

LINEAR REGRESSION WITH INTERVAL CENSORED DATA

BY GANG LI AND CUN-HUI ZHANG¹

ACER/Excel Inc. and Rutgers University

This paper concerns linear regression with interval censored data. M -estimators for the regression coefficients are derived. Asymptotic consistency and normality of the M -estimators are obtained via an exponential inequality for U -statistics. Asymptotically efficient estimators are provided under mild conditions.

1. Introduction. Consider the linear model

$$(1.1) \quad Y_i = \beta^t Z_i + \varepsilon_i, \quad 1 \leq i \leq n,$$

where Z_i are known covariate vectors and β is an unknown parameter vector. We are interested in estimating β from the current status data such that

$$(1.2) \quad X_i = (\delta_i, T_i, Z_i), \quad 1 \leq i \leq n,$$

are observed instead of the usual (Y_i, Z_i) , where $\delta_i = I_{\{Y_i \leq T_i\}}$ for some censoring times T_i . This observation scheme is also called the binary choice model [Coslett (1987)] and Case 1 interval censoring [Groeneboom and Wellner (1992)]. We assume that the errors $\{\varepsilon_i\}$ are independent of $\{Z_i, T_i\}$ and are i.i.d. variables with a common distribution $F(t)$.

Situations involving interval censoring arise commonly in engineering and medicine. In some clinical settings (e.g., final follow-up after treatment), an examination at time T_i determines whether or not an endpoint Y_i has occurred. In reliability studies, destructive tests are often used to find whether an item (e.g., fire extinguisher) has failed. In rodent bioassay experiments, the presence or absence of a tumor may only be detected through sacrifices. For certain designs of experiments, T_i are i.i.d. random variables with a known probability density. In general, T_i and Z_i are associated. In Case 2 interval censoring, two censoring variables $T'_i < T''_i$ are observed with the knowledge that Y_i is inside the interval $[T'_i, T''_i]$. In both cases of interval censoring and with known β ($= 0$), the estimation of the common distribution F of ε_i in (1.1) has been considered by many authors under parametric and nonparametric assumptions; see, for example, Ayer, Brunk, Ewing, Reid and Silverman (1955), Brunk (1970), Groeneboom and Wellner (1992) and Peto et al. (1980) among others. Linear and hazards regression models for interval censored data have been considered by Finkelstein (1986), Finkel-

Received September 1994; revised March 1998.

¹Supported in part by the National Science Foundation and the National Security Agency.

AMS 1991 subject classifications. Primary 62G20, 62G05; secondary 60F05.

Key words and phrases. Linear regression, censoring, information, efficient estimation, U -statistics.

stein and Wolfe (1986), Huang (1996), Huang and Wellner (1996), Jewell and Shiboski (1990), Lin and Ying (1996), Rabinowitz, Tsiatis and Aragon (1995), and Shiboski and Jewell (1992). The binary choice model has been considered by Cosslett (1987), Han (1987), Klein and Spady (1993), Manski (1975, 1985) and Sherman (1993) in the econometrics literature. In particular, Klein and Spady (1993) considered an estimated likelihood function and obtained asymptotic efficiency for their estimators under certain regularity conditions.

In this paper we provide some simple estimators for β as well as some asymptotically efficient ones under (1.1) and (1.2). Recent results on U -statistics [Arcones and Giné (1993)] are used to obtain expansions of estimating functions. Consequently, the asymptotic normality and efficiency of our estimators are obtained under much weaker conditions compared with previous results (cf. Section 2.4).

2. Estimation of β . We shall provide our estimators and their asymptotic properties in this section. Three types of estimators are considered: (1) simple M -estimators, (2) estimators with efficient asymptotic variance at a given error distribution $F = F_0$ and (3) asymptotically efficient estimators. The asymptotic expansions of estimating functions and proofs of the main results are given in Sections 3 and 4.

2.1. *A simple method.* Let $X = (\delta, T, Z)$ be a random vector related to (ε, T, Z) as in (1.1) and (1.2). Set $T_b = T - b^{\text{tr}}Z$. Since $E[\delta|T, Z] = F(T_\beta) = F(T_b + (b - \beta)^{\text{tr}}Z)$ and F is an increasing function,

$$(2.1) \quad (b - \beta)^{\text{tr}} \text{Cov}(\delta, Z|T_b) = \text{Cov}(F(T_b + (b - \beta)^{\text{tr}}Z), (b - \beta)^{\text{tr}}Z|T_b) \geq 0 \quad \text{a.s.}$$

for all b . This fact motivates the estimating equation $\xi_n(\hat{\beta}_n) = 0$ with the U -statistics

$$(2.2) \quad \xi_n(b) = \frac{1}{n(n-1)(n-2)} \times \sum_{(i,j,k) \in I_3^n} \psi(T_{i,b})(\delta_i - \delta_j)(Z_i - Z_k)K_{nij}(b)K_{nik}(b),$$

where $T_{i,b} = T_i - b^{\text{tr}}Z_i$, $\psi(\cdot)$ is a weight function and $K_{nij}(b) = a_n^{-1}K((T_{i,b} - T_{j,b})/a_n)$ with bandwidth $a_n \rightarrow 0$. In this paper, $K(t)$ is always a twice continuously differentiable density function with bounded support, and $I_i^n = \{\mathbf{i} = (i_1, \dots, i_l): i_j \leq n, i_j \neq i_{j'} \text{ for } j \neq j'\}$.

Let us take a brief look at $E\xi_n(b)$ to explain the use of (2.2) based on (2.1). Suppose iid observations (1.2) are taken and the conditional density $P\{T_i \in dt|Z_i = z\} = g(t|z) dt$ exists. Let $g(t; b) = Eg(t + b^{\text{tr}}Z|Z)$ be the density of T_b . Under certain smoothness conditions on $g(\cdot|z)$, $E[h(\delta_j, Z_j)K_{nij}(b)|X_i] \approx E[h(\delta_i, Z_i)|T_{i,b}]g(T_{i,b}; b)$ as $n \rightarrow \infty$ (i.e., $a_n \rightarrow 0$) for all bounded Borel functions h . Since $\xi_n(b)$ are U -statistics, this and the SLLN imply

$$(2.3) \quad \xi_n(b) \approx E\xi_n(b) \approx \xi_\infty(b), \quad \xi_\infty(b) = E\psi(T_b)g^2(T_b; b)\text{Cov}(\delta, Z|T_b).$$

By (2.1) $b = \beta$ is the only solution of $\xi_\infty(b) = 0$, provided that $\psi(t) > 0$ for all t and the inequality in (2.1) is strict with positive probability for $b \neq \beta$. Thus, the solution of $\xi_n(b) = 0$ should be close to the regression coefficient β .

Asymptotic properties of $\xi_n(b)$ are relatively easy to find. Recently, Arcones and Giné (1993) extended the Bernstein inequality to completely degenerate U -statistics. These inequalities are used in Section 3, with the Hoeffding decomposition, to obtain expansions of the estimating functions $\xi_n(b)$ and their derivatives $(\partial/\partial b)\xi_n(b)$, which in turn imply the asymptotic normality of $\hat{\beta}_n$.

The solution $\xi_n(b) = 0$ is in general not unique. For definiteness and to rule out unbounded solution sequences, we define our estimator $\hat{\beta}_n$ by

$$(2.4) \quad \xi_n(\hat{\beta}_n) = 0, \quad \|\hat{\beta}_n - b_n\| = \min\{\|b - b_n\|: \xi_n(b) = 0\},$$

where $b_n = O_p(1)$ are (possibly inconsistent; e.g., $b_n = 0$) initial estimates of β . Under mild conditions on $\{\alpha_n\}$, it will be shown in Theorem 2.1 below that

$$(2.5) \quad P\{\hat{\beta}_n \text{ is the unique solution of } \xi_n(b) = 0 \text{ for } b \in B_M\} \rightarrow 1$$

for all $M > 0$, where $B_M = \{b: \|b - \beta\| \leq M\}$. Thus, $\hat{\beta}_n$ do not depend on the choice of b_n asymptotically.

Suppose for some finite M_0 and $0 < \alpha \leq 1$,

$$(2.6) \quad \|Z\| \leq M_0, \quad g(t|z) \in \text{Lip}(M_0, \alpha) \text{ in } t.$$

A function $f(u, v)$ belongs to the Lipschitz class $\text{Lip}(M, \alpha)$ in u if $|f(u, v)| \leq M$ and $|f(u, v) - f(u', v)| \leq M\|u - u'\|^\alpha$ for all (u, u', v) , $0 \leq \alpha \leq 1$. Define

$$(2.7) \quad \begin{aligned} \Lambda_\psi &= \int w_\psi(t)v(t; \beta)g(t; \beta)F(dt), \\ \Sigma_\psi &= Ew_\psi^2(T_\beta)F(T_\beta)(1 - F(T_\beta))v(T_\beta; \beta), \end{aligned}$$

where $w_\psi(t) = \psi(t)g^2(t; \beta)$ and $v(t; b) = \text{Cov}(Z|T_b = t)$.

THEOREM 2.1. *Suppose (2.6) holds, Λ_ψ is of full rank, and $(b - \beta)^{\text{tr}} \xi_\infty(b) \neq 0$ for all $b \neq \beta$ with the $\xi_\infty(b)$ in (2.3). Let $\hat{\beta}_n$ be given by (2.4) with the $\xi_n(b)$ in (2.2) such that $\|\psi\|_\infty < \infty$ and $\psi' = (d\psi/dt) \in \text{Lip}(M_0, \alpha_0)$ for some $\alpha_0 > 0$. Choose a_n such that $a_n(n/\log n)^{2/3} \rightarrow \infty$, $a_n^{1-\alpha}\sqrt{n/\log n} \rightarrow \infty$ and $a_n^{1+\alpha}\sqrt{n} \rightarrow 0$. Then, $\hat{\beta}_n$ is asymptotically unique in the sense of (2.5) and*

$$(2.8) \quad \begin{aligned} \sqrt{n}(\hat{\beta}_n - \beta) &= -\Lambda_\psi^{-1} \left\{ n^{-1/2} \sum_{i=1}^n \rho_\psi(X_i; \beta) \right\} + o_p(1) \\ &\rightarrow_{\mathcal{D}} N(0, \Lambda_\psi^{-1} \Sigma_\psi \Lambda_\psi^{-1}), \end{aligned}$$

where $\rho_\psi(X; b) = \psi(T_b)(\delta - F(T_b))(Z - E[Z|T_b])g^2(T_b; b)$.

Theorem 2.1 is proved in Section 3. For small sample size, one may choose $\psi(t) = 1$ so that (2.2) is location invariant. The conditions on the bandwidth

α_n hold with $\alpha_n = \tau_0/\sqrt{n}$, $0 < \tau_0 < \infty$, for $0 < \alpha \leq 1$. By (2.1), (2.3) and (2.7), Λ_ψ is of full rank and $(b - \beta)^{\text{tr}} \xi_{\infty}(b) > 0$ for all $b \neq \beta$ if $\psi(t) > 0 \forall t$, F is strictly increasing and $E \text{Var}(\eta^{\text{tr}} Z|T_b) > 0 \forall \eta, b$.

2.2. *Efficient estimation at F_0 .* Suppose $F(t)$ has a bounded continuous density $f(t) = F'(t)$. Cosslett (1987) showed that the (semiparametric) minimum Fisher information for the estimation of β with unknown F is

$$(2.9) \quad I_* = E \text{Cov}(\rho_*(X; \beta)) = E \text{Cov}(Z|T_\beta) f(T_\beta) \psi_*(T_\beta) g^2(T_\beta; \beta),$$

where $\psi_*(t) = f(t)/[F(t)\{1 - F(t)\}g^2(t; \beta)]$, and that the efficient score is $\rho_*(X; b) = \rho_{\psi_*}(X; b)$ via (2.8). See also Bickel, Klaassen, Ritov and Wellner (1993). If $\psi \approx \psi_*$ in (2.2), then $\Lambda_\psi^{-1} \Sigma_\psi \Lambda_\psi^{-1} \approx I_*^{-1}$ and $\hat{\beta}_n$ in Theorem 2.1 is nearly efficient, as $I_* = \Lambda_{\psi_*} = \Sigma_{\psi_*}$.

