CURRENT STATUS LINEAR REGRESSION

(SUPPLEMENTARY MATERIAL)

By Piet Groeneboom and Kim Hendrickx

Delft University of Technology and Hasselt University

We give the proofs of the results stated in Sections 3 and 4 of the manuscript. Entropy results are used in our proofs. Before we prove the results we first give some definitions and an equicontinuity lemma needed in the proofs.

Consider a class of functions $\mathcal{F}$ on $\mathcal{R}$ and let $L_2(Q)$ be the L_2-norm defined by a probability measure Q on $\mathcal{R}$, i.e. for $g \in \mathcal{F}$,

$$\|g\|_{L_2} = \int g^2 dQ.$$

For any probability measure Q on $\mathcal{R}$ let $N_B(\zeta, \mathcal{F}, L_2(Q))$ be the minimal number N for which there exists pairs of functions $\{[g_j^L, g_j^U], j = 1, \ldots, N\}$ such that $\|g_j^U - g_j^L\|_{L_2} \le \zeta$ for all $j = 1, \ldots, N$ and such that for each $g \in \mathcal{F}$ there is a $j \in \{1, \ldots, N\}$ such that $g_j^L \le g \le g_j^U$. The $\zeta-$entropy with bracketing of $\mathcal{F}$ (for the $L_2(Q)-$distance) is defined as $H_B(\zeta, \mathcal{F}, L_2(Q)) = \log(N_B(\zeta, \mathcal{F}, L_2(Q)))$.

Lemma S0.1 (Equicontinuity Lemma, Theorem 5.12, p.77 in [4]). *Let $\mathcal{F}$ be a fixed class of functions with envelope F in $L_2(P) = \{f : \int f^2 dP < \infty\}$. Suppose that*

$$\int_0^1 H_B^{1/2}(u, \mathcal{F}, L_2(P))\, du \le \infty,$$

where H_B is the entropy with bracketing of $\mathcal{F}$ for the L_2-norm. Then, for all $\eta > 0$ there exists a $\delta > 0$ such that

$$\limsup_{n\to\infty} P\left(sup_{[\delta]}|\sqrt{n}\int (f-g)d(\mathbb{P}_n - P_0)| > \eta\right) < \eta,$$

where,

$$[\delta] = \{(f,g) : \|f - g\| \le \delta\}.$$

S1. Behavior of the maximum likelihood estimator. In this section we prove Lemma 3.1. We first prove in Lemma S1.1 some entropy bounds needed in the proofs.

Lemma S1.1. *Let*

$$\mathcal{F} = \{(t,x) \mapsto F(t - \beta' x) : F \in \mathcal{F}_0, \beta \in \Theta\},$$

where $\mathcal{F}_0$ is the set of subdistribution functions on $[a,b]$, where $[a,b]$ contains all values $t - \beta' x$ for $\beta \in \Theta$, and (t,x) in the compact neighborhood over which we let them vary. Then,

$$\sup_{\varepsilon > 0} \varepsilon H_B(\varepsilon, \mathcal{F}, L_2(P_0)) = O(1), \tag{S1.1}$$

Furthermore, let

$$\mathcal{G} = \{(t,x) \mapsto g(t-\beta' x) : g \in \mathcal{G}_0, \beta \in \Theta\},$$

where $\mathcal{G}_0$ is a set functions of uniformly bounded variation, then

$$\sup_{\varepsilon>0} \varepsilon H_B\left(\varepsilon, \mathcal{G}, L_2(P_0)\right) = O(1). \tag{S1.2}$$

Proof. We only prove the result for the class $\mathcal{F}$ since the proof for the class $\mathcal{G}$ can be obtained similarly.

Fix $\varepsilon > 0$. We first note that Θ can be covered by N neighborhoods with diameter at most ε^2 where N is of order ε^{-2d}. Let $\{\beta_1,\dots,\beta_N\}$ denote elements of each of these neighborhoods. Consider an ε-bracket $[F_j^L, F_j^U]$, $j=1,\dots,N'$ covering the class $\mathcal{F}_0$ such that,

$$\left\{\int \{F_j^U(u) - F_j^L(u)\}^2 f_{T-\beta' X}(u)\,du\right\}^{1/2} < \varepsilon.$$

