The Annals of Statistics
1993, Vol. 21, No. 1, 76–99

A LARGE SAMPLE STUDY OF RANK ESTIMATION FOR CENSORED REGRESSION DATA[1]

BY ZHILIANG YING

University of Illinois

Large sample approximations are developed to establish asymptotic linearity of the commonly used linear rank estimating functions, defined as stochastic integrals of counting processes over the whole line, for censored regression data. These approximations lead to asymptotic normality of the resulting rank estimators defined as solutions of the linear rank estimating equations. A second kind of approximations is also developed to show that the estimating functions can be uniformly approximated by certain more manageable nonrandom functions, resulting in a simple condition that guarantees consistency of the rank estimators. This condition is verified for the two-sample problem, thereby extending earlier results by Louis and Wei and Gail, as well as in the case when the underlying error distribution has increasing failure rate, which includes most parametric regression models in survival analysis. Techniques to handle the delicate tail fluctuations are provided and discussed in detail.

1. Introduction. Let $T_1, \dots, T_n$ be a sequence of positive random variables, usually representing survival (failure) times of n patients (items) in a medical (industrial life) study. Let $X_1, \dots, X_n$ be their corresponding $(p \times 1)$ covariates sequence. The accelerated life model [cf. Cox and Oakes (1984) and Kalbfleisch and Prentice (1980)] is to relate the logarithms of survival times, $Y_i = \log T_i$, to their covariates through a system of linear regression equations

$$Y_i = \beta^T X_i + \varepsilon_i, \qquad i = 1, \dots, n, \tag{1.1}$$

where β is a $p \times 1$ parameter vector and the ε_i are conditional on X_i, independent and identically distributed (i.i.d.) random errors with a common distribution function F. The regression model (1.1) for survival data is often complicated by the so-called right-censorship: There exist (log) censoring times C_i, such that we can only observe $Y_i \wedge C_i$, $\delta_i = I_{\{Y_i \le C_i\}}$ and X_i, $i = 1, \dots, n$. Here, we shall assume that conditional on X_i, ε_i and C_i are independent.

For the censored regression (1.1), when the common error distribution F of ε_i is modeled as a member from a specific parametric family of distributions,

Received February 1991; revised April 1992.
[1]Partially supported by the National Science Foundation.
AMS 1991 *subject classifications*. Primary 62J05; secondary 60F05, 62F12, 62G05.
Key words and phrases. Accelerated life model, censored regression, log-rank statistic, asymptotic linearity, consistency, asymptotic normality, counting process.

one can apply the usual maximum likelihood method for statistical inference; compare Lawless (1982). Simple formulas are often obtainable by applying the widely used large sample theory such as the asymptotic normality of the maximum likelihood estimators. Without any parametric assumption, the classical linear rank statistics for testing $\beta = \beta_0$ may be extended and used in testing hypotheses for censored regression data; compare Prentice (1978) and Cuzick (1985). Examples of such extended linear rank statistics include the log-rank and the Gehan (1965) statistics, and the G^ρ family, compare Harrington and Fleming (1982).

In the absence of the right-censorship, it is well known that the linear rank statistics can also be used as estimating functions to construct R-estimates; compare Hodges and Lehmann (1963), Adichie (1967) and Jurečková (1971). Because of discontinuity, these estimating functions are much more difficult to analyze than those of the maximum likelihood in the parametric case. The approaches given by the aforementioned articles are based upon a monotonicity property of the R-statistics and a standard contiguity argument. Extensions of the two-sample log-rank and Gehan statistics are obtained by Louis (1981) and Wei and Gail (1983), who showed that in these two cases, the monotonicity is preserved even in the presence of a right-censorship and therefore the classical method applies. Earlier, Buckley and James (1979) proposed an extension of the least squares estimating equation to handle the censored regression. Ritov (1990) linked this type of estimating equations to a class of weighted log-rank forms and developed certain asymptotic properties.

For the general censored linear regression, the linear rank (in particular, the log-rank) estimating functions are not only discontinuous, but also nonmonotone, and therefore, neither the usual Taylor expansion method, often applied in analyzing a maximum likelihood estimator, nor the contiguity argument can be used. Recently, Tsiatis (1990) and Lai and Ying (1992) proposed certain modifications to the linear rank estimating functions and investigated their large sample properties. Their approaches are based upon establishing local asymptotic linearity properties of these functions to show that the resulting estimators are asymptotically normal. Specifically, Tsiatis (1990) studied the log-rank statistic truncated at T^*, that is,

$$\xi^{TS}(b) = \sum_{i=1}^{n} \int_{-\infty}^{T^*} \left(X_i - \frac{Z^x(b,t)}{Z(b,t)} \right) dN_i(b,t), \tag{1.2}$$

where T^* is a prespecified constant such that

$$\liminf_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} P\{Y_i \wedge C_i - \beta^T X_i \geq T^* + \eta\} > 0 \tag{1.3}$$

for some $\eta > 0$. The role of T^* is to avoid the usual technical difficulty of a possible tail instability. Here and in the sequel, we shall use the notation

$$
\begin{aligned}
&N_i(b,t) = I_{\{Y_i \wedge C_i - b^T X_i \le t,\, \delta_i = 1\}}, \qquad N(b,t) = \sum_{i=1}^n N_i(b,t),\\
&N^x(b,t) = \sum_{i=1}^n X_i N_i(b,t),\\
&\tilde{N}_i(b,t) = I_{\{Y_i \wedge C_i - b^T X_i \ge t,\, \delta_i = 1\}}, \qquad \tilde{N}(b,t) = \sum_{i=1}^n \tilde{N}_i(b,t),\\
&\tilde{N}^x(b,t) = \sum_{i=1}^n X_i \tilde{N}_i(b,t),\\
&Z_i(b,t) = I_{\{Y_i \wedge C_i - b^T X_i \ge t\}}, \qquad Z(b,t) = \sum_{i=1}^n Z_i(b,t),\\
&Z^x(b,t) = \sum_{i=1}^n X_i Z_i(b,t).
\end{aligned} \tag{1.4}
$$

His main result is to show that there exists a nonrandom $p \times p$ matrix A^{TS} such that

$$
\frac{1}{\sqrt{n}} \xi^{TS}(b) = \frac{1}{\sqrt{n}} \xi^{TS}(\beta) + A^{TS} \sqrt{n}\,(b - \beta) + o_p(1), \tag{1.5}
$$

uniformly in $\|b - \beta\| \le B/\sqrt{n}$, for every fixed constant B. Here and in the sequel, we shall follow the convention that $\|v\|$ of a vector v denotes its Euclidean norm and $\|V\|$ of a matrix $V = (v_{ij})$ denotes $\sqrt{\Sigma v_{ij}^2}$. Equation (1.5) means that the function $\xi^{TS}(b)$ is asymptotically linear in the $n^{-1/2}$ neighborhood of the true regression parameter β. From (1.5), it follows that a solution of $\xi^{TS}(b)/\sqrt{n} \approx 0$ exists in the $n^{-1/2}$ neighborhood of β that is also asymptotically normal. Instead of truncating at T^*, the approach of Lai and Ying (1992), for the case of log-rank statistic, is to put a weight function w_n and to consider

$$
\xi^{LY}(b) = \sum_{i=1}^n \int_{-\infty}^{\infty} w_n(b,t) \left(X_i - \frac{Z^x(b,t)}{Z(b,t)} \right) dN_i(b,t). \tag{1.6}
$$

The weight function $w_n(b,t)$ is constructed from the data and typically takes value 1 if the risk size $Z(b,t)$ is $\ge n^\delta$ for some $\delta < 1$ and 0 if $Z(b,t)$ is much smaller than n^δ.

In spite of these efforts, many important issues still remain to be resolved. Listed below are five major ones which seem to be necessary for any comprehensive resolution.

(a) Replacing T^* in (1.2) by ∞. A major drawback of Tsiatis' (1990) result is that it requires a known T^*, a rather unrealistic assumption. In fact, a reasonable choice of T^* can only be made if one has some knowledge of F and β. Moreover, if T^* is chosen to be too small, a substantial portion of the information in the data set may be lost. Lai and Ying's (1992) approach, which though constructs w_n adaptively from the data and can deal with truncated data, is also rather unpleasant for the simple, such as the log-rank, estimating functions since one has to choose an appropriate weight function w_n to dampen the tail instability. In view of this, it is important to study the *original* log-rank estimating function

$$\xi(b) = \sum_{i=1}^{n} \int_{-\infty}^{\infty} \left(X_i - \frac{Z^x(b,t)}{Z(b,t)} \right) dN_i(b,t). \tag{1.7}$$

(b) Approximating $n^{-1}\xi(\cdot)$ by a nonrandom function $m(\cdot)$. In order to address the issue of consistency, the global behavior of the random function $n^{-1}\xi(\cdot)$ has to be studied. This can be done by developing uniform approximation of $n^{-1}\xi$ by a nonrandom function m, so that consistency of the rank estimate is ensured by showing that the much simpler nonrandom m has a unique root. Note that $m(\beta) = 0$.

(c) Establishing asymptotic linearity

$$\frac{1}{\sqrt{n}}\xi(b) \approx \frac{1}{\sqrt{n}}\xi(\beta) + A\sqrt{n}(b - \beta) \tag{1.8}$$

for b in any shrinking neighborhood. This will ensure that any consistent root is also asymptotically normal, provided that the slope matrix A is nonsingular. Note that if (b) is settled, the consistency reduces to verifying that m has a single root. However, linearity in the $n^{-1/2}$ neighborhood is not sufficient to ensure normality from consistency.

(d) Checking under what condition m has a unique root. In view of (b), this is crucial to proving consistency.

(e) Checking whether the slope matrix A in (1.8) is nonsingular. This is crucial to usefulness of (1.8).