Let F_0 be a distribution function with a density f_0 . By (2.9) the weight function $\psi_0(t) = f_0(t)/\{F_0(t)(1 - F_0(t))g^2(t; \beta)\}$ should be used when the error distribution F is known to be close to F_0 . Since $g(t; \beta)$ is unknown, we consider the estimating function

$$(2.10) \quad \bar{\xi}_n(b) = \frac{1}{\binom{n}{3} 3! \sum_{\mathbf{i} \in I_3^n} \frac{f_0(T_{i_1, b})(\delta_{i_1} - \delta_{i_2})(Z_{i_1} - Z_{i_3})K_{ni_1i_2}(b)K_{ni_1i_3}(b)}{c_n + F_0(T_{i_1, b})\{1 - F_0(T_{i_1, b})\}\hat{g}_{n, \mathbf{i}}^2(T_{i_1, b})},$$

with $\hat{g}_{n, \mathbf{i}}^2(T_{i_1, b}) = \sum_{\mathbf{j} \in I_2^n(\mathbf{i})} K_{ni_1j_1} K_{ni_1j_2} / \{\binom{n}{2} 2\}$ and $c_n > 0$, where $I_k^n(\mathbf{i}) = \{\mathbf{j} \in I_k^n: \{\mathbf{j}\} \cap \{\mathbf{i}\} = \emptyset\}$ with $\{\mathbf{i}\}$ being the set of the components of \mathbf{i} . Define

$$(2.11) \quad \bar{\xi}_n(\hat{\beta}_n) = 0, \quad \|\hat{\beta}_n - b_n\| = \min\{\|b - b_n\|: \bar{\xi}_n(b) = 0\},$$

where b_n are as in (2.4). Set $w_0 = f_0/\{F_0(1 - F_0)\}$ and for $v = v(t; \beta)$ and $g = g(t; \beta)$,

$$(2.12) \quad \Lambda_0 = \int w_0 v g dF, \quad \Sigma_0 = \int w_0^2 F(1 - F) v g dt, \\ \bar{\xi}_{\infty}(b) = E w_0(T_b) \text{Cov}(\delta, Z|T_b).$$

THEOREM 2.2. *Suppose (2.6) holds, $f'_0 \in \text{Lip}(\alpha_0, M_0)$ for some $\alpha_0 > 0$, Λ_0 and Σ_0 are finite and of full rank, and $(b - \beta)^{\text{tr}} \bar{\xi}_{\infty}(b) \neq 0$ for all $b \neq \beta$. Let $\hat{\beta}_n$ be given by (2.11) with the $\bar{\xi}_n(b)$ in (2.10). Choose a_n and $c_n \rightarrow 0$ such that $\log n / \{n a_n^2 c_n^2\} \rightarrow 0$ and $a_n^{1+\alpha} \sqrt{n} / c_n \rightarrow 0$. Then, (2.5) and (2.8) hold with $\Lambda_\psi = \Lambda_0$ and $\Sigma_\psi = \Sigma_0$ in (2.12) and $\rho_\psi = w_0(T_\beta)(\delta - F(T_\beta)) \times (Z - E[Z|T_\beta])$.*

The proof of Theorem 2.2 is omitted as it is similar to but simpler than that of Theorem 2.3 below. If $F = F_0$, then $\Lambda_0 = \Sigma_0 = I_*$. The conditions on a_n and c_n hold for $a_n = \tau_0(\log n)^2/\sqrt{n}$ and $c_n = \tau_2/\log n$ for all $0 < \alpha \leq 1$, although this choice is somewhat conservative. Theorem 2.2 requires smoothness conditions on the nominal distribution F_0 , but not on the true error distribution F .

2.3. *Efficient estimation.* Although the definition of I_* and the choice of ψ is not clear without smoothness assumptions on F , (2.9) suggests the use of a data-drive ψ when f is believed to exist and the sample size is adequate for an estimation of ψ_* .

We shall use the following notation throughout: for any Borel function \tilde{K} , let $\tilde{K}_n(t) = \tilde{K}(t/\tilde{a}_n)/\tilde{a}_n$ and $\tilde{K}_{nij} = \tilde{K}_{nij}(b) = \tilde{K}_n(T_{i,b} - T_{j,b})$, where the band-width \tilde{a}_n is specified when \tilde{K} is given. For example, $(\partial/\partial b)K_{nij} = (Z_j - Z_i)K'_{nij}/a_n$ for the kernel in (2.2) with $K' = (d/dt)K$. Also, $\mathbf{X}_{(i_1, \dots, i_l)} = (X_{i_1}, \dots, X_{i_l})$ and $\mathbf{X}_{i,j} = (X_i, X_{i+1}, \dots, X_j)$.

Among many methods of estimating the efficient weight function ψ_* , consider

$$\hat{F}_n(T_{i,b})\{1 - \hat{F}_n(T_{i,b})\}\hat{g}_n^2(T_{i,b}; b) = n^{-2} \sum_{j,k} \delta_j(1 - \delta_k)K_{nij}K_{nik}$$

and the local least squares estimator

$$\begin{aligned} \hat{f}_n(T_{i,b}; b) &= \frac{\sum_j \delta_j (T_{j,b} - \bar{T}_{i,b}) K_{nij}^{[0]}}{\sum_j (T_{j,b} - \bar{T}_{i,b})^2 K_{nij}^{[0]}} \\ (2.13) \quad &= \frac{\sum_{j,k} \delta_j (K_{nij}^{[1]} K_{nik}^{[0]} - K_{nij}^{[0]} K_{nik}^{[1]})}{a'_n \sum_{j,k} \{K_{nij}^{[2]} K_{nik}^{[0]} - K_{nij}^{[1]} K_{nik}^{[1]}\}}, \end{aligned}$$

where $\bar{T}_{i,b} = \sum_j T_{j,b} K_{nij}^{[0]}/\sum_j K_{nij}^{[0]}$, $K^{[l]}(t) = (-t)^l K^{[0]}(t)$, $K^{[0]}$ is a continuously differentiable density with bounded support and bandwidth a'_n . Replacing $\psi(T_{i,b})$ in (2.2) by a modification of $\hat{f}_n(T_{i,b}; b)/\{\hat{F}_n(T_{i,b})\{1 - \hat{F}_n(T_{i,b})\}\hat{g}_n^2(T_{i,b}; b)\}$, we arrive at

$$\begin{aligned} \hat{\xi}_n(b) &= \sum_{\mathbf{i} \in I_5^n} \frac{\tilde{h}_{n,0}(\mathbf{X}_i; b)/a'_n}{\binom{n}{5} 5! \{c_n + \hat{D}_{n,i}\}}, \\ (2.14) \quad \hat{D}_{n,i} = \hat{D}_{n,i}(b) &= \sum_{\mathbf{j} \in I_4^n(\mathbf{i})} \frac{\tilde{h}_n(\mathbf{X}_j, T_{i,b}; b)}{\binom{n-5}{4} 4!}, \end{aligned}$$

where

$$\begin{aligned} \tilde{h}_{n,0}(\mathbf{X}_{1,5}; b) &= (\delta_1 - \delta_2)(Z_1 - Z_3)K_{n12}K_{n13}\delta_4\{K_{n14}^{[1]}K_{n15}^{[0]} - K_{n14}^{[0]}K_{n15}^{[1]}\}, \\ (2.15) \quad & \end{aligned}$$

and

$$\tilde{h}_n(\mathbf{X}_{6,9}, T_{1,b}; b) = \delta_6(1 - \delta_7)K_{n16}K_{n17}\{K_{n18}^{[2]}K_{n19}^{[0]} - K_{n18}^{[1]}K_{n19}^{[1]}\}.$$

Note that terms with ties are removed, sums renormalized, and the denominator is stabilized by adding $0 \leq c_n \rightarrow 0$. The choice of $(K^{[0]}, a'_n) = (K, a_n)$ is possible. Define $\hat{\beta}_n$ by

$$\hat{\xi}_n(\hat{\beta}_n) = 0, \quad \|\hat{\beta}_n - \tilde{\beta}_n\| = \inf\{\|b - \tilde{\beta}_n\|: \hat{\xi}_n(b) = 0\},$$

where the initial estimates satisfy $\|\tilde{\beta}_n - \beta\| = O_p(n^{-1/2})$ (i.e., those in Theorem 2.1). For $K^{[0]}$ in (2.13), define $\tilde{K}(t) = \int_t^\infty (u - \mu_1)K^{[0]}(u) du$, $\tilde{K}' = (d/dt)\tilde{K} = K^{[1]} + \mu_1 K^{[0]}$,

$$\begin{aligned} \mu_1 &= \int uK^{[0]}(u) du, \\ (2.18) \quad \mu_2 &= \int (t - \mu_1)^2 K^{[0]}(t) dt = \int \tilde{K}(t) dt, \\ \tilde{\mu}_2 &= \int |\tilde{K}'(u)|^2 du = \int (t - \mu_1)^2 \{K^{[0]}(t)\}^2 dt. \end{aligned}$$

THEOREM 2.3. (i) *Let $0 < \gamma_2/\alpha < \gamma_1 \leq \gamma_0 \leq 1/3$ and $\gamma_2 < (1 + \alpha)\gamma_0 - 1/2$. Suppose*

$$(2.19) \quad \lim_{n \rightarrow \infty} \sup_{|u| \leq M} E \frac{\text{Cov}(Z|T_\beta)\{f(T_\beta + ua_n) - f(T_\beta)\}^2}{c_n + F(T_\beta)\{1 - F(T_\beta)\}} = 0 \quad \forall 0 < M < \infty,$$

(2.6) holds, I_* in (2.9) is finite and positive definite, and for some finite matrix σ^2 ,

$$(2.20) \quad \sigma_n^2 \rightarrow \sigma^2, \quad \sigma_n^2 = \frac{\tilde{\mu}_2/\mu_2^2}{n(\alpha'_n)^3} E \frac{v(T_\beta; \beta)D^2(T_\beta; \beta)}{g(T_\beta; \beta)\{c_n + D(T_\beta; \beta)\}^2},$$

where $D(t; \beta) = \mu_2 g^4(t; \beta)F(t)(1 - F(t))$. Let $\hat{\beta}_n$ be given by (2.17) with $\hat{\xi}_n(b)$ in (2.14) such that $n^{\gamma_0}a_n \rightarrow \tau_0$, $n^{\gamma_1}a'_n \rightarrow \tau_1$ and $n^{\gamma_2}c_n \rightarrow \tau_2$ for some $0 < \tau_j < \infty$. Then,

$$(2.21) \quad \sqrt{n}(\hat{\beta}_n - \beta) \rightarrow_{\mathcal{D}} N(0, I_*^{-1}(I_* + \sigma^2)I_*^{-1}).$$

Furthermore, the set $\{b: \hat{\xi}_n(b) = 0, \|b - \beta\| \leq M/\sqrt{n}\}$ contains a single point $\hat{\beta}_n$ with probability tending to 1 as $n \rightarrow \infty$ and then $M \rightarrow \infty$.

(ii) *If in addition $\sigma^2 = 0$, then $\hat{\beta}_n$ are asymptotically efficient, and with ρ_* as in (2.9),*

$$(2.22) \quad \begin{aligned} \sqrt{n}(\hat{\beta}_n - \beta) &= -I_*^{-1} \left\{ n^{-1/2} \sum_{i=1}^n \rho_*(X_i; \beta) \right\} + o_p(1) \\ &\rightarrow_{\mathcal{D}} N(0, I_*^{-1}). \end{aligned}$$

Theorem 2.3 is proved in Section 4. The last statement of Theorem 2.3(i) asserts that $\hat{\beta}_n$ in (2.17) does not depend on the choice of $\tilde{\beta}_n$ asymptotically, as long as $\tilde{\beta}_n$ is $n^{-1/2}$ -consistent.