for $j=1,\dots,N'$. The existence of such an ε-net is assured by the fact that $f_{T-\beta' X}$ is bounded above (uniformly in β). The number N' is of order $\exp(C/\varepsilon)$ for some constant C (See e.g. [4], p. 18). Let β_j be chosen such that $\|\beta_j - \beta\| < \varepsilon^2$, where $\|\cdot\|$ denotes the Euclidean norm. Then:

$$t - \beta_j' x - \varepsilon^2 R \le t - \beta' x = t - \beta_j' x - (\beta-\beta_j)' x \le t - \beta_j' x + \varepsilon^2 R,$$

where R is the maximum of the values $\|x\|$. This implies that for each $F \in \mathcal{F}_0$ and $\beta \in \Theta$,

$$F_i^L(t - \beta_j' x - \varepsilon^2 R) \le F(t-\beta' x) \le F_i^U(t - \beta_j' x + \varepsilon^2 R),$$

for some $i = 1,\dots,N'$ and $j=1,\dots,N$. The result of Lemma S1.1 follows if we can show that,

$$\left\{\int \{F_i^U(t - \beta_j' x + \varepsilon^2 R) - F_i^L(t - \beta_j' x - \varepsilon^2 R)\}^2 dG(t,x)\right\}^{1/2} \lesssim \varepsilon. \tag{S1.3}$$

By the triangle inequality we get that the left-hand side of the above equation is bounded by:

$$\begin{aligned}
&\left\{\int \{F(t - \beta_j' x - \varepsilon^2 R) - F_i^L(t - \beta_j' x - \varepsilon^2 R)\}^2\, dG(t,x)\right\}^{1/2} \\
&\quad + \left\{\int \{F_i^U(t - \beta_j' x + \varepsilon^2 R) - F(t - \beta_j' x + \varepsilon^2 R)\}^2\, dG(t,x)\right\}^{1/2} \\
&\quad + \left\{\int \{F(t - \beta_j' x + \varepsilon^2 R) - F(t - \beta_j' x - \varepsilon^2 R)\}^2\, dG(t,x)\right\}^{1/2} \\
&\lesssim \varepsilon + \left\{\int \{F(u + \varepsilon^2 R) - F(u - \varepsilon^2 R)\}^2 f_{T-\beta_j' x}(u)\right\}^{1/2}.
\end{aligned}$$

Let $u_0 = a - \varepsilon^2 R < u_1, \dots < u_m = b + \varepsilon^2 R$, be points such that $u_k - u_{k-1} = \varepsilon^2$, $k = 1,\dots,m-1$,

$u_m - u_{m-1} \le \varepsilon^2$. Then:

$$\begin{aligned}
&\int \{F(u+\varepsilon^2 R) - F(u-\varepsilon^2 R)\}^2 f_{T-\beta_j X}(u)\,du\\
&\le \int \{F(u+\varepsilon^2 R) - F(u-\varepsilon^2 R)\} f_{T-\beta_j X}(u)\,du \le M\int \{F(u+\varepsilon^2 R) - F(u-\varepsilon^2 R)\}\,du\\
&= M\int_{a+\varepsilon^2 R}^{b+\varepsilon^2 R} F(u)\,du - M\int_{a-\varepsilon^2 R}^{b-\varepsilon^2 R} F(u)\,du\\
&\le M\int_{a-\varepsilon^2 R}^{a+\varepsilon^2 R} F(u)\,du + M\int_{b-\varepsilon^2 R}^{b+\varepsilon^2 R} F(u)\,du \lesssim \varepsilon^2,
\end{aligned}$$

where M is an upper bound for $f_{T-\beta_j' X}$, and where we extend the function F by a constant value outside $[a,b]$. This completes the proof of (S1.3) since we have shown that there exist positive constants A_1, A_2, A_3 and C such that

$$\begin{aligned}
H_B\left(\varepsilon, \mathcal{F}, L_2(P_0)\right) &\le \log N + \log N' \le d\log(A_1/\varepsilon^2) + A_2\log(\exp(C/\varepsilon)) \le A_3/\varepsilon\\
&= O(\log(1/\varepsilon)) + O(1/\varepsilon) = O(1/\varepsilon), \qquad \varepsilon \downarrow 0.
\end{aligned}$$

□

Proof of Lemma 3.1. Let h denote the Hellinger distance on the class of densities $\mathcal{P}$ defined by

$$\mathcal{P} = \left\{p_{\beta,F}(t,x,\delta) = \delta F(t-\beta' x) + (1-\delta)\{1-F(t-\beta' x)\} : F \in \mathcal{F}_0, \beta \in \Theta\right\},$$

w.r.t. the product of counting measure on $\{0,1\}$ and the measure dG of (T,X), where $\mathcal{F}_0$ is the class of right-continuous subdistribution functions.