This paper tackles all five issues raised above for, in fact, more general weighted log-rank estimating functions. Specifically, all the results will be established for the estimating functions without the unpleasant upper limit T^*. It will be shown that the asymptotic linearity (1.8) holds in any shrinking neighborhood in the most general sense one would hope for and that the random estimating function $n^{-1}\xi$ can indeed be approximated uniformly by its nonrandom limit. It will also be shown that the slope matrix A is always nonnegative definite and, under an extremely mild condition, is actually positive definite. Finally, it will be verified that the nonrandom limit function m has a unique root in the case of the two-sample problem, thereby extending the results of two earlier papers by Louis (1981) and Wei and Gail (1983) to the general weighted log-rank estimators, and the case when the error distribution has an increasing failure rate.

The paper is organized as follows. In Section 2, we prove that asymptotic linearity holds almost surely for the log-rank estimating functions when b is close to β. An immediate consequence of this result is that there exists a fixed neighborhood of β, within which the rank estimator exists and is strongly consistent and that any consistent rank estimator is also asymptotically normal. These results are extended in Section 3 to the more general weighted log-rank estimating functions ξ_ϕ. In Section 4, we show that within any bounded region, $n^{-1}\xi(b)$ can be uniformly approximated by a nonrandom function. By defining rank estimator as a minimizer of $\|\xi(b)\|$, such an approximation leads to a simple condition that guarantees the consistency and, together with the results of Section 2, the asymptotic normality of the rank estimator. Similar results are also proved for the rank estimators using the weighted log-rank estimating functions ξ_ψ. The paper concludes in Section 5 with two important special cases.

2. Asymptotic linearity of the log-rank estimating function and asymptotic normality of the resulting estimator. In this section, we establish the asymptotic linearity of the log-rank estimating function $\xi(b)$ defined by (1.7) for b in a neighborhood of the true parameter β. This result is then used to show the existence of a fixed neighborhood that guarantees the consistency and asymptotic normality of the resulting rank estimator. Because of its delicacy, we shall provide sufficient technical details in our proof. Since all of our developments are conditional on the covariates X_i, we shall assume that the X_i are nonrandom.

First we introduce the following conditions.

CONDITION 1. The covariates are uniformly bounded, and without loss of generality we assume that $\sup_i\|X_i\| \le 1$.

CONDITION 2. The error density f and its derivative f' are bounded and $\int(f'(t)/f(t))^2 f(t)\,dt < \infty$.

CONDITION 3. The C_i have uniformly bounded densities g_i, that is, there exists B_c such that $|g_i(t)| \le B_c$ for all t and i.

CONDITION 4. $\sup_i E|\min\{\varepsilon_i, C_i\}|^{\theta_0} < \infty$, for some $\theta_0 > 0$.

Condition 1 is the same as the Condition (D) in Tsiatis [(1990), page 358]. As will be commented following Lemma 1, this condition can certainly be relaxed to $\sup_{i\le n}\|X_i\| = O(n^\varepsilon)$ for every $\varepsilon > 0$. Condition 3 is the same as the Condition (B) in Tsiatis [(1990), page 357] and seems to be the most restrictive among all four conditions.

We shall use F and G_i to denote the distribution functions of ε_i and C_i and use $\overline{F}$ and $\overline{G}_i$ to denote their survival functions. Let $\lambda = f/\overline{F}$ be the hazard rate of ε_i. Moreover, let

$$\Gamma_{n,k}(t) = \frac{1}{n}\sum_{i=1}^{n} X_i^k \overline{G}_i(t + \beta^T X_i), \qquad k = 0, 1, 2, \tag{2.1}$$

where $X_i^0 = 1$, $X_i^1 = X_i$ and $X_i^2 = X_i X_i^T$, and

$$\text{(2.2a)} \qquad A_n = \int_{-\infty}^{\infty} \left[\Gamma_{n,2}(t) - \frac{\Gamma_{n,1}(t)\Gamma_{n,1}^T(t)}{\Gamma_{n,0}(t)} \right] \frac{\lambda'(t)}{\lambda(t)}\, dF(t),$$

$$\text{(2.2b)} \qquad \Sigma_n = \int_{-\infty}^{\infty} \left[\Gamma_{n,2}(t) - \frac{\Gamma_{n,1}(t)\Gamma_{n,1}^T(t)}{\Gamma_{n,0}(t)} \right] dF(t).$$

For matrices A_n, Σ_n and $R_n(t) = \Gamma_{n,2}(t) - \Gamma_{n,1}(t)\Gamma_{n,1}^T(t)/\Gamma_{n,0}(t)$ we claim that the following properties hold:

1. The $R_n(t)$ are nonnegative definite, and nonincreasing in the sense that $t \geq t'$ implies $R_n(t') - R_n(t)$ is nonnegative definite [denoted hereafter by $R_n(t') \geq R_n(t)$].
2. A_n is nonnegative definite.
3. If the eigenvalues of $R_n(t)$ are bounded away from 0 for all large n and some t in the support of f, then the eigenvalues of A_n and Σ_n are also bounded away from 0 for all large n.
4. The eigenvalues of A_n are bounded away from 0 if and only if the eigenvalues of Σ_n are bounded away from 0.

To justify these claims, note that we can write

$$R_n(t) = \frac{1}{n} \sum_{i=1}^{n} \left(X_i - \frac{\sum_{i=1}^n X_i \overline{G}_i(t + \beta^T X_i)}{\sum_{i=1}^n \overline{G}_i(t + \beta^T X_i)} \right) \times \left(X_i - \frac{\sum_{i=1}^n X_i \overline{G}_i(t + \beta^T X_i)}{\sum_{i=1}^n \overline{G}_i(t + \beta^T X_i)} \right)^T \overline{G}_i(t + \beta^T X_i) \geq 0.$$

Moreover, for $t' \leq t$,

$$\begin{aligned} R_n(t') &\geq \frac{1}{n} \sum_{i=1}^{n} \left(X_i - \frac{\sum_{i=1}^n X_i \overline{G}_i(t' + \beta^T X_i)}{\sum_{i=1}^n \overline{G}_i(t' + \beta^T X_i)} \right) \\ &\qquad \times \left(X_i - \frac{\sum_{i=1}^n X_i \overline{G}_i(t' + \beta^T X_i)}{\sum_{i=1}^n \overline{G}_i(t' + \beta^T X_i)} \right)^T \overline{G}_i(t + \beta^T X_i) \\ &\geq \frac{1}{n} \sum_{i=1}^{n} \left(X_i - \frac{\sum_{i=1}^n X_i \overline{G}_i(t + \beta^T X_i)}{\sum_{i=1}^n \overline{G}_i(t + \beta^T X_i)} \right) \\ &\qquad \times \left(X_i - \frac{\sum_{i=1}^n X_i \overline{G}_i(t + \beta^T X_i)}{\sum_{i=1}^n \overline{G}_i(t + \beta^T X_i)} \right)^T \overline{G}_i(t + \beta^T X_i), \end{aligned}$$

where the last inequality follows from an ANOVA-type decomposition. Now write

$$A_n = \int_{-\infty}^{\infty} R_n(t)\left[f'(t) + \frac{f^2(t)}{1 - F(t)}\right] dt$$
$$= \int_{-\infty}^{\infty} R_n(t)\frac{f^2(t)}{1 - F(t)}\, dt + \int_{-\infty}^{\infty} f(t)\, d(-R_n(t)),$$

which clearly implies 2–4.

THEOREM 1. *Under Conditions* 1–4, *the log-rank estimating function* $\xi(b)$ *is asymptotically linear in the sense that, for every sequence* $d_n > 0$ *with* $d_n \to 0$ *a.s.*,

$$\sup_{\|b-\beta\|\le d_n} \left\{\|\xi(b) - \xi(\beta) - A_n n(b - \beta)\|/\left(\sqrt{n} + n\|b - \beta\|\right)\right\} = o(1) \quad a.s.$$

In particular, if $0 < \tilde{d}_n \to 0$ *in probability, then*

$$\sup_{\|b-\beta\|\le \tilde{d}_n} \left\{\|\xi(b) - \xi(\beta) - A_n n(b - \beta)\|/\left(\sqrt{n} + n\|b - \beta\|\right)\right\} = o_p(1).$$

Theorem 1 differs and improves the result of Tsiatis [(1990), Theorems 3.1 and 3.2] in several ways. First, it is for the usual log-rank statistic ξ, which is not truncated at the tail. Second, it provides the asymptotic linearity for ξ in any shrinking neighborhood of β, rather than in the $n^{-1/2}$ neighborhood. As we have mentioned earlier, this is crucial for translating consistency into asymptotic normality for the rank estimator. It is also useful for conducting statistical inferences in some situations in which neighborhoods larger than $n^{-1/2}$ are needed; compare Wei, Ying and Lin (1990). Finally, the asymptotic linearity holds almost surely.

The "a.s. linearity" implies "in probability linearity" because of a well-known result that $\tilde{d}_n = o_p(1)$ if and only if for every subsequence n_k there exists a sub-subsequence n_{kl} such that $\tilde{d}_{n_{kl}} = o(1)$ a.s.

With the additional assumption that A_n is eventually nonsingular, which has been shown to be almost always satisfied, Theorem 1 can be used to characterize the local behavior (near β) of the resulting rank estimator. This is given by the following corollary.

COROLLARY 1. *Suppose that Conditions* 1–4 *are satisfied and that all the eigenvalues of* A_n *are bounded away from zero for all large* n.

(i) *There exists a closed neighborhood* $\mathscr{N}$ *containing* β *as its interior point such that* $\hat{\beta}$, *defined as a solution of* $\|\xi(\hat{\beta})\| = \min_{b\in\mathscr{N}}\|\xi(b)\|$, *is strongly consistent.*

(ii) *For any* $\mathscr{N}$ *containing* β *as its interior point defining* $\hat{\beta}$ *as in* (i), *if* $\hat{\beta}$ *is consistent, then*

$$\sqrt{n}\,\Sigma_n^{-1/2}A_n(\hat{\beta} - \beta) \to_{\mathscr{D}} N(0, I_p).$$

In particular, if $\Gamma_{n,k}(t) \to \Gamma_k(t)$ *exist for* $k = 0, 1, 2$ *and all* t, *then*

$$A_n \to A = \int_{-\infty}^{\infty} \left[\Gamma_2(t) - \frac{\Gamma_1(t)\Gamma_1^T(t)}{\Gamma_0(t)} \right] \frac{\lambda'(t)}{\lambda(t)}\, dF(t), \tag{2.3}$$

$$\Sigma_n \to \Sigma = \int_{-\infty}^{\infty} \left[\Gamma_2(t) - \frac{\Gamma_1(t)\Gamma_1^T(t)}{\Gamma_0(t)} \right] dF(t), \tag{2.4}$$

and

$$\sqrt{n}\,(\hat{\beta} - \beta) \to_{\mathscr{D}} N(0, A^{-1}\Sigma A^{-1}).$$

We need a few lemmas for the proofs of Theorem 1 and Corollary 1. Without loss of generality, we shall assume $\beta = 0$ throughout the rest of this section.