2.4. Remarks. The (partial) likelihood function for the data $\mathbf{X}_{1,n}$ in (1.2) is

$$L_n(\beta, F) = \prod_{i=1}^n \{F(T_{i,\beta})\}^{\delta_i} \{1 - F(T_{i,\beta})\}^{1-\delta_i}.$$

Since the nonparametric maximum likelihood estimator (MLE) $\hat{F}_{n,\beta}$ of F for known β is well understood, one may consider the profile MLE of β which maximizes the profile likelihood $L_n(\beta, \hat{F}_{n,\beta})$. The profile MLE is asymptotically consistent, but its rate of convergence is unknown. Since $\hat{F}_{n,\beta}$ is given by local average with a data-driven bandwidth of order $n^{-1/3}$, the properties of the profile MLE should be similar to (2.17) with bandwidth $a_n = a'_n = (\tau_0 + o(1))n^{-1/3}$. By Theorem 2.3, (2.17) with such bandwidth is asymptotically normal but inefficient, since $\sigma^2 = \tilde{\mu}_2 \int v(t; \beta) I\{g(t; \beta) > 0\} dt / (\mu_2^2 \tau_0^3) \neq 0$. In general, a wider bandwidth a'_n than those of order $n^{-1/3}$ is needed for consistent estimation of $f = F'$, on which the efficient score ρ_* depends.

In Theorem 2.3, we may choose $1/6 \leq \gamma_1 < \gamma_0 = 1/3$ and $0 < \gamma_2 < (2\alpha - 1)/6$, or more generally (a_n, a'_n, c_n) satisfying (4.2) and (4.3) below. This allows the use of the bandwidth of optimal order for the estimation of F and f . For example, $\gamma_0 = \gamma_1 = 1/(2\alpha_0 + 3)$ for $f \in \text{Lip}(\alpha_0, M_0)$. Condition (2.20) holds with $\sigma^2 = 0$ if $n(a'_n)^3 c_n \rightarrow \infty$. The regularity conditions on $g(t|z)$ and F in Theorem 2.3(ii) are weaker than those of Klein and Spady (1993), who assumed the fourth continuous differentiability of $g(t|z)$, the boundedness of T_β , and $\varepsilon < F(T_\beta) < 1 - \varepsilon$ for some $\varepsilon > 0$. In Theorems 2.1 and 2.2 the only assumptions are (2.6) for any $\alpha > 0$, the uniqueness conditions and the finiteness of the matrices in the asymptotic variance.

As mentioned before, exponential inequalities for U -statistics are used in the proofs. The advantage of this approach is that the entropy for the estimating functions as U -statistics is of much smaller order than those of typical classes of estimated density, regression or weight functions in semi-parametric models.

3. Proof of Theorem 2.1. The basic idea is to treat (2.2) and $(\partial/\partial b)\xi_n^{\text{tr}}$ as U -statistics indexed by b and to obtain expressions in the following theorem, which leads to the local asymptotic linearity of $\xi_n(b)$ [(3.1), (3.2) and (3.5) below] and then the conclusions. The following notation is used in this and next sections: $X = X_1$ (e.g., $T_b = T_{1,b}$), $\tilde{M} = O(1)$, \tilde{M}_p is a random sequence with $(E\|\tilde{M}_p\|^p)^{1/p} = \tilde{M}$, \tilde{M}_∞ is a uniformly bounded sequence of random variables or functions, $h_b = h(b) = h(T_b; b)$ for all Borel functions $h(t; b)$, $g_{i,b}(t) = g(t + b^{\text{tr}} Z_i | Z_i)$, $F_{i,b}(t) = F(t + (b - \beta)^{\text{tr}} Z_i)$, the convolution is $(h * \tilde{K})(t) = \int h(t - u)\tilde{K}(u) du$ and $(h * \tilde{K})_b = (h * \tilde{K})(T_b)$.

3.1. *Asymptotic normality via expansions of $\xi_n(b)$ and $(\partial/\partial b)\xi_n^{\text{tr}}(b)$.*

THEOREM 3.1. *Suppose conditions of Theorem 2.1 hold. Then,*

$$(3.1) \quad \limsup_{n \rightarrow \infty} \Pr \left\{ \sup_{\|b - \beta\| \leq M} \left\| \frac{\partial}{\partial b} \xi_n^{\text{tr}}(b) - \frac{\partial}{\partial b} E \xi_n^{\text{tr}}(b) \right\| > \varepsilon \right\} = 0$$

$$\forall 0 < \varepsilon < M < \infty,$$

$$(3.2) \quad \lim_{\varepsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \sup_{\|b - \beta\| \leq \varepsilon a_n^{1-\alpha}} \left\| \frac{\partial}{\partial b} E \xi_n^{\text{tr}}(b) - \Lambda_\psi \right\| = 0,$$

$$(3.3) \quad \lim_{n \rightarrow \infty} \Pr \left\{ \left\| \xi_n(\beta) - \sum_{i=1}^n \rho_\psi(X_i; \beta) / n \right\| > \varepsilon \sqrt{n} \right\} = 0 \quad \forall \varepsilon > 0$$

and

$$(3.4) \quad \lim_{n \rightarrow \infty} \sup_b \frac{\|E \xi_n(b) - \xi_\infty(b)\|}{\|b - \beta\| + n^{-1/2}} = 0, \quad \lim_{b \rightarrow \beta} \frac{\|\xi_\infty(b) - \Lambda_\psi(b - \beta)\|}{\|b - \beta\|} = 0.$$

REMARK. It follows from Theorem 3.1 that

$$(3.5) \quad \xi_n(b) = \frac{1}{n} \sum_{i=1}^n \rho_\psi(X_i; \beta) + \Lambda_\psi(b - \beta) + \varepsilon_n(b)(\|b - \beta\| + n^{-1/2})$$

with $P\{\sup_{\|b - \beta\| \leq \varepsilon'} |\varepsilon_n(b)| > \varepsilon\} \rightarrow 0$ as $(n, \varepsilon') \rightarrow (\infty, 0+)$ for all $\varepsilon > 0$. This asymptotic linearity of the estimating functions is commonly used to derive asymptotic properties of the resulting estimators, but it does not imply (2.5). Theorem 3.1 provides the convergence of the derivative $(\partial/\partial b)\xi_n(b)$, which implies (2.5). The ball $\|b - \beta\| \leq Mn^{-1/2}$ is eventually inside the ball $\|b - \beta\| \leq \varepsilon a_n^{1-\alpha}$ for fixed $0 < \varepsilon < M < \infty$, as $a_n^{1-\alpha}\sqrt{n}/\log n \rightarrow \infty$. By our proofs, (3.5) remains valid if the conditions on $\{a_n\}$ and ψ in Theorem 2.1 are replaced by $a_n n^{3/4}/\log n \rightarrow \infty$, $a_n^{1+\alpha}\sqrt{n} \rightarrow 0$ and $\psi \in \text{Lip}(M_0, \alpha_0)$.

PROOF OF THEOREM 2.1. It follows from Theorem 3.1 and the positive definiteness of Λ_ψ that $\xi_n(\hat{\beta}_n) = 0$ for some $\hat{\beta}_n = O_p(n^{-1/2})$ and such solution is unique, as $(\partial/\partial b)\xi_n^{\text{tr}}(b) = \Lambda_\psi + o(1)$ for $\|b - \beta\| \leq Mn^{-1/2}$ and $\xi_n^{\text{tr}}(\beta) = O_p(n^{-1/2})$. Since $(b - \beta)^{\text{tr}} \xi_\infty(b) \neq 0$, (2.5) follows from (3.1) and (3.4). Finally, (3.5) implies $(1 + o(1))(\hat{\beta}_n - \beta) = -(1 + o(1))\Lambda_\psi^{-1} \sum_{i=1}^n \rho_\psi(X_i; \beta) / n$, which implies (2.8) as $\text{Cov}(\rho_\psi(X; \beta)) = \Sigma_\psi$. \square

3.2. *The Arcones–Giné inequality.* The proof of Theorem 3.1 is based on the Arcones and Giné (1993) extension of the Bernstein inequality to completely degenerate U -statistics. A combination of their inequality and the Hoeffding decomposition is given here, providing a bound on the tail probability of mean zero U -statistics to be conveniently used in our proofs. Let E_{i_1, \dots, i_l}^* be the conditional expectation given X_{i_1}, \dots, X_{i_l} .

LEMMA 3.1. *Let $f_n(x_1, \dots, x_k; b)$ be vector-valued Borel functions such that*

$$(3.6) \quad \|E_{i_1, \dots, i_l}^* f_n(\mathbf{X}_{1, k}; b)\| \leq c_{n, l}, \quad E \|E_{i_1, \dots, i_l}^* f_n(\mathbf{X}_{1, k}; b)\|^2 \leq \sigma_{n, l}^2,$$

for all $1 \leq i_1 < \dots < i_l \leq k$, $1 \leq l \leq k$. Suppose

$$(3.7) \quad c_{n, l} \leq (n/t_n)^{(l+1)/2}, \quad \sigma_{n, l}^2 \leq (n/t_n)^l, \quad 1 \leq l \leq k.$$

Then there exist finite positive constants c_1^* and c_2^* depending on k only such that

$$(3.8) \quad \Pr \left\{ \left\| \frac{\sum_{\mathbf{i} \in I_k^n} \{f_n(\mathbf{X}_i; b) - Ef_n(\mathbf{X}_i; b)\}}{\binom{n}{k} k!} \right\| > \varepsilon \right\} \leq c_1^* \exp(-c_2^* \varepsilon^2 t_n)$$

for all $0 < \varepsilon \leq 1$. Furthermore, if $\|f_n(\mathbf{x}; b) - f_n(\mathbf{x}; b')\| \leq M_0 n^{M_0} \|b - b'\|$ for some $M_0 < \infty$ and $t_n / \log n \rightarrow \infty$, then for all $0 < \varepsilon < M < \infty$,

$$(3.9) \quad \Pr \left\{ \sup_{\|b - \beta\| \leq M} \left\| \frac{\sum_{\mathbf{i} \in I_k^n} \{f_n(\mathbf{X}_i; b) - Ef_n(\mathbf{X}_i; b)\}}{\binom{n}{k} k!} \right\| > \varepsilon \right\} \rightarrow 0.$$

PROOF. Assume $Ef_n = 0$. We first prove (3.9) based on (3.8). Due to the smoothness condition on $f_n(\mathbf{x}; \cdot)$, there exist sets of vectors $\{b_j: \|b_j - \beta\| \leq M, j \leq N_n\}$ with $(\log N_n) / \log n = O(1)$ such that $\min_{1 \leq j \leq N_n} \|f_n(x; b) - f_n(x; b_j)\| \leq \varepsilon/2$ for all $\|b - \beta\| \leq M$. Thus, by (3.8) the left-hand side of (3.9) is bounded by

$$N_n \sup_{j \leq N_n} \Pr \left\{ \left\| \frac{\sum_{\mathbf{i} \in I_k^n} f_n(\mathbf{X}_i; b_j)}{\binom{n}{k} k!} \right\| > \varepsilon/2 \right\} \leq N_n c_1^* \exp(-c_2^* \min\{(\varepsilon/2)^2, 1\} t_n) \rightarrow 0.$$

Let us prove (3.8). By the Hoeffding decomposition,

$$(3.10) \quad \frac{1}{\binom{n}{k} k!} \sum_{\mathbf{i} \in I_k^n} f_n(\mathbf{X}_i; b) = \sum_{l=0}^k \frac{\binom{k}{l}}{\binom{n}{l} l!} \sum_{\mathbf{i} \in I_l^n} f_{n,l}(\mathbf{X}_i; b),$$

where $f_{n,l}$ are completely degenerate kernels. Furthermore, since $Ef_n = 0$, $f_{n,l}$ is a sum of $2^l - 1$ terms of the form $\pm \mathbf{E}_{i_1, \dots, i_j}^* f_n(X_1, \dots, X_k; b)$, $1 \leq j \leq l$. Thus, $\|f_{n,l}\| \leq \tilde{c}_l (n/t_n)^{(l+1)/2}$ and $\|\text{Cov}(f_{n,l})\| \leq \tilde{c}_l (n/t_n)^l$ for some universal constants \tilde{c}_l and (3.8) follows from Proposition 2.3(c) of Arcones and Giné (1993).

