - \beta = O_p(n^{-d})$ by (3.8), (3.12) follows from Lemma 5(i), (4.22), (4.27), (4.28) and (4.32), noting that $\hat{h}'_{n,1}/\hat{h}_{n,1} - h'/h = (\hat{h}'_{n,1} - h')/\hat{h}_{n,1} + h'(h - \hat{h}_{n,1})/(h\hat{h}_{n,1})$ and that $m_1(n)/n \to K/2$ a.s. Since $\lim_{n\to\infty} n^{-1}N_{m_1(n)}(b_{n,1}, s) = \frac{1}{2}K\bar{F}(s)G_0(s)(>0)$ a.s. for $s \in [s_1, s_2]$ with $\tau_0 < s_1 < s_2 < \tau$, it follows from (3.12) and (3.13) that $\sup_{s\in[s_1,s_2]:\, h(s)\ge 2\delta_n^{1/4}} |\tilde{\psi}_{n,1}(s) - h'(s)/h(s)| \to_P 0$. Using this and (3.9) and noting that $\psi_{n,1}(s) = \int_{-\alpha}^{\alpha} w(s)\tilde{\psi}_{n,1}(s - \delta_n)\,dt$ for $|s| \le \delta_n^{-1/2}$, we obtain (3.6) with $d_n = 3\delta_n^{1/4}$. Moreover, by (3.12) and (3.13),

$$P\left\{ |\tilde{\psi}_{n,1}(s)| \le 2\frac{|h'(s)|}{h(s)} I_{\{h(s)\ge \delta_n^{1/4}/2,\, \bar{F}(s)\ge \delta_n^{1/5}/(2K)\}} \text{ for all } s \right\} \to 1. \tag{4.33}$$

Note that $|\psi_{n,1}(s)| \le \int_{-\alpha}^{\alpha} w(t)|\tilde{\psi}_{n,1}(s - \delta_n t)|\,dt \le \sup_{|t|\le\alpha} |\tilde{\psi}_{n,1}(s - \delta_n t)|$ and that for all large n,

$$\begin{aligned} &h(s - \delta_n u) \ge \delta_n^{1/4}/2 \text{ and } \bar{F}(s - \delta_n u) \ge \delta_n^{1/5}/(2K) \text{ for some } u \in [-\alpha, \alpha] \\ &\quad \Rightarrow \inf_{|t|\le\alpha} 2h(s - \delta_n t) \ge h(s) \ge \delta_n^{1/4}/3 \\ &\qquad\qquad \text{and } \sup_{|t|\le\alpha} |h''(s - \delta_n t)| = O(\delta_n^{-3/10}) \qquad [\text{cf. } (4.29)] \\ &\quad \Rightarrow \sup_{|t|\le\alpha} \frac{|h'(s - \delta_n t)|}{h(s - \delta_n t)} \le \frac{2\{|h'(s)| + O(\delta_n^{7/10})\}}{h(s)} = \frac{2|h'(s)|}{h(s)} + O(\delta_n^{9/20}). \end{aligned}$$

Hence it follows from (4.33) that

$$P\left\{ |\psi_{n,1}(s)| \le 4\frac{h'(s)}{h(s)} I_{\{f(s)>0\}} + \delta_n^{1/3} \text{ for all } s \right\} \to 1. \tag{4.34}$$

Since $\sup_i |x_i| \le B$ and $1 - \lambda > \frac{1}{2}$, we obtain from (4.31) that with probability 1,

$$\sup_{\substack{|b-\beta| \le m^{-d/2} \\ s \notin [\tau_0 - \varepsilon, \tau + \varepsilon]}} N_m(b, s) \le \sum_{i=1}^{m} \Big[P\{t_i - \beta x_i < \tau_0 - \varepsilon + Bm^{-d/2}\} + P\{c_i - \beta x_i > \tau + \varepsilon - Bm^{-d/2}\} \Big] + o(m^{1-\lambda}) = o(m^{1-\lambda}),$$

by (2.30). Since $\#_{n,1}(s) = N_{m_1(n)}(b_{n,1}, s)$, it then follows from (3.13) that for all $\varepsilon > 0$.

$$P\Big\{ \sup_{s \notin [\tau_0 - \varepsilon, \tau + \varepsilon]} |\tilde{\psi}_{n,1}(s)| \ne 0 \Big\} \le P\Big\{ \sup_{s \notin [\tau_0 - \varepsilon, \tau + \varepsilon]} n^{-1} \#_{n,1}(s) \ge \delta_n^{1/5} \Big\} \to 0.$$

Since $|\psi_{n,1}(s)| \le \sup_{|t| \le \alpha} |\tilde{\psi}_{n,1}(s - \delta_n t)|$, we then obtain that

$$\sup_{s \notin [\tau_0 - \varepsilon, \tau + \varepsilon]} |\psi_{n,1}(s)| \to_P 0$$

for every $\varepsilon > 0$. Combining this with (4.26) and (4.34), (3.7) follows. □

REFERENCES

AMEMIYA, T. (1985). *Advanced Econometrics*. Harvard Univ. Press.

BEGUN, J., HALL, W. J., HUANG, W. and WELLNER, J. A. (1983). Information and asymptotic efficiency in parametric-nonparametric models. *Ann. Statist.* **11** 432–452.

CUZICK, J. (1985). Asymptotic properties of censored linear rank tests. *Ann. Statist.* **13** 133–141.

GILL, R. (1980). *Censoring and Stochastic Integrals*. Math. Centre Tract 124. Mathematisch Centrum, Amsterdam.

GOLDBERGER, A. S. (1981). Linear regression after selection. *J. Econometrics* **15** 357–366.

HÁJEK, J. and ŠIDÁK, Z. (1967). *Theory of Rank Tests*. Academic, New York.

JUREČKOVÁ, J. (1969). Asymptotic linearity of a rank statistic in regression parameter. *Ann. Math. Statist.* **40** 1889–1900.

KALBFLEISCH, J. G. and PRENTICE, R. L. (1980). *The Statistical Analysis of Failure Time Data*. Wiley, New York.

KEIDING, N. and GILL, R. D. (1990). Random truncation models and Markov processes. *Ann. Statist.* **18** 582–602.

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

LAI, T. L. and YING, Z. (1989). Linear rank statistics in regression analysis with censored or truncated data. Technical Report, Dept. Statistics, Stanford Univ.

LAI, T. L. and YING, Z. (1991). Estimating a distribution function with truncated and censored data. *Ann. Statist.* **19** 417–442.

LAWLESS, J. F. (1982). *Statistical Models and Methods for Lifetime Data*. Wiley, New York.

PRENTICE, R. L. (1978). Linear rank tests with right censored data. *Biometrika* **65** 167–179.

REBOLLEDO, R. (1980). Central limit theorems for local martingales. *Z. Wahrsch. Verw. Gebiete* **51** 269–286.

RITOV, Y. (1990). Estimation in a linear regression model with censored data. *Ann. Statist.* **18** 303–328.

TOBIN, J. (1958). Estimation of relationships for limited dependent variables. *Econometrica* **26** 24–36.

TSIATIS, A. A. (1990). Estimating regression parameters using linear rank tests for censored data. *Ann. Statist.* **18** 354–372.

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA 94305

DEPARTMENT OF STATISTICS
UNIVERSITY OF ILLINOIS
CHAMPAIGN, ILLINOIS 61820