LEMMA 1. *Let* V_i *be a bounded sequence of constants. Under Conditions* 1–4 *for every* $\gamma \in [0, 1)$, $B > 0$, $K > 0$ *and* $\theta > 0$, *with probability* 1,

$$\sup_{\|b\| \le B,\, E\nu(b,t) \le Kn^{1-\gamma}} \left| \sum_{i=1}^{n} V_i [\nu_i(b,t) - E\nu_i(b,t)] \right| = o(n^{(1-\gamma)/2+\theta}), \tag{2.5a}$$

$$\sup_{\|b-b'\| + |t-t'| \le Kn^{-\gamma}} \left| \sum_{i=1}^{n} V_i [\nu_i(b,t) - E\nu_i(b,t) - \nu_i(b',t') + E\nu_i(b',t')] \right| = o(n^{(1-\gamma)/2+\theta}), \tag{2.5b}$$

where ν_i *is any one of* Z_i, N_i *or* $\tilde{N}_i$ *defined by* (1.4) *and* $\nu = \sum_{i=1}^{n} \nu_i$.

The preceding lemma is a special case of Theorem 1 of Lai and Ying (1988). The proof given there essentially uses Bennett's (1962) exponential inequality, or, more precisely, its extension by Alexander (1984). It says that the order of these weighted empirical processes, centered at their means, are bounded by their standard deviations multiplied by n^{θ}. Note that if $\sup_{i \le n} |V_i| = o(n^{\varepsilon})$ for every $\varepsilon > 0$, then (2.5) still holds since ε can always be absorbed into θ. Continuing to track orders in this way, we can discover that all approximations we shall establish hold when Condition 1 is relaxed to $\sup_{i \le n} \|X_i\| = o(n^{\varepsilon})$.

LEMMA 2. *Let* $t_b(\alpha) = \inf\{t\colon EZ(b,t) \le n^{1-\alpha}\}$ *with* $0 < \alpha < 1$. *Then under Conditions* 1–4, *for every* $\theta > 0$ *and* $B > 0$,

$$\sup_{\|b\| \le B,\, t \ge t_b(\alpha)} \left\| \int_{t_b(\alpha)}^{t} \sum_{i=1}^{n} \left[X_i - \frac{Z^x(b,s)}{Z(b,s)} \right] dN_i(b,s) - \int_{t_b(\alpha)}^{t} \sum_{i=1}^{n} \left[X_i - \frac{EZ^x(b,s)}{EZ(b,s)} \right] dEN_i(b,s) \right\| = o(n^{(1-\alpha)/2+\theta}) \quad a.s. \tag{2.6}$$

In particular, the left-hand side of (2.6) *is* $o(n^{1/2})$ *a.s.*

PROOF. From Lemma 1 and the fact that $E\tilde{N}(b,t) \le EZ(b,t) \le n^{1-\alpha}$ for $t \ge t_b(\alpha)$, we get, for every $\theta > 0$,

$$\begin{aligned}(2.7)\quad \sup_{\|b\|\le B,\, t\ge t_b(\alpha)} &\big\{|\tilde{N}(b,t) - E\tilde{N}(b,t)| + \|\tilde{N}^x(b,t) - E\tilde{N}^x(b,t)\| \\ &+ |Z(b,t) - EZ(b,t)| + \|Z^x(b,t) - EZ^x(b,t)\|\big\} \\ &= o(n^{(1-\alpha)/2+\theta}) \quad \text{a.s.}\end{aligned}$$

Now $dN_i(b,t) = -d\tilde{N}_i(b,t)$, which implies that

$$\begin{aligned}&\text{l.h.s. of } (2.6) \\ &\le \sup_{\|b\|\le B,\, t\ge t_b(\alpha)} \|\tilde{N}^x(b,t) - E\tilde{N}^x(b,t) - \tilde{N}^x(b,t_b) + E\tilde{N}^x(b,t_b)\| \\ (2.8)\quad &+ \sup_{\|b\|\le B,\, t\ge t_b} \left\| \int_{t_b}^{t} \frac{Z^x(b,s)}{Z(b,s)}\, d\big[\tilde{N}(b,s) - E\tilde{N}(b,s)\big] \right\| \\ &+ \sup_{\|b\|\le B,\, t\ge t_b} \left\| \int_{t_b}^{t} \left[\frac{Z^x(b,s)}{Z(b,s)} - \frac{EZ^x(b,s)}{EZ(b,s)}\right] dE\tilde{N}(b,s) \right\|.\end{aligned}$$

From (2.7), the first term on the right-hand side of (2.8) is $o(n^{(1-\alpha)/2+\theta})$ a.s. The second term is also $o(n^{(1-\alpha)/2+\theta})$ by applying the integration by parts formula [cf. Gill (1980), page 153] together with (2.7) and the fact that the total variation

$$(2.9)\quad \sup_b \int_{-\infty}^{\infty} \left| d\frac{Z^x(b,s)}{Z(b,s)} \right| \le 2p \sup_b \int_{-\infty}^{\infty} \frac{-dZ(b,s)}{Z(b,s)} = O(\log n),$$

where $\int |dv(t)|$ of a vector function $v(t)$ denotes the sum of the total variations of all its components. Recall that p is the dimension of X_i.

Thus it remains to control the last term on the right-hand side of (2.8). Choose $1 > \alpha^* > (1+\alpha)/2$ and $t_b(\alpha^*) = \inf\{t\colon\ EZ(b,t) \le n^{1-\alpha^*}\}$. Then $E\tilde{N}(b,t_b(\alpha^*)) \le EZ(b,t_b(\alpha^*)) \le n^{1-\alpha^*} = o(n^{(1-\alpha)/2})$, implying

$$\begin{aligned}(2.10)\quad &\sup_{\|b\|\le B,\, t\ge t_b(\alpha)} \left\| \int_{t_b(\alpha)}^{t} \left[\frac{Z^x(b,s)}{Z(b,s)} - \frac{EZ^x(b,s)}{EZ(b,s)}\right] dE\tilde{N}(b,s) \right\| \\ &\le -\int_{t_b(\alpha)}^{t_b(\alpha^*)} \left\| \frac{Z^x(b,s)}{Z(b,s)} - \frac{EZ^x(b,s)}{EZ(b,s)} \right\| dEZ(b,s) + o(n^{(1-\alpha)/2}).\end{aligned}$$

Since

$$\begin{aligned}\left\| \frac{Z^x(b,s)}{Z(b,s)} - \frac{EZ^x(b,s)}{EZ(b,s)} \right\| &\le \left\| \frac{Z^x(b,s)(Z(b,s) - EZ(b,s))}{Z(b,s)EZ(b,s)} \right\| \\ &\quad + \left\| \frac{Z^x(b,s) - EZ^x(b,s)}{EZ(b,s)} \right\| \\ &\le \frac{|Z(b,s) - EZ(b,s)|}{EZ(b,s)} + \frac{\|Z^x(b,s) - EZ^x(b,s)\|}{EZ(b,s)},\end{aligned}$$

it follows from (2.7) again that the second term on the right-hand side of (2.10) is of the order $o(n^{(1-\alpha)/2+\theta})$ a.s., noting that $\sup_b \int_{t_b(\alpha)}^{t_b(\alpha^*)} |dE\,Z(b,s)/EZ(b,s)| = O(\log n)$. Hence the right-hand side of (2.8) is $o(n^{(1-\alpha)/2+\theta})$ and the desired conclusion (2.6) follows. □

LEMMA 3. *Suppose that Conditions* 1–4 *are satisfied. Define*

$$\xi(b,t) = \int_{-\infty}^{t} \sum_{i=1}^{n} \left[X_i - \frac{Z^x(b,s)}{Z(b,s)} \right] dN_i(b,s),$$

$$\zeta(b,t) = \int_{-\infty}^{t} \sum_{i=1}^{n} \left[X_i - \frac{EZ^x(b,s)}{EZ(b,s)} \right] dE\,N_i(b,s).$$

(i) *For every* $B > 0$ *and every* $\theta > 0$,

$$\sup_{\|b\| \le B,\, t \in R^1} \|\xi(b,t) - \zeta(b,t)\| = o(n^{1/2+\theta}) \quad a.s. \tag{2.11}$$

(ii) *There exists* $\theta_0 > 0$ *such that*

$$\sup_{\|b\| \le n^{-1/3},\, t \in R^1} \|\xi(b,t) - \xi(0,t) - \zeta(b,t)\| = o(n^{1/2-\theta_0}) \quad a.s. \tag{2.12}$$

The proof given below is basically in the same spirit as that of Theorem 2 of Lai and Ying [(1988), pages 346–348] for a slightly more general setup. However, we will present all the key steps here, partly for its completeness and partly to show that the condition (3.1) there is not needed in our setting.