3.3. Proof of Theorem 3.1. It follows from (2.2) that

$$(3.11) \quad \xi_n(b) = \sum_{\mathbf{i} \in I_3^n} \frac{h_n(\mathbf{X}_i; b)}{\binom{n}{3} 3!}, \quad E\xi_n(b) = Eh_n(b),$$

with $h_n(b) = h_n(\mathbf{X}_{1,3}; b) = \psi_b h_{n,0}(b)$, $\psi_b = \psi(T_{1,b})$ and $h_{n,0}(b) = (\delta_1 - \delta_2)(Z_1 - Z_3)K_{n12}K_{n13}$ and

$$\begin{aligned} \frac{\partial h_n^{\text{tr}}(b)}{\partial b} &= \frac{\psi_b}{\alpha_n} (\delta_1 - \delta_2)(Z_2 - Z_1)(Z_1 - Z_3)^{\text{tr}} K'_{n12} K_{n13} \\ (3.12) \quad &+ \frac{\psi_b}{\alpha_n} (\delta_1 - \delta_2)(Z_3 - Z_1)(Z_1 - Z_3)^{\text{tr}} K_{n12} K'_{n13} - Z_1 \psi'(T_{1,b}) h_{n,0}^{\text{tr}}(b) \\ &= h'_{n,1}(b) + h'_{n,2}(b) + h'_{n,3}(b), \quad \text{say.} \end{aligned}$$

LEMMA 3.2. *Let $\tilde{K}_j(t)$ be Borel functions and $\tilde{a}_j \geq \tilde{a}$. Let $M_p = \max_j \|\tilde{K}_j\|_p$ with respect to the Lebesgue measure. Then for $1 \leq i_1 < \dots < i_l \leq k$ and $l \geq 1$,*

$$\begin{aligned} E_{i_1, \dots, i_l}^* \prod_{j=2}^k \left| \frac{1}{\tilde{a}_j} \tilde{K}_j \left(\frac{T_{1,b} - T_{j,b}}{\tilde{a}_j} \right) \right| &\leq (\|g(\cdot; b)\|_{\infty} M_1)^{k-l} (M_{\infty}/\tilde{a})^{l-1}, \\ E \left(E_{i_1, \dots, i_l}^* \prod_{j=2}^k \left| \frac{1}{\tilde{a}_j} \tilde{K}_j \left(\frac{T_{1,b} - T_{j,b}}{\tilde{a}_j} \right) \right| \right)^2 &\leq (\|g(\cdot; b)\|_{\infty} M_1)^{2k-l-1} (M_{\infty}/\tilde{a})^{l-1}. \end{aligned}$$

LEMMA 3.3. *Let $h_n(b)$, $\xi_{\infty}(b)$, $\rho_{\psi}(X_1; b)$ and Λ_{ψ} be as in (3.11), (2.3), (2.8) and (2.7), respectively. Set $\eta = b - \beta$ and $a_{n,b} = \|\eta\| + \alpha_n$. Suppose (2.6) holds, $F \in \text{Lip}(M_0, \alpha_1)$ and $\psi \in \text{Lip}(M_0, \alpha_0)$ for some $0 < \alpha_0 \leq \min(\alpha, 1/2)$ and $0 \leq \alpha_1 \leq 1$. Then, uniformly in b*

$$\begin{aligned} (3.13) \quad \|Eh_n(b) - \xi_{\infty}(b)\| &\leq \tilde{M} \|\psi\|_{\infty} \alpha_n^{\alpha} a_{n,b}, \\ \|\xi_{\infty}(b) - \Lambda_{\psi} \eta\| &\leq \tilde{M} \|\eta\|^{1+\alpha_0}, \end{aligned}$$

$$\begin{aligned} (3.14) \quad E \|E_1^* h_n(b) - \rho_{\psi}(X_1; \beta)\|^2 &\leq \tilde{M} a_{n,b}^{2\alpha_0}, \\ E \|E_i^* h_n(b)\|^2 &\leq \tilde{M} (a_{n,b} + a_n^{2\alpha}), \quad i = 2, 3. \end{aligned}$$

If in addition $\psi \in \text{Lip}(M_0, 1)$, then

$$(3.15) \quad \|E(\partial/\partial b) h_n^{\text{tr}}(b) - \Lambda_{\psi}\| \leq \tilde{M} a_n^{\alpha-1} a_{n,b},$$

$$(3.16) \quad \sup_b E \|E_i^* (\partial/\partial b) h_n^{\text{tr}}(b)\|^2 = O(a_n^{2(\alpha-1)} + a_n^{\alpha_1-1}), \quad 1 \leq i \leq 3.$$

The proof of Lemma 3.2 is standard and omitted. The proof of Lemma 3.3 is given after the proof of Theorem 3.1. Some parts of Lemma 3.3 are used in Section 4.

PROOF OF THEOREM 3.1. By (3.11), (3.2) follows directly from (3.15). By (3.12) and Lemma 3.2, (3.6) holds for $c_{n,l} = O(a_n^{-l}) = o(1)(m/\log n)^{(l+1)/2}$, $1 \leq l \leq 3$, and $\sigma_{n,l}^2 = O(a_n^{-l-1}) = o(1)(n/\log n)^l$, $l \geq 2$, for $f_n = (\partial/\partial b) h_n^{\text{tr}}(b)$. This and (3.16) imply (3.6) and (3.7) for all $1 \leq l \leq k = 3$ with $t_n/\log n \rightarrow \infty$ by $a_n^{2(\alpha-1)} + a_n^{-1} = o(n/\log n)$, so that (3.1) holds via Lemma 3.1. Note that the smoothness condition for (3.9) holds easily with crude bounds. Similarly,

the $\sigma_{n,l}^2$ part of (3.6) is of the order $o(n^l)$ for $f_n = \sqrt{n} \{h_n(b) - \rho_\psi(X_1; \beta)\}$ and all $0 \leq l \leq 3$ by Lemma 3.2 and (3.14), so that (3.3) holds by (3.10). Finally (3.4) follows from (3.13) as $a_n^{1+\alpha} \sqrt{n} \rightarrow 0$. \square

PROOF OF LEMMA 3.3. We shall prove (3.13), (3.15), (3.16) and then (3.14). Assume $\|\psi\|_\infty \leq 1$. Let \tilde{E}_i^* be the conditional expectation given (X_i, \mathbf{Z}) , where $\mathbf{Z} = (Z_1, Z_2, \dots)$. For all $\tilde{K} \in L_1(dx)$ and bandwidth $\tilde{a}_n > 0$,

$$(3.17) \quad \begin{aligned} \tilde{E}_i^* \tilde{K}_{nij} &= (g_{j,b} * \tilde{K}_n)(T_{i,b}), \\ \tilde{E}_i^* \delta_j \tilde{K}_{nij} &= \tilde{E}_i^* F_{j,b}(T_{j,b}) \tilde{K}_{nij} = ((F_{j,b} g_{j,b}) * \tilde{K}_n)(T_{i,b}), \end{aligned}$$

as $g_{j,b}$ is the conditional density of $T_{j,b}$ given Z_j and $E[\delta_j | T_j, Z_j] = F_{j,b}(T_{j,b})$. By (2.6),

$$(3.18) \quad \{g(t; b), (g\mu)(t; b), (g^2v)(t; b)\} \in \text{Lip}(\tilde{M}, \alpha) \quad \text{in } (t, b)$$

with $\mu(t; b) = EZg(t + b^{\text{tr}}Z|Z)/g(t; b) = E[Z|T_b = t]$ and $v(t; b)$ as in (2.7), so that

$$(3.19) \quad \begin{aligned} g_{j,b} * \tilde{K}_n &= \mu(\tilde{K}) g_{j,b} + \tilde{M}_\infty \tilde{a}_n^\alpha, \\ (F_{j,b} g_{j,b}) * \tilde{K}_n &= g_{j,b} (F_{j,b} * \tilde{K}_n) + \tilde{M}_\infty \tilde{a}_n^\alpha, \end{aligned}$$

where $\mu(\tilde{K}) = \int \tilde{K}(u) du$ ($\mu(K) = 1$). Since F is a distribution, by (2.6),

$$(3.20) \quad \begin{aligned} E[F(T-t+\Delta) - F(T-t)|Z] \\ \leq \|g\|_\infty \int F(u+\Delta) - F(u) du = \|g\|_\infty \Delta \end{aligned}$$

for all $\Delta > 0$ and t [i.e., $F \in \text{Lip}(M_0, 1)$ in $L_1(g(u|z) du)$ uniformly in z]. For example,

$$(3.21) \quad \begin{aligned} & \left(| \{F(T_\beta) - F_{2,b}\} \{g_{2,b} - g_{2,b}(T_b)\} | * |\tilde{K}_n| \right)_b \\ & \leq \tilde{M} a_n^\alpha \left(|F(T_\beta) - F_{2,b}| * |\tilde{K}_n| \right)_b \\ & \leq a_n^\alpha \min(|\tilde{M}_1| a_{n,b}, |\tilde{M}_\infty| a_{n,b}^{\alpha_1}) \leq a_n^\alpha |\tilde{M}_2| a_{n,b}^{(1+\alpha_1)/2} \end{aligned}$$

by (3.19) and (3.20). The last inequality above follows from $E\|U\|^2 \leq \|U\|_\infty E\|U\|$ for all random vectors U . Since $\mu(K) = 1$, it follows from (3.17)–(3.21) that

$$(3.22) \quad \begin{aligned} & E h_n(b) \\ &= E \tilde{E}_1^* h_n(b) \\ &= E \psi_b(Z - Z_3) \left(\left(\{F(T_\beta) - F_{2,b}\} g_{2,b} \right) * K_n \right)_b (g_{3,b} * K_n)_b \\ &= E \psi_b(Z - Z_3) \left((F(T_\beta) - F_{2,b}) * K_n \right)_b (g_{2,b} g_{3,b})(T_b) + \tilde{M} a_n^\alpha a_{n,b} \\ &= E \psi_b g_b(Z - \mu_b) \left((F(T_\beta) - F_{2,b}) * K_n \right)_b g_{2,b}(T_b) + \tilde{M} a_n^\alpha a_{n,b} \\ &= \xi_\infty(b) + \tilde{M} a_n^\alpha a_{n,b}, \end{aligned}$$

as $E[(Z - \mu_b)(F_{2,b} * K_n)_b | T_b, Z_2] = E[(F_{2,b} * K_n)_b | T_b, Z_2]E[Z - \mu_b | T_b] = 0$ and by (2.3) $\xi_\infty(b) = E\psi_b g_b^2(Z - \mu_b)F(T_\beta) = E\psi_b g_b^2(Z - \mu_b)\{F(T_\beta) - F(T_b)\}$. This gives the first part of (3.13). The second part of (3.13) follows from

$$\begin{aligned} \xi_\infty(b) &= E \int \int_u^{u + \eta^{\text{tr}} Z} dF(t) \psi(u) g_b^2(u) (Z - \mu(u; b)) g(u + b^{\text{tr}} Z / Z) du \\ &= E \int \int_{t - \eta^{\text{tr}} Z}^t \psi(u) g_b^2(u) (Z - \mu(u; b)) g(u + b^{\text{tr}} Z / Z) du dF(t) \\ &= \Lambda_\psi \eta + \tilde{M} \|\eta\|^{1 + \alpha_0} \end{aligned}$$

by (3.18). The rest of the proof is similar and we omit some details.

For (3.15) consider $h'_{n,j}(b)$ in (3.12). As in (3.22), $Eh'_{n,3}(b) = \tilde{M} a_n^\alpha \|\psi'\|_\infty$. Since $\mu(K') = 0$, $Eh'_{n,2}(b) = \tilde{M} a_n \|\mathbf{g}_{3,b} * K'_n\|_\infty / a_n = \tilde{M} a_n^{\alpha-1} a_{n,b}$. Since $F_{2,b} * K'_n / a_n = (dF_{2,b}) * K'_n = \int K'_n(\cdot + \eta^{\text{tr}} Z_2 - u) dF(u)$ and $F(T_\beta) * K'_n = 0$,

$$\begin{aligned} (3.23) \quad Eh'_{n,1}(b) &= -E\psi_b h(\mathbf{Z}) \mathbf{g}_{2,b} \mathbf{g}_{3,b} ((F(T_\beta) - F_{2,b}) * K'_n)_b / a_n \\ &\quad + \tilde{M} a_n^{\alpha-1} a_{n,b} \\ &= \int E\psi_b h(\mathbf{Z}) \mathbf{g}_{2,b} \mathbf{g}_{3,b} K'_n(T_b + \eta^{\text{tr}} Z_2 - u) dF(u) \\ &\quad + \tilde{M} a_n^{\alpha-1} a_{n,b} \\ &= \int \psi(u) v(u; b) g^3(u; b) dF(u) + \tilde{M} a_n^\alpha \\ &\quad + \tilde{M} a_n^{\alpha-1} a_{n,b}, \end{aligned}$$

where $h(\mathbf{Z}) = (Z - Z_2)(Z - Z_3)^{\text{tr}}$. Thus (3.15) holds by (2.7) and (3.18) as $a_n^\alpha \leq a_n^{\alpha-1} a_{n,b}$.