We have (see, e.g., the "basic inequality" Lemma 4.5, p. 51 of [4]):

$$h^2\left(p_{\beta,\hat F_{n,\beta}}, p_{\beta,F_\beta}\right) \le \int \frac{2p_{\beta,\hat F_{n,\beta}}}{p_{\beta,\hat F_{n,\beta}} + p_{\beta,F_\beta}}\, d\left(\mathbb{P}_n - P_0\right),$$

where we use the convexity of the set of densities of this type for (temporarily) fixed β. Hence we get, for $\epsilon \in (0,1]$:

$$\begin{aligned}
&\mathbb{P}\left\{\sup_{\beta\in\Theta} h\left(p_{\beta,\hat F_{n,\beta}}, p_{\beta,F_\beta}\right) \ge \epsilon\right\}\\
&= \mathbb{P}\left\{\sup_{\beta\in\Theta,\, h\left(p_{\beta,\hat F_{n,\beta}}, p_{\beta,F_\beta}\right)\ge\epsilon} \left\{\int\left\{\frac{2p_{\beta,\hat F_{n,\beta}}}{p_{\beta,\hat F_{n,\beta}} + p_{\beta,F_\beta}} - 1\right\} d\left(\mathbb{P}_n - P_0\right) - h^2\left(p_{\beta,\hat F_{n,\beta}}, p_{\beta,F_\beta}\right)\right\} \ge 0\right.\\
&\qquad\qquad \left., \sup_{\beta\in\Theta} h\left(p_{\beta,\hat F_{n,\beta}}, p_{\beta,F_\beta}\right) \ge \epsilon\right\}\\
&\le \mathbb{P}\left\{\sup_{\beta\in\Theta,\, F\in\mathcal{F}_0,\, h\left(p_{\beta,F}, p_{\beta,F_\beta}\right)\ge\epsilon} \left\{\int\left\{\frac{2p_{\beta,F}}{p_{\beta,F} + p_{\beta,F_\beta}} - 1\right\} d\left(\mathbb{P}_n - P_0\right) - h^2\left(p_{\beta,F}, p_{\beta,F_\beta}\right)\right\} \ge 0\right\}
\end{aligned}$$

$$\leq \sum_{s=0}^{\infty} \mathbb{P}\left\{\sup_{\substack{\beta\in\Theta,\, F\in\mathcal{F}_0,\\ 2^{2s}\epsilon\leq h\left(p_{\beta,F},\,p_{\beta,F_\beta}\right)\leq 2^{s+1}\epsilon}} \sqrt{n}\int\left\{\frac{2p_{\beta,F}}{p_{\beta,F}+p_{\beta,F_\beta}}-1\right\}d\left(\mathbb{P}_n-P_0\right)\geq \sqrt{n}\,2^{2s}\epsilon^2\right\},$$

We can now use Theorem 5.13 in [4], taking $\epsilon = Mn^{-1/3}, \alpha = 1, \beta = 0$ and $T = \sqrt{n}\,2^{2s}\epsilon^2 = M2^{2s}n^{-1/6}$, together with Lemma S1.1 for the entropy of the set of densities to conclude:

$$\sum_{s=0}^{\infty} \mathbb{P}\left\{\sup_{\substack{\beta\in\Theta,\, F\in\mathcal{F}_0,\\ 2^{2s}\epsilon\leq h\left(p_{\beta,F},\,p_{\beta,F_\beta}\right)\leq 2^{s+1}\epsilon}} \sqrt{n}\int\left\{\frac{2p_{\beta,F}}{p_{\beta,F}+p_{\beta,F_\beta}}-1\right\}d\left(\mathbb{P}_n-P_0\right)\geq \sqrt{n}\,2^{2s}\epsilon^2\right\}$$
$$\leq \sum_{s=0}^{\infty} c_1\exp(-c_2M2^{2s})$$

for constants $c_1, c_2 > 0$. Since the sum can be made arbitrarily small for M sufficiently large, we find:

$$\sup_{\beta\in\Theta} h\left(p_{\beta,\hat{F}_{n,\beta}}, p_{\beta,F_\beta}\right) = O_p\left(n^{-1/3}\right).$$