PROOF OF LEMMA 3. From the definitions of $\xi(b,t)$ and $\zeta(b,t)$ and Lemma 2,

$$\begin{aligned}
&\sup_{\|b\| \le B,\, t \in R^1} \|\xi(b,t) - \zeta(b,t)\| \\
&\quad \le \sup_{\|b\| \le B,\, t \le t_b(\alpha)} \{\|N^x(b,t) - EN^x(b,t)\|\} \\
&\qquad + \sup_{\|b\| \le B,\, t \le t_b(\alpha)} \left\{ \int_{-\infty}^{t} \frac{\|Z^x(b,s) - EZ^x(b,s)\| + |Z(b,s) - EZ(b,s)|}{Z(b,s)} \times dN(b,s) \right\} \\
&\qquad + \sup_{\|b\| \le B,\, t \le t_b(\alpha)} \left\{ \left\| \int_{-\infty}^{t} \frac{EZ^x(b,s)}{EZ(b,s)} \, d[N(b,t) - EN(b,t)] \right\| \right\} + o(n^{1/2}) \\
&\quad = o(n^{1/2+\theta}) \quad \text{a.s.},
\end{aligned}$$

where the last equality follows from (2.5a) (with $\gamma = 0$),

$$\sup_b \int_{-\infty}^{\infty} dN(b,s)/Z(b,s) = O(\log n)$$

and

$$\sup_b \int_{-\infty}^{t_b(\alpha)} \left| d\frac{EZ^x(b,s)}{EZ(b,s)} \right| = O(\log n).$$

For (ii), let $t_0 = \inf\{t: EZ(0,t) \le n^{1-\alpha_0}\}$ with $0 < \alpha_0 < 1/6$. Note that from Conditions 1–3 it is easy to check that for any $\alpha < \alpha_0$, $\sup\{t_b(\alpha): \|b\| \le n^{-1/3}\} \le t_0$ for all large n. Thus, in view of Lemma 2, it suffices to show (2.12) with t restricted to $t \le t_0$. Note also that $\zeta(0,t) = 0$ for all t. This can be used to verify that

$$\begin{aligned}
&\xi(b,t) - \xi(0,t) - \zeta(b,t) \\
&\quad = [N^x(b,t) - EN^x(b,t) - N^x(0,t) + EN^x(0,t)] \\
&\qquad - \int_{-\infty}^{t} \frac{Z^x(0,s)}{Z(0,s)}\, d[N(b,s) - EN(b,s) - N(0,s) + EN(0,s)] \\
&\qquad - \int_{-\infty}^{t} \left[\frac{Z^x(b,s)}{Z(b,s)} - \frac{EZ^x(b,s)}{EZ(b,s)} - \frac{Z^x(0,s)}{Z(0,s)} + \frac{EZ^x(0,s)}{EZ(0,s)} \right] dN(b,s) \\
&\qquad - \int_{-\infty}^{t} \left[\frac{EZ^x(b,s)}{EZ(b,s)} - \frac{EZ^x(0,s)}{EZ(0,s)} \right] d[N(b,s) - EN(b,s)] \\
&\qquad - \int_{-\infty}^{t} \left[\frac{Z^x(0,s)}{Z(0,s)} - \frac{EZ^x(0,s)}{EZ(0,s)} \right] d[EN(b,s) - EN(0,s)] \\
&\quad = Q_1(b,t) + Q_2(b,t) + Q_3(b,t) + Q_4(b,t) + Q_5(b,t), \quad \text{say.}
\end{aligned} \tag{2.13}$$

From Lemma 1, for every $\theta > 0$,

$$\sup_{\|b\| \le n^{-1/3},\, t \le t_0} \|Q_1(b,t)\| = o(n^{1/3+\theta}) \quad \text{a.s.} \tag{2.14}$$

From Lemma 1, the integration by parts formula and (2.9), for every $\theta > 0$,

$$\sup_{\|b\| \le n^{-1/3},\, t \le t_0} \|Q_2(b,t)\| = o(n^{1/3+\theta}) \quad \text{a.s.} \tag{2.15}$$

By a tedious but otherwise straightforward manipulation to express it in an appropriate form so that Lemma 1 can be used, it can be shown that for every $\theta > 0$,

$$\begin{aligned}
&\left\| \frac{Z^x(b,s)}{Z(b,s)} - \frac{EZ^x(b,s)}{EZ(b,s)} - \frac{Z^x(0,s)}{Z(0,s)} + \frac{EZ^x(0,s)}{EZ(0,s)} \right\| \\
&\qquad \le \frac{n^{1/2-(1/3-\alpha_0)+\theta}}{Z(b,s)}
\end{aligned} \tag{2.16}$$

for all large n and all $s \le t_0$ and $\|b\| \le n^{-1/3}$. From (2.16) and the fact that $\sup_{\|b\| \le n^{-1/3}} \int_{-\infty}^{t_0} dN(b,s)/Z(b,s) = O(\log n)$, we have

$$(2.17) \qquad \sup_{\|b\| \le n^{-1/3},\, t \le t_0} \|Q_3(b,t)\| = o(n^{1/2}) \quad \text{a.s.}$$

Now $\sup_{\|b\| \le n^{-1/3},\, t \le t_0} |EZ(b,t)/EZ(0,t) - 1| = o(1)$ a.s., which can be used to show that for all large n, $\|a\| + \|b\| \le n^{-1/3}$ and $\max\{t, s\} \le t_0$,

$$(2.18) \qquad \left\| n^{-\alpha} \frac{EZ^x(b,t)}{EZ(b,t)} - n^{-\alpha} \frac{EZ^x(a,s)}{EZ(a,s)} \right\| \le K_0(\|b-a\| + |t-s|)$$

for some $K_0 > 0$. Therefore, we can apply Lemma 3 of Lai and Ying (1988) to get for every $\theta > 0$

$$(2.19) \qquad \sup_{\|b\| \le n^{-1/3},\, t \le t_0} \|Q_4(b,t)\| = o(n^{(1-1/3)/2+\alpha+\theta}) \quad \text{a.s.}$$

Finally, since

$$\left\| \frac{Z^x(0,s)}{Z(0,s)} - \frac{EZ^x(0,s)}{EZ(0,s)} \right\| \le \frac{\|Z^x(0,s) - EZ^x(0,s)\|}{EZ(0,s)} + \frac{|Z(0,s) - EZ(0,s)|}{EZ(0,s)},$$

we can apply Lemma 1 to get for every $\theta > 0$,

$$(2.20) \qquad \begin{aligned} &\sup_{\|b\| \le n^{-1/3},\, t \le t_0} \|Q_5(b,t)\| \\ &= o(n^{1/2+\alpha+\theta}) \int_{-\infty}^{t_0} \frac{1}{n} \sum_{i=1}^{n} |\overline{G}_i(s + b^T X_i) f(s + b^T X_i) \\ &\qquad\qquad - \overline{G}_i(s) f(s)|\, ds \\ &= o(n^{1/2+\alpha-1/3+\theta}) \int_{-\infty}^{\infty} [f(s) + |f'(s)|]\, ds \quad \text{a.s.} \end{aligned}$$

Therefore (2.12) follows from (2.13)–(2.15), (2.17), (2.19) and (2.20). □

LEMMA 4. *Under Condition* 2, $f(t) = o(\overline{F}^{1/2}(t))$ *as* $t \uparrow \tau_F$, *where* $\tau_F = \sup\{t\colon \overline{F}(t) > 0\}$.

PROOF. By the Cauchy–Schwarz inequality,

$$f^2(t) = \left[\int_t^{\tau_F} \frac{f'(s)}{f(s)}\, dF(s) \right]^2 \le \int_t^{\tau_F} \left[\frac{f'(s)}{f(s)} \right]^2 dF(s) \overline{F}(t) = o(\overline{F}(t))$$

as $t \uparrow \tau_F$. □

LEMMA 5. *Suppose that Conditions* 1–3 *are satisfied. Then*

$$\int_{-\infty}^{t}\sum_{i=1}^{n}\left[X_i - \frac{EZ^x(b,s)}{EZ(b,s)}\right] dEN_i(b,s) = n\left[A_n(t)b + \eta_n(t,b)\right], \tag{2.21}$$

where $\sup\{\|\eta_n(t,b)\|/\|b\|: t \in R^1, 0 < \|b\| \le d_n\} \to 0$ *for any* $d_n \downarrow 0$, *and where*

$$A_n(t) = \int_{-\infty}^{t}\left[\Gamma_{n,2}(s) - \frac{\Gamma_{n,1}(s)\Gamma_{n,1}^T(s)}{\Gamma_{n,0}(s)}\right]\frac{\lambda'(s)}{\lambda(s)}\, dF(s). \tag{2.22}$$

PROOF. Let $\tilde{t}_{1,b} = \inf\{t: EZ(b,t) \le \|b\|^{4/3}\}$. Then

$$\int_{\tilde{t}_{1,b}}^{\infty}\sum_{i=1}^{n}\left\|X_i - \frac{EZ^x(b,s)}{EZ(b,s)}\right\| dEN_i(b,s) \le 2n\|b\|^{4/3}. \tag{2.23}$$

Moreover, let $\tilde{t}_{2,b} = \inf\{t: EZ(b,t) \le \|b\|^{1/3}\}$. Then

$$\begin{aligned}
&\int_{\tilde{t}_{2,b}}^{t}\sum_{i=1}^{n}\left[X_i - \frac{EZ^x(b,s)}{EZ(b,s)}\right]\overline{G}_i(s + b^TX_i)f(s + b^TX_i)\, ds\\
&= \int_{\tilde{t}_{2,b}}^{t}\sum_{i=1}^{n}\left[X_i - \frac{EZ^x(b,s)}{EZ(b,s)}\right]\overline{G}_i(s + b^TX_i)\\
&\qquad \times \overline{F}(s + b^TX_i)\left[\lambda(s + b^TX_i) - \lambda(s)\right] ds\\
&= n\int_{\tilde{t}_{2,b}}^{t}\frac{1}{n}\sum_{i=1}^{n}\left[X_i - \frac{EZ^x(b,s)}{EZ(b,s)}\right]X_i^T\overline{G}_i(s + b^TX_i)\\
&\qquad \times \frac{\overline{F}(s + b^TX_i)}{\overline{F}(s + b^{*T}X_i)}\frac{\lambda'(s + b^{*T}X_i)}{\lambda(s + b^{*T}X_i)}\, dF(s + b^{*T}X_i)b\\
&= n\eta_n(\tilde{t}_{2,b}, t, b), \quad \text{say},
\end{aligned} \tag{2.24}$$

where b^* lies between 0 and b. Now $\overline{F}(s + \Delta) = \overline{F}(s) - f(s)\Delta - (f'(s^*)/2)\Delta^2$ for some $s^* \in [s, s + \Delta]$ by the Taylor expansion and $f(s) = o(\overline{F}^{1/2}(s))$ by Lemma 4. Therefore

$$\sup_{|\Delta| \le \|b\|,\ \overline{F}(s) \ge \|b\|^{4/3}}\left|\frac{\overline{F}(s + \Delta)}{\overline{F}(s)} - 1\right| \to 0 \quad \text{as } \|b\| \to 0. \tag{2.25}$$