By Lemma 3.2, $E_i^* h'_{n,3}(b) = \tilde{M}_\infty \|\psi'\|_\infty$ for all i . Since $\mu(K') = 0$, $\mathbf{g}_{j,b} * K'_n = \tilde{M}_\infty a_n^\alpha$ by (3.19), which implies $\|\tilde{E}_i^* h'_{n,j}(b)\| \leq \tilde{M}_\infty a_n^{\alpha-1} + |\tilde{M}_1| \wedge |\tilde{M}_\infty a_n^{\alpha_1-1}|$, for all i and $j = 1, 2$ by (3.12) and (3.20). For example, by (3.19) and (3.20) and with $\tilde{K}'(t) = K'(-t)$,

$$\begin{aligned} \tilde{E}_2^* h'_{n,1}(b) &= \tilde{M}_\infty \tilde{E}_2^* (\{(F_{1,b} - \delta_2) \psi_b \mathbf{g}_{1,b} \mathbf{g}_{3,b}\} * \tilde{K}'(T_{2,b}) / a_n + \tilde{M}_\infty a_n^{\alpha-1} \\ &= \tilde{M}_\infty \tilde{E}_2^* (F_{1,b} * \tilde{K}'(T_{2,b}) / a_n + \tilde{M}_\infty a_n^{\alpha-1} \\ &= \tilde{M}_2 a_n^{(\alpha_1-1)/2} + \tilde{M}_\infty a_n^{\alpha-1}. \end{aligned}$$

In general we compute first the convolution without F and then the one with F . For \tilde{E}_3^* , we split $F(T_{1,\beta}) - F(T_{2,\beta})$ into the difference of $F(T_{j,\beta}) - F(T_{3,\beta})$. Thus (3.16) holds.

Finally, $E_3^* h_n(\beta) = \tilde{M}_2 a_n^{\alpha \wedge ((1 + \alpha_1)/2)}$ by (3.19) and (3.20),

$$E_2^* h_n(\beta) = E_2^* \psi_b g_\beta (Z - \mu_\beta) (F(T_\beta) - \delta_2) K_{n12}(\beta) + \tilde{M}_\infty a_n^\alpha = \tilde{M}_\infty a_n^\alpha$$

as $E_2^*[(Z - \mu_\beta)K_{n12}(\beta)|T_\beta] = 0$, and

$$E_1^*h_n(\beta) = \psi_\beta g_\beta^2 E_1^*(Z - \mu_\beta)((\delta - F_{2,\beta}) * K_n)_\beta + \tilde{M}_\infty a_n^\alpha$$

implies $E_1^*h_n(\beta) = \rho_\psi(X_1; \beta) + \tilde{M}_2 a_n^{\alpha \wedge (1+\alpha_1)/2}$. Hence, the proof of (3.14) is complete. \square

4. Proof of Theorem 2.3. The methods in Section 3 are to be used again.

4.1. *Asymptotic local linearity and normality.* Define

$$(4.1) \quad \tilde{\rho}_n(\mathbf{X}_{1,2}) = \frac{\{\delta - F(T_\beta)\}\{\delta_2 - F(T_{2,\beta})\}(Z - \mu_\beta)g^3(T_\beta; \beta)\tilde{K}'_{n12}(\beta)}{a'_n\{c_n + D(T_\beta; \beta)\}},$$

where $D(t; \beta)$ is as in (2.20) and $\tilde{K}'(t)$ is as in (2.18) with bandwidth a'_n .

THEOREM 4.1. Let $\hat{\xi}_n$ be as in (2.14) with (a_n, a'_n, c_n) such that

$$(4.2) \quad n^{1/3}a_n \rightarrow \tau'_0 \in (0, \infty], \frac{a_n}{a'_n} \leq M_0, c_n^{-2(k_0+2)} \left\{ \frac{\log n}{na_n^2} \right\}^{k_0+1} \rightarrow 0$$

for some integer $k_0 \geq 6$ and that

$$(4.3) \quad \begin{aligned} & \frac{\log n}{na_n^2 c_n^2} \left(1 + \frac{\log n}{na_n^3} \right) \rightarrow 0, \\ & \frac{a_n^\alpha}{c_n^2} + \frac{(a'_n)^\alpha}{c_n} + \frac{\sqrt{n} a_n^{1+\alpha}}{c_n} \rightarrow 0, \quad c_n \rightarrow 0. \end{aligned}$$

Suppose (2.6), (2.19) and (2.20) hold and I_* is finite and positive definite. Then,

$$(4.4) \quad \sup_{\|\eta\| \leq M/\sqrt{n}} \|(\partial/\partial b)\hat{\xi}_n^{\text{tr}}(b) - I_*\| = o_P(1)$$

with $\eta = b - \beta$, and for ρ_* and $\tilde{\rho}_n$ as in (2.9) and (4.1),

$$(4.5) \quad \hat{\xi}_n(\beta) = n^{-1} \sum_{i=1}^n \rho_*(X_i; \beta) + \frac{1}{\binom{n}{2} 2!} \sum_{i \in I_2^n} \tilde{\rho}_n(\mathbf{X}_i) + o_P(n^{-1/2}).$$

PROOF OF THEOREM 2.3. Clearly, (4.2) and (4.3) hold. Let $z^{\otimes 2} = zz^{\text{tr}}$. Since $F(T_{2,\beta}) - F(T_\beta) = \tilde{M}_\infty a'_n$ for $\tilde{K}'_{n12}(\beta) \neq 0$, by (4.1) and (2.20),

$$(4.6) \quad \begin{aligned} & E\tilde{\rho}_n^{\otimes 2}(\mathbf{X}_{1,2})/n \\ &= E \frac{\{F(T_\beta)(1 - F(T_\beta)) + \tilde{M}_\infty a'_n\}^2 v(T_\beta; \beta)g^6(T_\beta; \beta)\{\tilde{K}'_{n12}(\beta)\}^2}{n(a'_n)^2\{c_n + D(T_\beta; \beta)\}^2} \\ &= E \frac{\tilde{\mu}_2 v(T_\beta; \beta)D^2(T_\beta; \beta)/\{g(T_\beta; \beta)\mu_2^2\}}{n(a'_n)^3\{c_n + D(T_\beta; \beta)\}^2} + o(1) \\ &= \sigma_n^2 + o(1) \rightarrow \sigma^2, \end{aligned}$$

as $a'_n E_1^*(\tilde{K}'_{n12})^2 = g_\beta \tilde{\mu}_2 + \tilde{M}_\infty (a'_n)^\alpha$ by (3.17), (3.19) and (2.18). Since $E\rho_*^{\otimes 2}(X_1; \beta) = I_*$, (4.6) implies $\sqrt{n} \hat{\xi}_n(\beta) = O_p(1)$ by (4.5). The conclusions of Theorem 2.3(i) follow from (4.4) and the asymptotic normality of $\sqrt{n} \hat{\xi}_n(\beta)$ as $E\rho_*(X_i; \beta)\tilde{\rho}_n(\mathbf{X}_i) = 0$.

Let us prove the asymptotic normality of $\sqrt{n} \hat{\xi}_n(\beta)$. Let $\varepsilon > 0$. Let M_1 be large enough such that $\tilde{K}'_{nij} = 0$ for $|T_{i,\beta} - T_{j,\beta}| \geq M_1 a'_n$. Let the real line be divided into disjoint intervals $B_0, A_1, B_1, \dots, A_k, B_k$ from left to right such that the length of A_1, \dots, A_k is $M_1 a'_n$. It can be easily seen that $E\tilde{\rho}_n^{\otimes 2}(\mathbf{X}_{1,2})I\{T_{i,\beta} \in A_j\} = o(1), i = 1, 2$, as in (4.6), so that by (4.5) $\sqrt{n} \hat{\xi}_n(\beta) = \sum_{j=0}^k S_{n,j} + o_p(1)$, where $S_{n,j} = S_{n,j,1} + S_{n,j,2}$ with

$$S_{n,j,1} = \sum_{i=1}^n \frac{\rho_*(X_i; \beta)}{\sqrt{n}} I_{ij}, \quad S_{n,j,2} = \frac{\sqrt{n}}{\binom{n}{2} 2!} \sum_{\mathbf{i} \in I_2^n} \tilde{\rho}_n(\mathbf{X}_i) I_{i_1 j} I_{i_2 j},$$

$$I_{ij} = I_{\{T_{i,\beta} \in B_j\}}.$$

Since B_j are separated by intervals of length $M_1 a'_n$, $S_{n,j}$ are conditionally independent given (N_0, \dots, N_k) , where $N_j = \sum_i I_{ij}$. For fixed k , the strong law holds for N and

$$\begin{aligned} & E|\eta_0^{\text{tr}} S_{n,j,2}|^4 - 3(1 + o(1))(E|\eta_0^{\text{tr}} S_{n,j,2}|^2)^2 \\ &= \left(\frac{\sqrt{n}}{\binom{n}{2} 2!} \right)^4 \sum_{\mathbf{i} \in I_2^n} E|\eta_0^{\text{tr}} \tilde{\rho}_n(\mathbf{X}_i)|^4 I_{i_1 j} I_{i_2 j} \\ &\leq \tilde{M} n^{-4} \|\tilde{\rho}_n\|_\infty^4 = \tilde{M} (n(a'_n)^2 c_n)^{-4} = o(1) \quad \forall \|\eta_0\| = 1. \end{aligned}$$

Furthermore, we may choose B_j such that $E\|S_{n,j}\|^2 \leq \tilde{M}/k$. Hence, the Lindeberg condition holds as $n \rightarrow \infty$ and then $k \rightarrow \infty$, and the asymptotic normality of $\sum_0^k S_{n,j}$ and then $\sqrt{n} \hat{\xi}_n(\beta)$ follows. Finally, (2.22) follows directly from (4.4), (4.5) and (4.6). \square

4.2. *Proof of Theorem 4.1.* Our plan is to prove (4.5) and the following statements: for all $0 < \varepsilon < M < \infty$,

$$(4.7) \quad \lim_{n \rightarrow \infty} P \left\{ \sup_{\|b - \beta\| \leq M a_n} \left\| \frac{\partial}{\partial b} \hat{\xi}_n^{\text{tr}}(b) - E \frac{\partial}{\partial b} \hat{\xi}_n^{\text{tr}}(b) \right\| > \varepsilon \right\} = 0,$$

$$(4.8) \quad \sup_{\|b - \beta\| \leq M/\sqrt{n}} \left\| E \frac{\partial}{\partial b} \hat{\xi}_n^{\text{tr}}(b) - I_* \right\| \rightarrow 0$$

Throughout the proof, the convergence (to zero) of many functions of (a_n, a'_n, c_n) are derived from (4.2) and (4.3), but they are not necessarily referred to when used. Let $\tilde{K}_{nijk} = \tilde{K}_{nijk}(b) = \{K_{nij}^{[1]} K_{nik}^{[0]} - K_{nij}^{[0]} K_{nik}^{[1]}\}$. Since

$\tilde{K}_{nijk} + \tilde{K}_{nikj} = 0$, we may center δ_4 in (2.15) without changing the value of $\hat{\xi}_n(b)$ in (2.14). In the sequel, let

$$(4.9) \quad \begin{aligned} \tilde{h}_{n,0}(b) &= \tilde{h}_{n,0}(\mathbf{X}_{1,5}; b) \\ &= (\delta_1 - \delta_2)(Z_1 - Z_3)K_{n12}K_{n13}\{\delta_4 - F(T_{1,\beta})\}\tilde{K}_{n145}. \end{aligned}$$