We have:

$$\begin{aligned}
&h\left(p_{\beta,\hat{F}_{n,\beta}}, p_{\beta,F_\beta}\right)^2\\
&= \tfrac12\int\left\{p^{1/2}_{\beta,\hat{F}_{n,\beta}}(t,x,1)-p^{1/2}_{\beta,F_\beta}(t,x,1)\right\}^2 dG(t,x) + \tfrac12\int\left\{p^{1/2}_{\beta,\hat{F}_{n,\beta}}(t,x,0)-p^{1/2}_{\beta,F_\beta}(t,x,0)\right\}^2 dG(t,x)\\
&= \tfrac12\int\left\{\hat{F}_{n,\beta}(t-\beta'x)^{1/2}-F_\beta(t-\beta'x)^{1/2}\right\}^2 dG(t,x)\\
&\qquad + \tfrac12\int\left\{\left(1-\hat{F}_{n,\beta}(t-\beta'x)\right)^{1/2}-\left(1-F_\beta(t-\beta'x)\right)^{1/2}\right\}^2 dG(t,x),
\end{aligned}$$

and

$$\begin{aligned}
&\int\left\{\hat{F}_{n,\beta}(t-\beta'x)-F_\beta(t-\beta'x)\right\}^2 dG(t,x)\\
&= \int\left\{\hat{F}_{n,\beta}(t-\beta'x)^{1/2}-F_\beta(t-\beta'x)^{1/2}\right\}^2\left\{\hat{F}_{n,\beta}(t-\beta'x)^{1/2}+F_\beta(t-\beta'x)^{1/2}\right\}^2 dG(t,x)\\
&\leq 4\int\left\{\hat{F}_{n,\beta}(t-\beta'x)^{1/2}-F_\beta(t-\beta'x)^{1/2}\right\}^2 dG(t,x) \leq 8h\left(p_{\hat{F}_{n,\beta}}, p_{F_\beta}\right)^2
\end{aligned}$$

So we find

$$\sup_{\beta\in\Theta}\int\left\{\hat{F}_{n,\beta}(t-\beta'x)-F_\beta(t-\beta'x)\right\}^2 dG(t,x) = O_p\left(n^{-2/3}\right).$$

□

S2. Asymptotic behavior of the simple estimate based on the MLE $\hat{F}_{n,\beta}$, avoiding any smoothing. This section contains the proof of Theorem 4.1 stated in Section 4.1 of the manuscript. The proof is decomposed into three parts: (a) proof of existence of a root of $\psi_{1,n}$, (b) proof of consistency of $\hat\beta_n$ and (c) proof of asymptotic normality of $\sqrt{n}(\hat\beta_n - \beta_0)$. We first prove the properties given in Lemma 4.1 of the population version of the statistic $\psi_{1,n}^{(\epsilon)}$ defined by,

$$
\begin{aligned}
\psi_{1,\epsilon}(\beta) &= \int_{F_\beta(t-\beta' x)\in[\epsilon,1-\epsilon]} x\{\delta - F_\beta(t-\beta' x)\}\,dP_0(t,x,\delta)\\
&= \int_{F_\beta(t-\beta' x)\in[\epsilon,1-\epsilon]} x\{F_0(t-\beta_0 x) - F_\beta(t-\beta' x)\}\,dG(t,x). \qquad \text{(S2.1)}
\end{aligned}
$$

Proof of Lemma 4.1 . We first note that,

$$
\psi_{1,\epsilon}(\beta_0) = \int_{F_0(t-\beta_0' x)\in[\epsilon,1-\epsilon]} x\Big\{\mathbb{E}\{\Delta|(T,X)=(t,x)\} - F_0(t-\beta_0' x)\Big\}\,dG(t,x) = 0,
$$

since $\mathbb{E}\{\Delta|(T,X)=(t,x)\} = F_0(t-\beta_0' x)$. We next continue by showing result (i). Since

$$
\mathbb{E}\left(\Delta|T-\beta' X = t-\beta' x\right) = F_\beta(t-\beta' x),
$$