From (2.25), we have for all b close to 0,

$$\sup_{\tilde{t}_{2,b}\le t\le \tilde{t}_{1,b}} \{\|\eta_n(\tilde{t}_{2,b},t,b)\|/\|b\|\}$$

$$\le 3\int_{\tilde{t}_{2,b}}^{\tilde{t}_{1,b}} \frac{1}{n}\sum_{i=1}^{n} \overline{G}_i(s+b^TX_i)\left|\frac{\lambda'(s+b^{*T}X_i)}{\lambda(s+b^{*T}X_i)}\right| dF(s+b^{*T}X_i)$$

$$\le 3\Bigg[\frac{1}{n}\sum_{i=1}^{n}\int_{\tilde{t}_{2,b}}^{\tilde{t}_{1,b}}\overline{G}_i(s+bX_i)^2\,dF(s+b^{*T}X_i)$$

$$\times\int_{\tilde{t}_{2,b}}^{\tilde{t}_{1,b}}\left(\frac{\lambda'(s+b^{*T}X_i)}{\lambda(s+b^{*T}X_i)}\right)^2 dF(s+b^{*T}X_i)\Bigg]^{1/2}$$

$$\le 3\left[\int_{-\infty}^{\infty}\left(\frac{f'(t)}{f(t)}\right)^2 dF(t)\frac{1}{n}\sum_{i=1}^{n}\overline{F}(\tilde{t}_{2,b}-\|b\|)\overline{G}_i(\tilde{t}_{2,b}-\|b\|)\right]^{1/2},$$

which converges to 0 as $\|b\|\to 0$, noting that $\int(\lambda'(t)/\lambda(t))^2\,dF(t) = \int(f'(t)/f(t))^2\,dF(t)$ [Efron and Johnstone (1990) and Ritov and Wellner (1988)]. This in conjunction with (2.23) and (2.24) implies

$$\sup_{t\ge\tilde{t}_{2,b}}\{\|\eta_n(\tilde{t}_{2,b},t,b)\|/\|b\|\}\to 0 \quad \text{as } \|b\|\to 0. \tag{2.26}$$

From (2.26) and the fact that $\sup_{n,t\ge\tilde{t}_{2,b}}\|A_n(t)-A_n(\tilde{t}_{2,b})\|\to 0$ as $\|b\|\to 0$, it follows that we only need to show (2.21) for $t\le\tilde{t}_{2,b}$, that is,

$$\sup_{t\le\tilde{t}_{2,b},\,0<\|b\|\le d_n}\{\|\eta_n(t,b)\|/\|b\|\}\to 0. \tag{2.27}$$

Similar to (2.24), we have

$$\begin{aligned}\zeta(b,t) = \sum_{i=1}^{n}\int_{-\infty}^{t}&\left[X_i-\frac{EZ^x(b,s)}{EZ(b,s)}\right]X_i^T\overline{G}_i(s+b^TX_i)\\ &\times\frac{\overline{F}(s+b^TX_i)}{\overline{F}(s+b^{*T}X_i)}\frac{\lambda'(s+b^{*T}X_i)}{\lambda(s+b^{*T}X_i)}\,dF(s+b^{*T}X_i)b,\end{aligned}\tag{2.28}$$

where b^* lies between 0 and b. From the definition of $\tilde{t}_{2,b}$, we can easily show that, uniformly in $\{X_j\}$,

$$\sup_{s\le\tilde{t}_{2,b}}\left\{\left\|\frac{EZ^x(b,s)}{EZ(b,s)}-\frac{EZ^x(0,s)}{EZ(0,s)}\right\|+\left|\frac{\overline{F}(s+b^TX_i)}{\overline{F}(s+b^{*T}X_i)}-1\right|\right\}\to 0. \tag{2.29}$$

From (2.28) and (2.29), as $\|b\| \to 0$,

$$\zeta(b,t) = (1+o(1))\frac{1}{n}\sum_{i=1}^{n}\int_{-\infty}^{t+b^{*T}X_i}\left[X_i^2\overline{G}_i(s) - \frac{\Gamma_{n,1}(s)}{\Gamma_{n,0}(s)}X_i^T\overline{G}_i(s)\right]\frac{\lambda'(s)}{\lambda(s)}\,dF(s)nb$$

$$= (1+o(1))\frac{1}{n}\sum_{i=1}^{n}\int_{-\infty}^{t}\left[X_i^2\overline{G}_i(s) - \frac{\Gamma_{n,1}(s)}{\Gamma_{n,0}(s)}X_i^T\overline{G}_i(s)\right]\frac{\lambda'(s)}{\lambda(s)}\,dF(s)nb.$$

Therefore (2.27) holds. □

PROOF OF THEOREM 1. From (2.11) and (2.12),

$$(2.30)\qquad \sup_{t\in R^1,\,0<\|b\|\le B}\left\{\|\xi(b,t)-\xi(0,t)-\zeta(b,t)\|/(\sqrt{n}+n\|b\|)\right\}\to 0,\quad \text{as } n\to\infty.$$

Combining Lemma 5 with (2.30) we get Theorem 1. □

PROOF OF COROLLARY 1. From Lemma 5 and the assumption that all eigenvalues of A_n are bounded away from zero we conclude that there exists such a neighborhood $\mathscr{N}$ that $\liminf_{n\to\infty}\inf_{b\in\mathscr{N},\,\|b-\beta\|\ge\theta}\{n^{-1}\|\eta(b)\|\} > 0$ for every $\theta > 0$. From this and Lemma 3(i), conclusion (i) follows.

For (ii), since $\hat{\beta}$ is consistent, Theorem 1 implies that

$$\sqrt{n}\,\Sigma_n^{-1/2}A_n(\hat{\beta}-\beta) = (n\Sigma_n)^{-1/2}\xi(\beta) + o_p(1),$$

which is also asymptotically equivalent to $(n\Sigma_n)^{-1/2}\xi(\beta, t_\beta(\alpha))$ for any $\alpha \in (0,1)$. Choosing $\alpha < 1/6$, we can show that the latter is asymptotically equivalent to a sum of independent zero-mean random vectors. Thus the desired asymptotic normality follows from the classical Lindeberg central limit theorem. □

3. Extensions to the weighted log-rank estimating equations. In this section, we extend the results of Section 2 to the weighted log-rank estimating functions of the form

$$(3.1)\qquad \xi_\psi(b) = \int_{-\infty}^{\infty}\psi(b,t)\sum_{i=1}^{n}\left(X_i - \frac{Z^x(b,t)}{Z(b,t)}\right)dN_i(b,t),$$

where $\psi(b,t) = \psi_n(b,t)$ is left continuous and $\psi(\beta,t) \in \sigma\{Z_i(\beta,s), N_i(\beta,s);\ s\le t,\ i=1,\dots,n\}$. We shall impose the following boundedness and continuity conditions, which are satisfied by the commonly used linear rank statistics in survival analysis as will be commented.

CONDITION 5(a). For every $B > 0$, $\limsup_{n\to\infty}\sup_{\|b\|\le B}[|\psi(b,0)| + \int_{-\infty}^{\infty}|d\psi(b,t)|] < \infty$ a.s.

CONDITION 5(b). There exist $\alpha_0 > 0$ and $\theta_0 > 0$ such that

$$\sup_{t \le t(\alpha_0),\, \|b-\beta\| \le n^{-1/3}} |\psi(b,t) - \psi(\beta,t) - \mu_\psi^T(t)(b-\beta)| = o(n^{-1/2}) \quad \text{a.s.},$$

where $t(\alpha) = \inf\{t\colon EZ(\beta,t) \le n^{1-\alpha}\}$ and $\mu_\psi(t)$ is a $p \times 1$ vector satisfying $\int_{-\infty}^{t(\alpha_0)} |d\mu_\psi(t)| = o(n^{1/3-\theta_0})$ a.s. for some $\theta_0 > 0$.

CONDITION 5(c). There exists $\bar{\psi}_n$ such that for every $\theta > 0$ and every t,

$$\sup_{\|b-\beta\| \le d_n} \left\{ |\psi(b,t) - \bar{\psi}_n(t)| I_{\{EZ(\beta,t) \ge \theta\}} \right\} \to 0.$$

It can be verified that Conditions 5(a)–(c) are satisfied by $\psi(b,t) = Z(b,t)/n$ (Gehan) and $\psi(b,t) = \hat{S}^\rho(b,t)$, $\rho \ge 0(G^\rho)$, where

$$\hat{S}(b,t) = \prod_{s<t} \left(1 - \frac{\Delta N(b,s)}{Z(b,s)}\right) \tag{3.2}$$

is the Kaplan–Meier estimator of the survival function $\bar{F}(t)$ or more generally by $\psi(b,t) = \phi(\hat{S}_n(b,t))$ with $|\phi'(u)| + |\phi''(u)| \le u^{-k_0}$ for some $k_0 > 0$ and all $u \in (0,1)$. In fact, for $\psi(b,t) = Z(b,t)/n$ we have $\bar{\psi}_n(t) = EZ(\beta,t)/n$ and

$$\mu_{n,\psi}(t) = -\left[f(t)\Gamma_n(t) + \bar{F}(t) n^{-1} \sum_{i=1}^{n} g_i(t + \beta^T X_i) X_i \right],$$

while for $\psi(b,t) = \phi(\hat{S}_n(b,t))$, $\bar{\psi}_n(t) = \phi(\bar{F}(t))$ and

$$\mu_{n,\psi}(t) = -\phi'(\bar{F}(t))\bar{F}(t) \int_{-\infty}^{t} \Gamma_{n,1}(s) \frac{\lambda'(s)}{\lambda(s)} \frac{dF(s)}{\bar{F}(s)}.$$

Recalling the definition of $\xi(b,t)$, we can write $\xi_\psi(b) = \int_{-\infty}^{\infty} \psi(b,t)\, d\xi(b,t)$. Since Section 2 gives the asymptotic linearity of the $\xi(b,t)$, it is intuitively clear, via the integration by parts formula, that $\xi_\psi(b)$ should also be asymptotically linear. Theorem 2 below confirms that this is indeed the case.