We need a lemma. Let $\phi_n(t) = 1/(c_n + t)$. Define

$$(4.10) \quad \begin{aligned} \bar{D}_{n,i} &= \bar{D}_n(T_{i_1,b}; b) = E_{\mathbf{i}}^* \sum_{\mathbf{j} \in I_4^n(\mathbf{i})} \frac{\tilde{h}_n(\mathbf{X}_{\mathbf{j}}, T_{i_1,b}; b)}{\binom{n-5}{4} 4!} \\ &= E\tilde{h}_n(\mathbf{X}_{1,4}, t; b)|_{t=T_{i_1,b}}. \end{aligned}$$

Decompose $(\delta_1 - \delta_2)(\delta_4 - F(T_{1,\beta}))$ into the following three terms with different orders of degeneracy:

$$(4.11) \quad \begin{aligned} f_{0,0} &= \{F(T_{1,\beta}) - F(T_{2,\beta})\}\{F(T_{4,\beta}) - F(T_{1,\beta})\}, \\ f_{0,1} &= \{\delta_1 - F(T_{1,\beta}) - \delta_2 + F(T_{2,\beta})\}\{F(T_{4,\beta}) - F(T_{1,\beta})\} \\ &\quad + \{F(T_{1,\beta}) - F(T_{2,\beta})\}\{\delta_4 - F(T_{4,\beta})\}, \\ f_{0,2} &= \{\delta_1 - F(T_{1,\beta}) - \delta_2 + F(T_{2,\beta})\}\{\delta_4 - F(T_{4,\beta})\}. \end{aligned}$$

LEMMA 4.1. *Let $\tilde{h}_{n,0}(b)$, \bar{D}_n , I_* and $f_{0,s}$ be as in (4.9), (4.10), (2.9) and (4.11), respectively. Set $\phi_n(b) = \phi_n(\bar{D}_n(T_b; b))$,*

$$h_n^* = \tilde{\phi}_n(\beta)(Z_1 - Z_3)K_{n12}(\beta)K_{n13}(\beta)\tilde{K}_{n145}(\beta)$$

and $f_{n,0,s} = \sqrt{n}f_{0,s}h_n^*/a'_n$. Suppose conditions of Theorem 4.1 hold. Then

$$(4.12) \quad \lim_{n \rightarrow \infty} \sup_{\|b-\beta\| \leq M/\sqrt{n}} \left\| E \frac{\tilde{\phi}_n(b)}{a'_n} \frac{\partial \tilde{h}_{n,0}^{\text{tr}}(b)}{\partial b} - I_* \right\| = 0,$$

$$(4.13) \quad \lim_{n \rightarrow \infty} E \left\| E_{i_1, i_2}^* \{f_{n,0,2} - \sqrt{n} \tilde{\rho}_n(\mathbf{X}_{(1,4)})\} \right\|^2 / n^2 = 0 \quad \forall (i_1, i_2) \in I_2^5,$$

$$(4.14) \quad \lim_{n \rightarrow \infty} E \left\| E_i^* \{f_{n,0,1} - \sqrt{n} \rho_*(X_1; \beta)\} \right\|^2 / n = 0, \quad 1 \leq i \leq 5$$

and $\lim_{n \rightarrow \infty} E f_{n,0,0} = o(1)$.

The proof of Lemma 4.1 is deferred to the end. We prove (4.7), (4.8) and (4.5) in three steps.

STEP 1. *Proof of (4.7).* Let $\phi_n^{(k)}(t) = (d/dt)^k \phi_n(t)$. By (2.14)

$$(4.15) \quad \begin{aligned} \frac{\partial}{\partial b} \hat{\xi}_n^{tr}(b) &= \sum_{\mathbf{i} \in I_5^n} \left\{ \frac{\partial}{\partial b} \hat{D}_{n,\mathbf{i}}(b) \right\} \phi_n^{(1)}(\hat{D}_{n,\mathbf{i}}(b)) \frac{\tilde{h}_{n,0}^{tr}(\mathbf{X}_\mathbf{i}; b)}{\alpha'_n \binom{n}{5} 5!} \\ &+ \sum_{\mathbf{i} \in I_5^n} \phi_n(\hat{D}_{n,\mathbf{i}}(b)) \frac{\partial}{\partial b} \frac{\tilde{h}_{n,0}^{tr}(\mathbf{X}_\mathbf{i}; b)}{\alpha'_n \binom{n}{5} 5!}. \end{aligned}$$

Let $\tilde{\xi}_n(b)$ be the first sum above. We shall prove

$$(4.16) \quad P \left\{ \sup_{\|b - \beta\| \leq Ma_n} \|\tilde{\xi}_n(b) - E\tilde{\xi}_n(b)\| > \varepsilon \right\} \rightarrow 0 \quad \forall \varepsilon > 0.$$

The rest of the proofs are similar, so that most details are provided here. Taking the Taylor expansion of $\phi_n^{(1)}(\hat{D}_{n,\mathbf{i}})$ at $\bar{D}_{n,\mathbf{i}}$, we find that (4.16) follows from

$$(4.17) \quad \begin{aligned} \sup_{\|\eta\| \leq Ma_n} \|\tilde{\xi}_{n,k}(b) - E\tilde{\xi}_{n,k}(b)\| &= o_P(1), \\ \sup_{\|\eta\| \leq Ma_n} \|\tilde{R}_n(b) - E\tilde{R}_n(b)\| &= o_P(1), \end{aligned}$$

$1 \leq k \leq k_0$, where k_0 is as in (4.2), $\tilde{R}_n(b) = \tilde{\xi}_n(b) - \sum_{k=1}^{k_0} \tilde{\xi}_{n,k}(b)$ and

$$(4.18) \quad \tilde{\xi}_{n,k}(b) = \sum_{\mathbf{i} \in I_5^n} \left\{ \frac{\partial}{\partial b} \hat{D}_{n,\mathbf{i}}(b) \right\} (\hat{D}_{n,\mathbf{i}} - \bar{D}_{n,\mathbf{i}})^{k-1} \frac{\phi_n^{(k)}(\bar{D}_{n,\mathbf{i}}) \tilde{h}_{n,0}^{tr}(\mathbf{X}_\mathbf{i}; b)}{\alpha'_n \binom{n}{5} 5! (k-1)!}.$$

Consider (4.18) for a fixed k . By (2.14), $(\hat{D}_{n,\mathbf{i}} - \bar{D}_{n,\mathbf{i}})^{k-1} (\partial/\partial b) \hat{D}_{n,\mathbf{i}}$ is

$$(4.19) \quad \left\{ \sum_{\mathbf{j} \in I_4^n(\mathbf{i})} \frac{\tilde{h}_n(\mathbf{X}_\mathbf{j}, T_{i_1,b}; b) - \bar{D}_{n,\mathbf{i}}}{\binom{n-5}{4} 4!} \right\}^{k-1} \sum_{\mathbf{j} \in I_4^n(\mathbf{i})} \frac{\partial}{\partial b} \frac{\tilde{h}_n(\mathbf{X}_\mathbf{j}, T_{i_1,b}; b)}{\binom{n-5}{4} 4!},$$

which can be written as a sum of products such that the number of independent X -vectors in these products depends on the ties in the labels \mathbf{j} in the sums in (4.19). Thus, (4.18), (4.10), (4.9) and (2.14)–(2.16) imply that, with $\tilde{f}_{n,k,j}(\mathbf{x}) = \tilde{f}_{n,k,j}(\mathbf{x}; b)$,

$$(4.20) \quad \begin{aligned} \tilde{\xi}_{n,k}(b) &= \sum_{j=9}^{4k+5} \sum_{\mathbf{i} \in I_j^n} \frac{\tilde{f}_{n,k,j}(\mathbf{X}_\mathbf{i})}{\binom{n}{j} j!}, \\ \tilde{f}_{n,k,j}(\mathbf{X}_{1,j}) &= (\delta_1 - \delta_2)(\delta_4 - F(T_{1,\beta})) \tilde{h}_{n,k,j}. \end{aligned}$$

where $\tilde{h}_{n,k,j} = \tilde{h}_{n,k,j}(\mathbf{X}_{1,j}; b)$ is a finite sum of functions of the form

$$(4.21) \quad \tilde{v}_{n,j} \frac{\phi_n^{(k)}(\bar{D}_n(T_{1,b}; b))}{a'_n(k+1)!} \left(E_{1,i_1,\dots,i_l}^* \frac{Z_{i_0} - Z_1}{a_{n,i_0}} h(\delta_{6,m}) \prod_{i=6}^m \tilde{K}_{n1i}^{[i]} \right) \times (Z_1 - Z_3)^{\text{tr}} \prod_{i=2}^5 \tilde{K}_{n1i}^{[i]},$$

for smooth kernels $\tilde{K}^{[i]}$ with bounded support, finite $\|\tilde{K}^{[i]}\|_p$ as in Lemma 3.2 and bandwidth $a_{n,i} = a_n$ or a'_n . For example, $\{E_1^* \tilde{h}_n(\mathbf{X}_{6,9}, T_{1,b}; b)\} \times (\partial/\partial b) \tilde{h}_n(\mathbf{X}_{6,9}, T_{1,b}; b)$ is written as $E_{1,\dots,9}^* \tilde{h}_n(X_{10,13}, T_{1,b}; b) (\partial/\partial b) \tilde{h}_n(\mathbf{X}_{6,9}, T_{1,b}; b)$, while $K_{n16}^{[2]}(b) (\partial/\partial b) K_{n16}(b)$ is written as $(Z_6 - Z_1)(a'_n a_n^2)^{-1} \tilde{K}((T_{1,b} - T_{6,b})/a_n)$, with the extra a_n being part of \tilde{v}_{nj} , where $\tilde{K}(t) = K^{[2]} \times (ta_n/a'_n)(d/dt)K(t)$. The dependence of \tilde{K} on a_n/a'_n is omitted as we only need $\|\tilde{K}\|_p \leq \|K^{[2]}\|_\infty \|K'\|_p$. Here in (4.21), $6 \leq i_0 \leq 9$ is the index of the differentiated kernel of \tilde{h}_n in (2.16) and (4.19), $h(\delta_{6,m})$ is a $\{0, 1\}$ -valued function of $\delta_6, \dots, \delta_m$ and $0 \leq \tilde{v}_{n,j} \leq \tilde{M}/(na_n)^{4k+5-j}$ due to the reduction of the dimensionality by the ties in \mathbf{j} and $a_n/a'_n \leq \tilde{M}$.

Since the smoothness condition for (3.9) holds, the first part of (4.17) holds if (3.6) and (3.7) hold for $t_n/\log n \rightarrow \infty$ and $f_n = f_{0,s} \tilde{h}_{n,k,j}$ for all $s = 0, 1, 2$ and k, j , where $f_{0,s}$ are given by (4.11). The decomposition $f_{n,k,j} = \sum_{s=0}^2 f_{0,s} \tilde{h}_{n,k,j}$ allows us to consider components of different orders of degeneracy [cf. the index $r - 1$ in (4.22) below]. Let $f_n = f_{0,s} \tilde{h}_{n,k,j}$ for a fixed (s, k, j) and consider $\|b - \beta\| \leq Ma_n$. Since $f = F'$ is bounded and $\tilde{K}^{[i]}$ in (4.21) have bounded support, $|(F(T_{1,\beta}) - F(T_{2,\beta})) \tilde{K}_{n12}^{[2]}|/a_n \leq \tilde{M} |\tilde{K}_{n12}^{[2]}|(a_{n,2} = a_n)$ and $|(F(T_{4,\beta}) - F(T_{1,\beta})) \tilde{K}_{n14}^{[4]}|/a'_n \leq \tilde{M} |\tilde{K}_{n14}^{[4]}|$, so that $|f_{0,s}|/(a'_n a_n) \leq \tilde{M}/a_n^s$ on the set $f_n \neq 0$ by (4.11). Thus, by (4.20) and (4.21), $\|f_n\|$ is bounded by a sum of $\tilde{M} \tilde{v}_{n,j} \|\phi_n^{(k)}\|_\infty a_n^{-s} E_{i_1,\dots,i_l}^* \prod_{i=2}^m |\tilde{K}_{n1i}^{[i]}|$. Since $\|\phi_n^{(k)}\|_\infty \leq \tilde{M} c_n^{-k-1}$ and $\tilde{v}_{n,j} \leq \tilde{M}/(na_n)^d$, this and Lemma 3.2 imply (3.6) with

$$(4.22) \quad c_{n,l} = \frac{\tilde{M} c_n^{-k-1} I\{l \geq r\}}{(na_n)^d a_n^{s+(l-1)^+}}, \quad \sigma_{n,l} = \frac{\tilde{M} c_n^{-2(k+1)} I\{l \geq r\}}{(na_n)^{2d} a_n^{2s+(l-1)^+}}, \quad l \geq 0,$$

where $d = 4k + 5 - j$ and r is the (largest) integer such that $E \|E_{i_1,\dots,i_l}^* f_n\| = 0$ for all $l < r$ ($r = 0$ if $E f_n \neq 0$ and $r \geq 1$ if $E f_n = 0$). For $j = 4k + 5$, $r = s + k - 1$ due to the centering of \tilde{h}_n at $\bar{D}_{n,i}$ in (4.19) and δ_i at $E[\delta_i | T_i, Z_i] = F(T_{i,\beta})$ in (4.11). For $j < 4k + 5$, ties temper some centering of \tilde{h}_n , so that $r = s + (k - 1 - 2d)^+$. Thus, by (4.22), for $l \geq 1$ and $k \geq 1$,

$$(4.23) \quad c_{n,l}^2 \left(\frac{\log n}{n} \right)^{l+1} \leq \frac{\tilde{M} c_n^{-2(k+1)} (\log n)^{r+1}}{(na_n)^{2d} a_n^{2(s+r-1)} n^{r+1}} \leq \tilde{M} c_n^2 \left(\frac{\log n}{na_n^2 c_n^2} \right)^{k+2} \rightarrow 0,$$

as $l = r$, $d = 0$ (i.e., $r = k + s - 1$) and $s = 2$ are the worst cases due to $\log n/(na_n^2) \rightarrow 0$, $a_n/\log n \rightarrow 0$ and $\log n/(na_n^4) \rightarrow \infty$, respectively. Similarly,

for all $l \geq 1$,

$$(4.24) \quad \sigma_{n,l}^2 \left(\frac{\log n}{n} \right)^l \leq \tilde{M} c_n^{-4} \alpha_n^k \left(\frac{\log n}{n \alpha_n^2 c_n^2} \right)^{k-1} \left(\frac{\log n}{n \alpha_n^3} \right)^s \left(\frac{\log n}{n \alpha_n} \right)^{(l-r)^+},$$

as $d = 0$ is the worst case. For $k = 1$ and $\tilde{h}'_n = (\partial/\partial b)\tilde{h}_n(\mathbf{X}_{6,9}, T_{1,b}; b)$, (4.19) is a single sum and (4.21) can be replaced by $\phi_n^{(1)}(\bar{D}_n(T_{1,b}; b))(\alpha'_n)^{-1}\tilde{h}'_n(\mathbf{Z}_1 - \mathbf{Z}_3)^{\text{tr}} \prod_{i=2}^5 \tilde{K}_{n1i}^{[i]}$. Since \tilde{h}'_n has differentiated kernels (with $\int \tilde{K}^{[i_0]}(t) dt = 0$), $\|E_{1,\dots,5}^* \tilde{h}'_n\|_\infty \leq \tilde{M} \alpha_n^{\alpha-1}$ (instead of \tilde{M}/α_n) as in the proof of (3.16), so that for $k = 1$ and $l = r$ (consequently $d = 0$ and $r = s$) we can modify (4.22) and (4.24), with an extra factor $\alpha_n^{2\alpha}$, into

$$(4.25) \quad \sigma_{n,l}^2 \left(\frac{\log n}{n} \right)^l \leq \tilde{M} c_n^{-4} \alpha_n^{2\alpha+s \wedge 1} \left(\frac{\log n}{n \alpha_n^3} \right)^s \rightarrow 0$$

for $k = 1$ and $0 \leq l = r = s \leq 2$, as (4.11) implies $E_{i_1, \dots, i_s}^* f_n \neq 0$ only if $\{i_1, \dots, i_s\} \subseteq \{1, \dots, 5\}$. The bounds in (4.24) are $o(1)$ for $k \geq 2$ and $l \geq 1$ and for $k = 1$ and $l > r$. Thus, (4.23)–(4.25) imply (3.7) with $t_n/\log n \rightarrow \infty$ and by Lemma 3.1 the first part of (4.17) holds for all $k \geq 1$.

The proof of the second part of (4.17) is simpler. By (4.18), (4.9) and (2.14)–(2.16),

$$(4.26) \quad \begin{aligned} \|\tilde{R}(b)\| &\leq \tilde{M} \sum_{(\mathbf{i}, \mathbf{j}) \in I_9^n} (\hat{D}_{n,\mathbf{i}} - \bar{D}_{n,\mathbf{i}})^{k_0} \\ &\times \frac{\|\phi_n^{(k_0+1)}\|_\infty \|\tilde{h}_{n,0}(\mathbf{X}_\mathbf{i}; b)\|}{\alpha'_n \binom{n}{9} 9!} \left\| \frac{\partial}{\partial b} \tilde{h}_n(\mathbf{X}_\mathbf{j}, T_{i_1,b}; b) \right\| \end{aligned}$$

for the even integer k_0 in (4.2). By Lemma 3.2 and the same upper bounds for $\|f_n\|$, (4.22) holds with $k = k_0 + 1$ and $r = (k_0 - 2d)^+$ in the worst case $s = 2$. Thus, $c_{n,l}^2 (\log n/n)^{l+1} \leq \tilde{M} c_n^{-2(k_0+2)} \{\log n/(n \alpha_n^2)\}^{k_0+1} \rightarrow 0$ by (4.2) for $l \geq 1$ and

$$(4.27) \quad \sigma_{n,l}^2 \left(\frac{\log n}{n} \right)^l \leq \tilde{M} \alpha_n^{2+l \wedge 1} c_n^{-4} \left(\frac{\log n}{n \alpha_n^2 c_n^2} \right)^{k_0} \rightarrow 0, \quad l \geq 0$$

as $l = r$, $d = 0$ (i.e., $l = k_0$), $s = 2$ and $k_0 = 6$ are the worst cases. This proves the second part of (4.17) by Lemma 3.1. Thus, (4.16) holds.

The proof for the second sum in (4.15) is almost identical. For the terms in the Taylor expansion with $\phi_n^{(k)}(\bar{D}_{n,\mathbf{i}})(\hat{D}_{n,\mathbf{i}} - \bar{D}_{n,\mathbf{i}})^k$, $0 \leq k \leq 5$, (4.22) holds with $r = s + (k - 2d)^+$, so that $c_{n,l}^2 (\log n/n)^{l+1} \leq \tilde{M} c_n^{-2} (\log n)^3 / (n \alpha_n^2)^3 \rightarrow 0$ for $l \geq 1$ and $\sigma_{n,l}^2 (\log n/n)^l \leq \tilde{M} \alpha_n c_n^{-2} (\log n / (n \alpha_n^3))^s \rightarrow 0$, $l \geq 1$, as $l = r$, $d = 0$ (i.e., $r = k + s$) and $k = 0$, and then $s = 2$ for $c_{n,l}^2 (\log n/n)^{l+1}$, are the worst cases. The calculation for the remainder is straightforward with $k = 6$ and $r = (6 - 2d)^+$, easier than $k = k_0 + 1$ for (4.27).

STEP 2. *Proof of (4.8).* We first compute $\sigma_{n,0}^2 \geq (Ef_n)^2$ as in 4.1 to remove terms of smaller order. Since $Ef_n = 0$ for $r > 0$, we only need to consider $l = r = 0$. For $E\tilde{\xi}_{n,k}(b)$ in (4.17) $r = s + (k - 1 - 2d)^+ = 0$ implies $s = 0$ and $2d \geq k - 1$, so that $\sigma_{n,0}^2 = o(1)$ in (4.22) for $k \geq 2$ (i.e. $d \geq 1$) and in (4.25) for $k = 1$. For $E\tilde{R}_n(b)$ in (4.17), $\sigma_{n,0}^2 = o(1)$ in (4.27). Thus, $E\tilde{\xi}_n(b) = o(1)$ in (4.16). For the second sum in (4.15), $\sigma_{n,0}^2 = o(1)$ in (4.22) for the terms with $1 \leq k \leq 5$ (with $l = r = s + (k - 2d)^+ = 0$) and the remainder (with $k = 6$, $l = r = (6 - 2d)^+ = 0$, $s = 2$). Thus, for $\|b - \beta\| \leq Ma_n$ the mean converges to zero for all terms other than those with $k = 0$ for the second sum in (4.15). This and (4.12) imply $E \partial \hat{\xi}_n(b) / \partial b = E(\tilde{\phi}_n(b) / a'_n) (\partial \tilde{h}_{n,0}^{\text{tr}}(b) / \partial b) + o(1) = I_* + o(1)$.

STEP 3. *Proof of (4.5).* As in the proof of (4.7), it suffices to show

$$(4.28) \quad \sqrt{n} \|\xi_{n,k}(\beta)\| = o_P(1), \quad 1 \leq k \leq 5, \sqrt{n} E\|R_n\| = o(1),$$

$$(4.29) \quad \xi_{n,0}(\beta) = n^{-1} \sum_{i=1}^n \rho_*(X_i; \beta) + \frac{1}{\binom{n}{2} 2!} \sum_{\mathbf{i} \in I_2^n} \tilde{\rho}_n(\mathbf{X}_i) + o_P(n^{-1/2}),$$

where $R_n = \hat{\xi}_n(\beta) - \sum_{k=0}^5 \xi_{n,k}(\beta)$, and

$$\xi_{n,k}(b) = \sum_{\mathbf{i} \in I_5^n} (\hat{D}_{n,i} - \bar{D}_{n,i})^k \phi_n^{(k)}(\bar{D}_{n,i}) \frac{\tilde{h}_{n,0}(\mathbf{X}_i; b)}{a'_n \binom{n}{5} 5! k!}.$$

We shall first prove (4.28). Let $k \geq 1$. As in (4.20) and (4.21),

$$\begin{aligned} \sqrt{n} \xi_{n,k}(\beta) &= \sum_j \sum_{\mathbf{i} \in I_j^n} \frac{f_{n,k,j}(\mathbf{X}_i)}{\binom{n}{j} j!}, \\ f_{n,k,j}(\mathbf{X}_{1,j}) &= \sum_{s=0}^2 f_{0,s}(\beta) h_{n,k,j}, \quad k \geq 1, \end{aligned}$$

where $h_{n,k,j}$ is a sum of terms of the form of the modified (4.21) with $(Z_{i_0} - Z_1) / \{a_{n,i_0}(k - 1)!\}$ replaced by $\sqrt{n} / k!$. This means an extra factor of $a_n \sqrt{n}$ in the computation. Let (s, k, j) be fixed and $f_n = f_{0,s}(\beta) h_{n,k,j}$. As in (4.22) and (4.24), (3.6) holds with

$$(4.30) \quad \begin{aligned} \frac{\sigma_{n,l}^2}{n^l} &\leq \frac{\tilde{M} n a_n^2 c_n^{-2(k+1)} I\{l \geq r\}}{(n a_n)^{2d} a_n^{2s+(l-1)^+} n^l} \\ &\leq \frac{\tilde{M} c_n^{-4} a_n^{k+l \wedge 1} I\{l \geq s\}}{(n a_n^3)^s (n a_n)^{(l-r)^+}} \left(\frac{c_n^{-2}}{n a_n^2}\right)^{k-1} = o(1), \end{aligned}$$

for all $l \geq 1$ and $k \geq 1$ and for $l = 0$ and $k \geq 2$, where $r = s + (k - 2d)^+$. For $k = 1$ there is no tie, so that $r = k + s \geq 1$ and $\sigma_{n,0}^2 = 0$. For $E\|R_n\|$, we use the upper bound as in (4.26) and then take $l = 0$, $k = 6$, $r = (k - 2d)^+ = 0$

and $s = 2$ in the first inequality of (4.30) to yield $\sigma_{n,0}^2 \leq \tilde{M}c_n^{-14}/(n^5a_n^8) = (a_n/c_n^2)^7/(na_n^3)^5 \rightarrow 0$. This proves (4.28).