THEOREM 2. *Suppose that Conditions* 1–5 *are satisfied. Then the weighted log-rank estimating function $\xi_\psi(b)$ is asymptotically linear in the sense that, for every sequence $d_n > 0$ with $d_n \to 0$ a.s.,*

$$\begin{aligned} &\sup_{\|b-\beta\| \le d_n} \left\{ \|\xi_\psi(b) - \xi_\psi(\beta) - A_{n,\psi} n(b-\beta)\| / \left(\sqrt{n} + n\|b-\beta\|\right) \right\} \\ &\quad = o(1) \quad a.s., \end{aligned} \tag{3.3}$$

where $A_{n,\psi} = \int_{-\infty}^{\infty} \bar{\psi}_n(t)\, dA_n(t)$.

From Theorem 2 and similar to Corollary 1, we have the following corollary.

COROLLARY 2. *Suppose that Conditions* 1–5 *are satisfied and that all the eigenvalues of $A_{n,\psi}$ are bounded away from zero for all large n.*

(i) *There exists a closed neighborhood* $\mathscr{N}$, *containing* β *as an interior point such that* $\hat{\beta}_\psi$, *defined as a solution of* $\|\xi_\psi(\hat{\beta}_\psi)\| = \min_{b \in \mathscr{N}} \|\xi_\psi(b)\|$, *is strongly consistent.*

(ii) *For any* $\mathscr{N}$ *containing* β *as its interior point defining* $\hat{\beta}_\psi$ *as in* (i), *if* $\hat{\beta}_\psi$ *is consistent, then*

$$\sqrt{n}\,\Sigma_{n,\psi}^{-1/2} A_{n,\psi}\big(\hat{\beta}_\psi - \beta\big) \to_{\mathscr{D}} N(0, I_p),$$

where $\Sigma_{n,\psi} = \int_{-\infty}^{\infty} \bar{\psi}_n(t)[\Gamma_{n,2}(t) - \Gamma_{n,1}(t)\Gamma_{n,1}^T(t)/\Gamma_{n,0}(t)]\, dF(t)$. *In particular, if stability condition* $\Gamma_{n,k}(t) \to \Gamma_k(t)$ *holds for* $k = 0, 1, 2$ *and all* t, *then*

$$A_{n,\psi} \to A_\psi = \int_{-\infty}^{\infty} \bar{\psi}(t)\left[\Gamma_2(t) - \frac{\Gamma_1(t)\Gamma_1^T(t)}{\Gamma_0(t)}\right]\frac{\lambda'(t)}{\lambda(t)}\, dF(t), \tag{3.4}$$

$$\Sigma_{n,\psi} \to \Sigma_\psi = \int_{\infty}^{\infty} \bar{\psi}(t)\left[\Gamma_2(t) - \frac{\Gamma_1(t)\Gamma_1^T(t)}{\Gamma_0(t)}\right] dF(t), \tag{3.5}$$

and

$$\sqrt{n}\,\big(\hat{\beta}_\psi - \beta\big) \to_{\mathscr{D}} N\big(0, A_\psi^{-1}\Sigma_\psi A_\psi^{-1}\big).$$

Corollary 2 is easily proved by using the same argument as in the proofs of Corollary 1 and Theorem 2, and therefore its proof is omitted.

PROOF OF THEOREM 2. For notational simplicity, assume again that $\beta = 0$. Let $\xi(b, t)$ and $\zeta(b, t)$ be defined as in Lemma 3. Then

$$\begin{aligned} \xi_\psi(b) &= \int_{-\infty}^{\infty} \psi(b,t)\, d[\xi(b,t) - \xi(0,t) - A_n(t)nb] \\ &\quad + \int_{-\infty}^{\infty} \psi(b,t)\, d\xi(0,t) + \int_{-\infty}^{\infty} \psi(b,t)\, dA_n(t)nb. \end{aligned} \tag{3.6}$$

From (2.30) and Lemma 5, we have

$$\sup_{t \in R^1,\, \|b\| \le d_n} \big\{\|\xi(b,t) - \xi(0,t) - A_n(t)nb\|/(\sqrt{n} + n\|b\|)\big\} \to 0. \tag{3.7}$$

From (3.7), Condition 5(a) and the integration by parts formula,

$$\sup_{\|b\| \le d_n} \left\{\left\|\int_{-\infty}^{\infty} \psi(b,t)\, d[\xi(b,t) - \xi(0,t) - A_n(t)nb]\right\| \Big/ (\sqrt{n} + n\|b\|)\right\} \to 0 \quad \text{a.s.} \tag{3.8}$$

From Condition 5(c) and the fact that $\sup_n \int I_{\{EZ(0,t) \le \theta\}} |dA_n(t)| \to 0$, as $\theta \to 0$,

$$\sup_{\|b\| \le d_n} \left\|\int_{-\infty}^{\infty} \psi(b,t)\, dA_n(t) - \int_{-\infty}^{\infty} \bar{\psi}(t)\, dA_n(t)\right\| \to 0. \tag{3.9}$$

In view of (3.6), (3.8) and (3.9), it remains to show

$$(3.10)\quad \sup_{\|b\|\le d_n}\left\{\left\|\int_{-\infty}^{\infty}\psi(b,t)\,d\xi(0,t)-\int_{-\infty}^{\infty}\psi(0,t)\,d\xi(0,t)\right\|\Big/(\sqrt{n}+n\|b\|)\right\}\to 0\quad \text{a.s.}$$

From (2.11), $\sup_{t\in R^1}\|\xi(0,t)\|=o(n^{1/2+\theta})$ a.s. Thus, (3.10) holds with $\|b\|\le d_n$ replaced by $d_n\ge\|b\|\ge n^{-1/3}$. Now Lemma 2 implies that

$$\sup_{t\ge t(\alpha_0)}\|\xi(0,t)-\xi(0,t(\alpha_0))\|=o(n^{1/2-\varepsilon_0})\quad \text{a.s.},$$

for some $\varepsilon_0>0$. Thus in view of Condition 5(b),

$$(3.11)\quad \begin{aligned}&\sup_{\|b\|\le n^{-1/3}}\left\|\int_{-\infty}^{\infty}\psi(b,t)\,d\xi(0,t)-\int_{-\infty}^{\infty}\psi(0,t)\,d\xi(0,t)\right\|\\&\quad=o(n^{1/2})+\sup_{\|b\|\le n^{-1/3}}\left\|\int_{-\infty}^{t(\alpha_0)}\mu_{\psi}^{T}(t)b\,d\xi(0,t)\right\|\\&\quad=o(n^{1/2})\quad\text{a.s.}\end{aligned}$$

Therefore (3.10) holds and Theorem 2 follows. □

4. Approximations of the estimating functions in compact regions and consistency and asymptotic normality of the rank estimators. In the preceding two sections, asymptotic linearity properties have been established for the log-rank and the weighted log-rank estimating functions $\xi(b)$ and $\xi_{\psi}(b)$ for b in some neighborhood close to the true parameter β. These properties are then used to show that inside that neighborhood of β, the corresponding rank estimators are consistent and asymptotically normal. Unless a consistent auxiliary estimator can be obtained a priori, such small region is unknown in practice. Therefore, it is important in both theory and practice to know when a rank estimator defined in an arbitrarily chosen compact region will be consistent. In doing so, we shall in this section develop another kind of approximation. For an easy presentation, we shall assume that (C_i, X_i) are i.i.d. random vectors with H as the common marginal distribution of X_i and $g_x(\overline{G}_x)$ as the conditional density (survival) function of C_i given $X_i=x$. Conditional on (C_i, X_i), the ε_i are i.i.d. $\sim f$. Therefore, the observations (Y_i, X_i, δ_i) are i.i.d. random vectors.

THEOREM 3. *Suppose that Conditions* 1–4 *are satisfied. Define the "mean function"*

$$(4.1)\quad m(b,t)=E\left\{\int_{-\infty}^{t}\left[X_1-\frac{EX_1I_{\{Y_1-b^TX_1\ge s\}}}{EI_{\{Y_1-b^TX_1\ge s\}}}\right]dI_{\{Y_1-b^TX_1\le s,\,\delta_1=1\}}\right\}$$

and define $m(b)=m(b,\infty)$. *Then for all* $B>0$ *and* $\varepsilon>0$,

$$(4.2)\quad \sup_{\|b\|\le B}\|\xi(b)-nm(b)\|=o(n^{1/2+\varepsilon})\quad a.s.$$

The preceding theorem shows that within any bounded region, the estimating function $n^{-1}\xi(b)$ can be uniformly approximated by the nonrandom function $m(b)$ up to the order of $n^{-1/2+\varepsilon}$. To apply this theorem to rank estimation, suppose that there is a known compact region C_R containing β as an interior point. Let $\hat{\beta}$ be defined as a minimizer of $\|\xi(b)\|$, that is,

$$\|\xi(\hat{\beta})\| = \min_{b \in C_R} \|\xi(b)\|. \tag{4.3}$$

Then the following results concerning consistency and asymptotic normality of $\hat{\beta}$ hold.