For (4.29), we have $k = d = 0$, $r = s$ and $\sqrt{n} \xi_{n,0}(\beta) = \sum_s \sum_{\mathbf{i} \in I_s^g} f_{n,0,s}(\mathbf{X}_i) / \binom{n}{s}!$ where $f_{n,0,s} = f_{n,0,s}(\mathbf{X}_{1,5})$ is as in Lemma 4.1. By (4.30) $\sigma_{n,l}^2/n^l = o(1)$ for $f_n = f_{n,0,s}$ if $l > r = s$. The computation for $l = r = s$ is given in Lemma 4.1. These imply (4.29) by the Hoeffding decomposition (3.10). Hence, (4.5) holds.

4.3. *Proof of Lemma 4.1.* Consider $\|\eta\| = \|b - \beta\| \leq M/\sqrt{n} = o(a_n)$. By (4.9),

$$(4.31) \quad \begin{aligned} & E(\tilde{\phi}_n(b)/a'_n)(\partial \tilde{h}_{n,0}^{\text{tr}}(b)/\partial b) \\ &= E\psi_n(b)(\partial h_{n,0}^{\text{tr}}(b)/\partial b) + E\tilde{\psi}_n(b)h_{n,0}^{\text{tr}}(b), \end{aligned}$$

where $h_{n,0}$ is the factor of h_n in (3.11), $\psi_n(b) = \tilde{\phi}_n(b)E_1^*\{F(T_{4,\beta}) - F(T_\beta)\}\tilde{K}_{n145}(b)/a'_n$ and $\tilde{\psi}_n(b) = \tilde{\phi}_n(b)E_1^*\{F(T_{4,\beta}) - F(T_\beta)\}(\partial/\partial b)\tilde{K}_{n145}(b)/a'_n$.

Consider $\psi_n(b)$. The argument of (3.23) can be modified to show

$$(4.32) \quad \begin{aligned} & E_1^*\{F(T_{4,\beta}) - F(T_\beta)\}\tilde{K}_{n145}(b)/a'_n \\ &= g_b^2 E_1^*\{(F_{4,b} - F(T_\beta)) * \tilde{K}'_n\}_b/a'_n + \dots \\ &= g_b^2 (F * \tilde{K}'_n)_\beta/a'_n + \dots \\ &= g_\beta^2 (f * \tilde{K}_n)_\beta + \tilde{M}_\infty\{(a'_n)^\alpha + \|\eta\|/a'_n\}. \end{aligned}$$

Let $D_\beta = D(T_\beta; \beta)$ as in (2.20). By (4.10), (2.16), (2.6) and $\|f\|_\infty < \infty$, $\|\bar{D}_n(T_b; b) - D_\beta\|_\infty \leq \tilde{M}a_n^\alpha$ as in (3.17)–(3.19). This and (4.32) imply

$$(4.33) \quad \tilde{\phi}_n(b) = \frac{1 + \tilde{M}_\infty(a'_n)^\alpha/c_n}{c_n + D_\beta}, \quad \|\psi_n(b) - \tilde{\phi}_n(b)g_\beta^2(f * \tilde{K}_n)_\beta\|_\infty = o(1)$$

as $(a'_n)^\alpha/c_n + 1/(\sqrt{n}a_n c_n) \rightarrow 0$. Thus, in view of the proof of (3.15) [especially (3.23)],

$$\begin{aligned} E\psi_n(b) \frac{\partial}{\partial b} h_{n,0}^{\text{tr}}(b) &= a_n^{-1} E\psi_n(b) v_b g_b^2 (F_{2,b} * K'_n)_b + o(1) \\ &= E \frac{v_\beta g_\beta^4 (f * \tilde{K}_n)_\beta (f * K_n)_\beta}{c_n + D_\beta} \left\{ 1 + \frac{\tilde{M}_\infty(a'_n)^\alpha}{c_n} \right\} + o(1) \\ &= I_* + o(1) \end{aligned}$$

by (2.19) and (4.33). Simpler computation provides $\|\tilde{\psi}_n\|_\infty \leq \tilde{M}/(a'_n c_n)$ and $E\tilde{\psi}_n(b)h_{n,0}^{\text{tr}}(b) = O(a_n^{1+\alpha}/(a'_n c_n)) = o(1)$ by (3.13). These give (4.12) via (4.31).

For (4.13), it suffices to consider $(i_1, i_2) = (1, 4)$ and $(2, 4)$, due to the centering of δ_i in (4.11) and (4.1), and $E_{i,4}^* f_{n,0,2} = (-1)^{i+1}(\sqrt{n}/a'_n)\{\delta_i - F(T_{i,\beta})\}\{\delta_4 - F(T_{4,\beta})\}E_{i,4}^* h_n^*$ for $i = 1, 2$. By (3.17), (3.19) and Lemma 3.2,

$E_{1,4}^* h_n^* = \tilde{\phi}_n(\beta)(Z - \mu_\beta)g_\beta^3 \tilde{K}'_{n14} + \tilde{M}_2(a'_n)^{\alpha-1/2}/c_n$. Let $\|\cdot\|_1 = \text{trace}(\cdot)$. With (4.33), (2.20) and (4.1), these imply

$$E\|E_{1,4}^*\{f_{n,0,2} - \sqrt{n}\tilde{\rho}_n(\mathbf{X}_{(1,4)})\}\|^2/n^2 \leq \tilde{M}\{(a'_n)^\alpha/c_n\}^2\|\sigma_n^2\|_1 + \tilde{M}c_n^{-2}(a'_n)^{2\alpha-3}/n \rightarrow 0$$

as in (4.6). Also, since $\|E(Z_1 - Z_3)K_{n13}(\beta)|T_\beta\|_\infty = \tilde{M}a_n^\alpha$, $E(E_{2,4}^* f_{n,0,2})^2/n^2$ is of the order $c_n^{-2}a_n^{2\alpha}/\{n(a'_n)^3\} = o(1)$. Therefore (4.13) holds due to $E_{2,4}^* \tilde{\rho}_n(\mathbf{X}_{(1,4)}) = 0$.

For (4.14), we only need to consider $i = 1, 2, 4$. As in the proof of (3.14), (4.33) and (3.17)–(3.21) imply $E_1^* f_{n,0,1}/\sqrt{n} = \psi_n(\beta)E_1^* h_{n,0} = \psi_n(\beta)\rho_1(X_1, \beta) + \tilde{M}_\infty a_n^\alpha/c_n$ with $E_1^* \|\rho_1(X_1, \beta)\|^2 = \tilde{M}_\infty \|v_\beta\|_1 D_\beta$ in view of (2.8). This and (4.33) and (2.19) imply

$$\begin{aligned} n^{-1}E\|E_1^*\{f_{n,0,1} - \sqrt{n}\rho_*(X_1; \beta)\}\|^2 &\leq \tilde{M}E|\psi_n(\beta) - \psi_*(T_\beta)|^2\|v_\beta\|_1 \frac{D_\beta}{\mu_2} + o(1) \\ &\leq \tilde{M}E\left|\frac{(f^* \tilde{K}_n)_\beta - \mu_2 f(T_\beta)}{c_n + D_\beta} g_\beta^2 - \frac{\mu_2 f(T_\beta) g_\beta^2 c_n}{\{c_n + D_\beta\} D_\beta}\right|^2 \|v_\beta\|_1 \frac{D_\beta}{\mu_2} \\ &\quad + o(1)\|I_*\|_1 + o(1) = o(1) \end{aligned}$$

as $\rho_* = \psi_* \rho_1$, $\psi_* = \mu_2 f(T_\beta) g_\beta^2/D_\beta$ and $I_* < \infty$. For $i = 2, 4$, $\|E[(Z_1 - Z_3)K_{n13}(\beta)|T_\beta]\|_\infty = \tilde{M}a_n^\alpha$ implies

$$E\|E_i^* f_{n,0,1}\|^2/n \leq \tilde{M}c_n^{-2}a_n^{2\alpha}\|f\|_\infty E(E_i^* |K_{n12} \tilde{K}'_{n145})^2 = o(1),$$

so that (4.14) follows. Finally, by (3.13), (4.31), (4.32) and (4.33), we find $E f_{n,0,0} = E\sqrt{n} h_{n,0} \psi_n(\beta) = \sqrt{n} a_n^{1+\alpha}/c_n = o(1)$. \square

Acknowledgments. The authors thank the Associate Editor and a referee for pointing out references in the econometrics literature and for their suggestions to fully explore the methods in this paper.

REFERENCES

ARCONES, M. A. and GINÉ, E. (1993). Limit theorems for U -processes. *Ann. Probab.* **21** 1494–1542.
 AYER, M., BRUNK, H. D., EWING, G. M., REID, W. T. and SILVERMAN, E. (1955). An empirical distribution function for sampling with incomplete information. *Ann. Math. Statist.* **26** 641–647.
 BICKEL, P. J., KLAASSEN, C. A. J., RITOV, Y. and WELLNER, J. A. (1993). Efficient and adaptive estimation for semiparametric models. Johns Hopkins Univ. Press, Baltimore.
 BRUNK, H. D. (1970). Estimation of isotonic regression. In *Nonparametric Techniques in Statistical Inference* (M. L. Puri, ed.) 177–195. Cambridge Univ. Press.
 COSSLETT, S. R. (1987). Efficiency bounds for distribution-free estimator of the binary choice and the censored regression models. *Econometrica* **55** 559–585.

- FINKELSTEIN, D. (1986). A proportional hazards model for interval-censored failure time data. *Biometrics* **42** 845–854.
- FINKELSTEIN, D. and WOLFE, R. A. (1986). Isotonic regression for interval-censored survival data using an E-M algorithm. *Comm. Statist. Theory Methods* **15** 2493–2505.
- GROENEBOOM, P. and WELLNER, J. A. (1992). *Information Bounds and Nonparametric Maximum Likelihood Estimation*. Birkhäuser, Boston.
- HAN, A. K. (1987). Non-parametric analysis of a generalized regression model. *J. Econometrics* **35** 303–316.
- HUANG, J. (1996). Efficient estimation for the proportional hazards model with interval censoring. *Ann. Statist.* **24** 540–568.
- HUANG, J. and WELLNER, J. A. (1996). Interval censored survival data: a review of recent progress. In *Proceedings of the First Seattle Symposium on Biostatistics: Survival Analysis* (D. Lin and T. Fleming, eds.) 123–170. Springer, New York.
- JEWELL, N. P. and SHIBOSKI, S. C. (1990). Statistical analysis of HIV infectivity based on partner studies. *Biometrics* **46** 1133–1150.
- KLEIN, R. W. and SPADY, R. H. (1993). An efficient semiparametric estimator for binary response models. *Econometrica* **61** 387–421.
- LIN, D. Y. and YING, Z. (1996). Additive hazards regression with current status data. Preprint.
- MANSKI, C. F. (1975). Maximum score estimation of the stochastic utility model of choice. *J. Econometrics* **3** 205–228.
- MANSKI, C. F. (1985). Semiparametric analysis of discrete response: asymptotic properties of maximum score estimator. *J. Econometrics* **27** 313–333.
- PETO, R., PIKE, M. C., DAY, N. E., GRAY, R. G., LEE, P. N., PARISH, S., PETO, J., RICHARDS, S. and WAHRENDORF, J. (1980). Guidelines for simple sensitive significance tests for carcinogenic effects in long-term animal experiments. In *Evaluation of Carcinogenic Risk of Chemicals to Humans, Suppl. 2. Long-Term and Short-Term Screening Assays for Carcinogens: A Critical Appraisal* 311–85. International Agency for Research on Cancer, Lyon.
- RABINOWITZ, D., TSIATIS, A. and ARAGON, J. (1995). Regression with interval censored data. *Biometrika* **82** 501–513.
- SHERMAN, R. P. (1993). The limiting distribution of the maximum rank correlation estimator. *J. Econometrics* **61** 123–137.
- SHIBOSKI, S. C. and JEWELL, N. P. (1992). Statistical analysis of the time dependence of HIV infectivity based on partner study data. *J. Amer. Statist. Assoc.* **87** 360–372.

ACER / EXCEL, INC.
KENDLE
6 COMMERCE DRIVE
CRANFORD, NEW JERSEY 07016

RUTGERS UNIVERSITY
DEPARTMENT OF STATISTICS
HILL CENTER
BUSCH CAMPUS
PISCATAWAY, NEW JERSEY 08854
E-MAIL: czhang@stat.rutgers.edu