THEOREM 4. *Suppose that Conditions* 1–4 *are satisfied.*

(i) *If* $m(b) \neq 0$ *for all* $b \in C_r \setminus \{\beta\}$, *then* $\hat{\beta} \to \beta$ *a.s.*

(ii) *In addition to the assumption* $m(b) \neq 0$, *suppose furthermore that* A *defined in Corollary* 1 *is nonsingular. Then*

$$\|\hat{\beta} - \beta\| = o(n^{-1/2+\varepsilon}) \quad a.s. \text{ for every } \varepsilon > 0, \tag{4.4}$$

$$\sqrt{n}\,(\hat{\beta} - \beta) \to_{\mathscr{D}} N(0, A^{-1}\Sigma A^{-1}), \tag{4.5}$$

Σ *is the same as in Corollary* 1.

Recall that $m(\beta) = 0$. Therefore, Theorem 4(i) shows that if the nonrandom mean function has a unique zero, then $\hat{\beta}$ is strongly consistent. Although $m(b)$ can be evaluated in principle for any given joint distribution of $(\varepsilon_i, C_i, X_i)$, we will verify the assumption $m(b) \neq 0$ for $b \neq \beta$ for two important cases in Section 5. Furthermore, if the slope of $m(\cdot)$ at β is nonsingular, which, from the comments following (2.2), is almost always the case, then $\hat{\beta}$ is also asymptotically normal.

PROOF OF THEOREM 3. We first note that the approximation (2.5a) in Lemma 1 remains valid if we regard X_i as random variables. More precisely, for every $0 \le \gamma < 1$, $B > 0$, $K > 0$ and $\theta > 0$, we have

$$\sup_{|b| \le B,\, EZ(b,t) \le Kn^{1-\gamma}} \|L(b,t) - EL(b,t)\| = o(n^{(1-\gamma)/2+\theta}) \quad \text{a.s.},$$

where L is any of the empirical processes Z, Z^x, N, N^x, $\tilde{N}$ or $\tilde{N}^x$. This clearly leads to Lemma 3(i) with $\zeta(b,t)$ replaced by $nm(b,t)$, that is,

$$\sup_{t \in R^1,\, |b| \le B} \|\xi(b,t) - nm(b,t)\| = o(n^{1/2+\varepsilon}) \quad \text{a.s.} \qquad \square \tag{4.6}$$

PROOF OF THEOREM 4. Since $m(b)$ is continuous in b and $m(\beta) = 0$, (i) follows from (4.2). Moreover, since $A_n \to A$, (ii) follows from (i) and Theorem 1. $\square$

From Theorems 3 and 4, and analogous to the argument given in the proof of Theorem 2, we can develop results similar to Theorems 3 and 4 for the weighted log-rank estimating functions and their induced estimators. More precisely, let $\xi_\psi(b)$ be defined by (3.1). Suppose that $\psi(b,t)$ converges to a nonrandom function $\bar{\psi}(b,t)$ in the sense that, for every $B > 0$ and some ε_0 and $\alpha_0 \in (0, 1/2)$ with $t_b(\alpha) = \inf\{t\colon EZ(b,t) \le n^{1-\alpha}\}$,

$$\sup_{\|b\| \le B,\, t \le t_b(\alpha_0)} |\psi(b,t) - \bar{\psi}(b,t)| = O(n^{-\varepsilon_0}) \quad \text{a.s.} \tag{4.7}$$

Define $\hat{\beta}_\psi$ as a minimizer of $\|\xi_\psi(b)\|$, that is, $\|\xi_\psi(\hat{\beta}_\psi)\| = \min_{\|b\| \in C_R} \|\xi_\psi(b)\|$. Then we have the following results.

THEOREM 5. *Suppose that Conditions* 1–5 *and* (4.7) *are satisfied.*

(i) $\sup_{\|b\| \le B} \|\xi_\psi(b) - n \int \bar{\psi}(b,t)\, dm(b,t)\| = O(n^{1-\min\{\varepsilon_0, \alpha_0\}})$ *a.s.*

(ii) *If* $m_\psi(b) = \int \bar{\psi}(b,t)\, dm(b,t) \ne 0$ *for all* $b \in C_R \setminus \{\beta\}$, *then* $\hat{\beta}_\psi \to \beta$ *a.s.*

(iii) *In addition to the assumption of* (ii) *above, suppose furthermore that* A_ψ *defined by* (3.4) $[\bar{\psi}(t) = \bar{\psi}(\beta, t)]$ *is nonsingular and that* $\bar{\psi}(b,t)$ *is continuous at* $b = \beta$ *for every* t. *Then*

$$\|\hat{\beta}_\psi - \beta\| = o(n^{-1/2+\varepsilon}) \quad a.s. \text{ for every } \varepsilon > 0, \tag{4.8}$$

$$\sqrt{n}\left(\hat{\beta}_\psi - \beta\right) \to_{\mathscr{D}} N\left(0, A_\psi^{-1} \Sigma_\psi A_\psi^{-1}\right), \tag{4.9}$$

where Σ_ψ *is defined by* (3.5).

It can be shown that the condition (4.7) is satisfied for the Gehan statistic and the G^ρ family. Moreover, for these statistics, the continuity of $\bar{\psi}(b,t)$ at $b = \beta$ is straightforward. The proof of Theorem 5 uses the integral representation $\xi_\psi(b) = \int \psi(b,t)\, d\xi(b,t)$ and approximations developed for $\xi(b,t)$. Instead of giving details, we only sketch some main ideas of the proof. Again, assume that $\beta = 0$.

PROOF OF THEOREM 5. For (i), we can first show that

$$\sup_{|b| \le B} \left\{ \left\| \int_{t_b(\alpha_0)}^{\infty} \psi(b,t)\, d\xi(b,t) \right\| + n \left\| \int_{t_b(\alpha_0)}^{\infty} \bar{\psi}(b,t)\, dm(b,t) \right\| \right\} \tag{4.10}$$
$$= O(n^{1-\alpha_0}) \quad \text{a.s.}$$

and then apply (4.6) and (4.7) to get

$$\int_{-\infty}^{t_b(\alpha_0)} \psi(b,t)\, d\xi(b,t) - n \int_{-\infty}^{t_b(\alpha_0)} \bar{\psi}(b,t)\, dm(b,t) = O(n^{1-\varepsilon_0}) \quad \text{a.s.} \tag{4.11}$$

Combining (4.10) and (4.11) we get (i). From (i) and the fact that $\int \bar{\psi}(0,t)\, dm(0,t) = 0$, we get (ii). Finally, (iii) follows from (ii) and Theorem 2. □

When no censoring is present, there are many popular methods, such as the least squares, for handling linear regression. These methods are usually superior in terms of computation and analysis. However, the rank method still has its own merits; compare Draper (1988). For the censored linear regression, the rank method becomes even more competitive since the computational and analytical advantages enjoyed by the least squares or similar methods have disappeared. For example, Buckley and James' (1979) extension of the least squares estimate for censored data is as difficult to analyze as (if not more than) the rank estimate.

By properly choosing the weight (score) function ψ, one can obtain an asymptotically efficient estimator $\hat{\beta}_\psi$ that attains the semiparametric (F is unspecified) lower bound. Adaptive construction of such asymptotically efficient estimate is given in Lai and Ying (1991), which also handles left-truncated data.

Recently, Lin and Geyer (1992) developed a useful algorithm for computing minimizers in some nonstandard settings by applying the method of simulated annealing. They have demonstrated that their algorithm is particularly useful for evaluating rank estimators because it does not require any smoothness condition on the "loss function."

5. Two special cases. In the preceding section, approximations are used to develop conditions that guarantee consistency and asymptotic normality of the rank estimators defined as minimizers of estimating functions over arbitrarily fixed bounded regions. A key condition in this development is the uniqueness of the zero for the nonrandom limit $\int \bar{\psi}(b,t)\,dm(b,t)$ of $n^{-1}\xi_\psi(b)$. Although given the underlying probability structure, one can always evaluate $\int \bar{\psi}(b,t)\,dm(b,t)$ numerically to check whether it is satisfied, we shall in this section demonstrate that this condition indeed holds in two important situations: (1) the two-sample problem and (2) the case when $\{\varepsilon_i\}$ has an increasing failure rate. Without mentioning specifically, we shall always assume throughout the rest that the regularity Conditions 1–5 as well as (4.7) are satisfied. Moreover, we shall assume that $\psi(b,t) \ge 0$. This is satisfied for all the commonly used statistics in survival analysis. In fact, it is satisfied when $\psi(b,t) = \phi(\hat{S}(b,t))$, where $\phi(u) = \gamma(u) - (1-u)^{-1}\int_u^1 \gamma(s)\,ds$ with $\gamma(u) = -\tilde{f}'(\tilde{F}^{-1}(u))/\tilde{f}(\tilde{F}^{-1}(u))$ for some strongly unimodal density $\tilde{f}$. We refer to Prentice (1978) for justification of generating $\psi(b,t)$ in such a way.

In the two-sample problem, we have $X_i = 0$, $i = 1,\dots,n_1$ and $X_i = 1$, $i = n_1+1,\dots,n$. It will be assumed that $\psi(b,t)$ is nonincreasing. Let

$$K_1(t) = \frac{1}{n}\sum_{i=1}^{n_1} \bar{G}_i(t), \qquad K_2(t) = \frac{1}{n}\sum_{i=n_1+1}^{n} \bar{G}_i(t). \tag{5.1}$$

Then

$$\begin{aligned} n^{-1}\zeta(b) &= \int_{-\infty}^{\infty} \bar{\psi}(b,t)\Bigg\{K_2(t+b)f(t+b-\beta) \\ &\qquad - \frac{K_2(t+b)\bar{F}(t+b-\beta)[K_1(t)f(t) + K_2(t+b)f(t+b-\beta)]}{K_1(t)\bar{F}(t) + K_2(t+b)\bar{F}(t+b-\beta)}\Bigg\}\,dt \\ &= \int_{-\infty}^{\infty} \bar{\psi}(b,t)\frac{K_1(t)\bar{F}(t)K_2(t+b)\bar{F}(t+b-\beta)}{K_1(t)\bar{F}(t) + K_2(t+b)\bar{F}(t+b-\beta)} \\ &\qquad \times[\lambda(t+b-\beta) - \lambda(t)]\,dt. \end{aligned} \tag{5.2}$$

Now $K_1(t)\bar{F}(t)K_2(t+b)\bar{F}(t+b-\beta)/[K_1(t)\bar{F}(t)+K_2(t+b)\bar{F}(t+b-\beta)]$ is nonincreasing in t and strictly decreasing at those points for which $K_1(t)K_2(t+b)>0$ and $f(t+b-\beta)+f(t)>0$. Therefore, for $b>\beta$,

$$\begin{aligned}&\int_{-\infty}^{\infty}\bar{\psi}(b,t)\frac{K_1(t)\bar{F}(t)K_2(t+b)\bar{F}(t+b-\beta)}{K_1(t)\bar{F}(t)+K_2(t+b)\bar{F}(t+b-\beta)}\lambda(t+b-\beta)\,dt\\ &\quad\geq\int_{-\infty}^{\infty}\bar{\psi}(b,t)\frac{K_1(t)\bar{F}(t)K_2(t+b)\bar{F}(t+b-\beta)}{K_1(t)\bar{F}(t)+K_2(t+b)\bar{F}(t+b-\beta)}\lambda(t)\,dt,\end{aligned}\tag{5.3}$$

with strict inequality if $K_1(t)K_2(t+b)\psi(b,t)>0$ for some t in the support of $f(t+b-\beta)+f(t)$. Moreover, $\liminf n^{-1}\zeta(b)>0$ provided that $K_1(t)K_2(t+b)\psi(b,t)$ is bounded away from 0 for all large n and for some t in the support of $f(t+b-\beta)+f(t)$. The other direction of the inequality holds when $b<\beta$. Thus $\hat{\beta}$ is consistent and asymptotically normal. This extends Louis (1981) for the log-rank estimating function and Wei and Gail (1983) for the Gehan case. In fact, either in the log-rank or in the Gehan, it is easy to see that the left-hand side of (5.3) is monotone in b. For general ψ, however, the limiting function is usually not monotone, but, as is shown, still has a unique zero.

Now consider the general regression model but with monotone increasing λ. For simplicity, assume that $(\varepsilon_i, X_i, C_i)$ are i.i.d. random vectors and that λ is strictly increasing. It should be emphasized that almost all the parametric families commonly used in modeling the accelerated life regression model have increasing failure rate (IFR). In particular, it can be verified that the extreme value, the log-gamma, the normal and the logistic families have strictly increasing IFR. The Cauchy family, however, does not have IFR property.

We first consider the one-dimensional case and then generalize it to the multiple regression. Following the notation of Section 4, we have

$$\begin{aligned}m_\psi(b)&=\int_{-\infty}^{\infty}\bar{\psi}(b,t)\int\bar{G}_x(t+bx)f(t+(b-\beta)x)\,dH(x)\\ &\qquad\times\left[\frac{\int x\bar{G}_x(t+bx)f(t+(b-\beta)x)\,dH(x)}{\int\bar{G}_x(t+bx)f(t+(b-\beta)x)\,dH(x)}\right.\\ &\qquad\left.-\frac{\int x\bar{G}_x(t+bx)\bar{F}(t+(b-\beta)x)\,dH(x)}{\int\bar{G}_x(t+bx)\bar{F}(t+(b-\beta)x)\,dH(x)}\right]dt.\end{aligned}\tag{5.4}$$

Let

$$q_{t,b}(x)=\bar{G}_x(t+bx)f(t+(b-\beta)x)\Big/\int\bar{G}_w(t+bw)f(t+(b-\beta)w)\,dH(w),$$

$$r_{t,b}(x)=\bar{G}_x(t+bx)\bar{F}(t+(b-\beta)x)\Big/\int\bar{G}_w(t+bw)\bar{F}(t+(b-\beta)w)\,dH(w).$$

Then $q_{t,b}$ and $r_{t,b}$ are both density functions with respect to the common measure dH. Moreover, $q_{t,b}/r_{t,b}$ is strictly increasing (decreasing) in x when $b>\beta$ $(b<\beta)$. It follows then that for $b>(<)\beta$, $\int xq_{t,b}(x)\,dH(x)>(<)\int xr_{t,b}(x)\,dH(x)$. Hence $m_\psi(b)>(<)0$ if $b>(<)\beta$.

We now extend the above result to the multiple regression model. To show $m_\psi(b) \neq 0$ for $b \neq \beta$, it suffices to show $(b-\beta)^T m_\psi(b) \neq 0$. From (5.4) with $x \in R^p$ and letting $x_b = b^T x$ and $x_{b-\beta} = (b-\beta)^T x$,

$$(5.5)\quad \begin{aligned}(b-\beta)^T m_\psi(b) = \int_{-\infty}^{\infty} \bar{\psi}(b,t) \int \overline{G}_x(t+x_b) f(t+x_{b-\beta})\, dH(x) \\ \times \Bigg[\frac{\int x_{b-\beta}\overline{G}_x(t+x_b) f(t+x_{b-\beta})\, dH(x)}{\int \overline{G}_x(t+x_b) f(t+x_{b-\beta})\, dH(x)} \\ - \frac{\int x_{b-\beta}\overline{G}_x(t+x_b)\overline{F}(t+x_{b-\beta})\, dH(x)}{\int \overline{G}_x(t+x_b)\overline{F}(t+x_{b-\beta})\, dH(x)} \Bigg] dt.\end{aligned}$$

Let $\Delta_1 = b - \beta$ and $\Delta_2, \dots, \Delta_p$ form an orthogonal basis for R^p. Define linear transformation $\tilde{x} = Qx$ where $Q^T = (\Delta_1, \dots, \Delta_p)$. Replacing x inside the brackets [ ] in (5.5) by $Q^{-1}\tilde{x}$ and integrating out $\tilde{x}_2, \dots, \tilde{x}_p$, we get, for some nondecreasing function $\tilde{H}$,

$$(5.6)\quad \begin{aligned}(b-\beta)^T m_\psi(b) = \int_{-\infty}^{\infty} \bar{\psi}(b,t) \int \overline{G}_x(t+x_b) f(t+x_{b-\beta})\, dH(x) \\ \times \left[\frac{\int \tilde{x}_1 f(t+\tilde{x}_1)\, d\tilde{H}(\tilde{x}_1)}{\int f(t+\tilde{x}_1)\, d\tilde{H}(\tilde{x}_1)} - \frac{\int \tilde{x}_1 \overline{F}(t+\tilde{x}_1)\, d\tilde{H}(\tilde{x}_1)}{\int \overline{F}(t+\tilde{x}_1)\, d\tilde{H}(\tilde{x}_1)} \right] dt,\end{aligned}$$

reducing the problem to the previous one-dimensional situation. Hence $(b-\beta)^T m_\psi(b) > 0$ for all $b \neq \beta$.

REFERENCES

ADICHIE, J. N. (1967). Estimates of regression parameters based on rank tests. *Ann. Math. Statist.* **38** 894–904.

ALEXANDER, K. (1984). Probability inequalities for empirical processes and a law of the iterated logarithm. *Ann. Probab.* **12** 1041–1067. [Correction: *Ann. Probab.* **15** 428–430.]

BENNETT, G. (1962). Probability inequalities for sums of independent random variables. *J. Amer. Statist. Assoc.* **57** 33–45.

BUCKLEY, J. and JAMES, I. (1979). Linear regression with censored data. *Biometrika* **66** 429–436.

COX, D. R. and OAKES, D. (1984). *Analysis of Survival Data.* Chapman and Hall, London.

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

DRAPER, D. (1988). Rank-based robust analysis of linear models. I. Exposition and review (with discussion). *Statist. Sci.* **3** 239–271.

EFRON, B. and JOHNSTONE, I. M. (1990). Fisher's information in terms of the hazard rate. *Ann. Statist.* **18** 38–62.

GEHAN, E. A. (1965). A generalized Wilcoxon test for comparing arbitrarily singly-censored samples. *Biometrika* **52** 203–223.

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

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

HARRINGTON, D. P. and FLEMING, T. R. (1982). A class of rank test procedures for censored survival data. *Biometrika* **69** 553–566.

HETTMANSPERGER, T. P. (1984). *Statistical Inference Based on Ranks.* Wiley, New York.

HODGES, L. J., JR. and LEHMANN, E. L. (1963). Estimates of location based on rank tests. *Ann. Math. Statist.* **34** 598–611.

JUREČKOVÁ, J. (1971). Nonparametric estimate of regression coefficients. *Ann. Math. Statist.* **42** 1328–1338.

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

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. (1991). Rank regression methods for left truncated and right censored data. *Ann. Statist.* **19** 531–554.

LAI, T. L. and YING, Z. (1992). Linear rank statistics in regression analysis with censored or truncated data. *J. Multivariate Anal.* **40** 13–45.

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

LIN, D. Y. and GEYER, C. J. (1992). Computational methods for semiparametric linear regression with censored data. *Journal of Computational and Graphical Statistics* **1** 77–90.

LOUIS, T. A. (1981). Nonparametric analysis of an accelerated failure time model. *Biometrika* **68** 381–390.

MILLER, R. G. and HALPERN, J. (1982). Regression with censored data. *Biometrika* **69** 521–531.

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

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

RITOV, Y. and WELLNER, J. A. (1988). Censoring, martingales and the Cox model. In *Proc. Conf. Statist. Inference for Stochastic Processes, Contemporary Mathematics.* Amer. Math. Soc., Providence, RI.

SHORACK, G. R. and WELLNER, J. A. (1986). *Empirical Processes with Applications to Statistics.* Wiley, New York.

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

WEI, L. J. and GAIL, M. H. (1983). Nonparametric estimation for a scale-change with censored observations. *J. Amer. Statist. Assoc.* **78** 382–288.

WEI, L. J., YING, Z. and LIN, D-Y. (1990). Linear regression analysis of censored survival data based on rank tests. *Biometrika* **77** 845–851.

DEPARTMENT OF STATISTICS
UNIVERSITY OF ILLINOIS
101 ILLINI HALL
725 SOUTH WRIGHT STREET
CHAMPAIGN, ILLINOIS 61820