we get,

$$
\begin{aligned}
&\mathbb{E}_{\epsilon,\beta}\left[\mathrm{Cov}\left(\Delta, X|T-\beta' X\right)\right]\\
&\stackrel{\mathrm{def}}{=} \int_{F_\beta(t-\beta' x)\in[\epsilon,1-\epsilon]} \mathrm{Cov}\left(\Delta, X|T-\beta' X = u\right) f_{T-\beta' X}(u)\,du\\
&= \int_{F_\beta(u)\in[\epsilon,1-\epsilon]} \mathrm{Cov}\left\{X, F_0(u+(\beta-\beta_0)' X)\;\Big|\; T-\beta' X = u\right\} f_{T-\beta' X}(u)\,du\\
&= \int_{F_\beta(t-\beta' x)\in[\epsilon,1-\epsilon]} x\Big\{F_0(t-\beta' x+(\beta-\beta_0)' x) - F_\beta(t-\beta' x)\Big\}\,dG(t,x)\\
&= \int_{F_\beta(t-\beta' x)\in[\epsilon,1-\epsilon]} x\left\{F_0(t-\beta_0' x) - F_\beta(t-\beta' x)\right\}\,dG(t,x) = \psi_{1,\epsilon}(\beta).
\end{aligned}
$$

For the second result (ii), we write

$$
\begin{aligned}
&(\beta-\beta_0)'\mathrm{Cov}\left(\Delta, X|T-\beta' X = u\right)\\
&\quad= \mathrm{Cov}\left(F_0(T-\beta' X+(\beta-\beta_0)' X), (\beta-\beta_0)' X|T-\beta' X = u\right)
\end{aligned}
$$

which is positive for all β, following from the fact that F_0 is an increasing function. Indeed, using Fubini's theorem, one can prove that for any random variables X and Y such that XY, X and Y are integrable, we have

$$
\mathrm{Cov}\left\{X,Y\right\} = EXY - EXEY = \int \{\mathbb{P}(X\ge s, Y\ge t) - \mathbb{P}(X\ge s)\mathbb{P}(Y\ge t)\}\,ds\,dt.
$$

Denote $Z_1 = (\beta-\beta_0)' X$ and $Z_2 = F_0(u+(\beta-\beta_0)' X) = F_0(u+Z_1)$. For simplicity of notation we no longer write the condition $T-\beta' X = u$ but note that the results below hold conditional on

$T-\beta'X=u$. Using the monotonicity of the function F_0, we have

$$\begin{aligned}\mathbb{P}(Z_1\ge z_1,Z_2\ge z_2)&=\mathbb{P}(Z_1\ge\max\{z_1,\tilde z_2\})\ge\mathbb{P}(Z_1\ge\max\{z_1,\tilde z_2\})\mathbb{P}(Z_1\ge\min\{z_1,\tilde z_2\})\\&=\mathbb{P}(Z_1\ge z_1)\mathbb{P}(Z_2\ge z_2),\end{aligned}$$

where

$$\tilde z_2=F_0^{-1}(z_2)-u.$$

We conclude that,

$$\begin{aligned}&\mathrm{Cov}\left(F_0(T-\beta'X+(\beta-\beta_0)'X),(\beta-\beta_0)'X|T-\beta'X=u\right)\\&\qquad=\int\{\mathbb{P}(Z_1\ge z_1,Z_2\ge z_2)-\mathbb{P}(Z_1\ge z_1)\mathbb{P}(Z_2\ge z_2)\}\,dz_1\,dz_2\ge0,\end{aligned}$$

and hence (ii) follows from the assumption that the covariance $\mathrm{Cov}(X,F_0(u+(\beta-\beta_0)'X)|T-\beta'X=u)$ is not identically zero for u in the region $A_{\epsilon,\beta}$, for each $\beta\in\Theta$, implying:

$$\begin{aligned}&\mathbb{E}_{\epsilon,\beta}\left[\mathrm{Cov}\left(\Delta,X|T-\beta'X\right)\right]\\&=\int_{F_\beta(u)\in[\epsilon,1-\epsilon]}\mathrm{Cov}\left(F_0(T-\beta'X+(\beta-\beta_0)'X),(\beta-\beta_0)'X|T-\beta'X=u\right)f_{T-\beta'X}(u)du\ge0.\end{aligned}$$

[Uniqueness of β_0:]

We next show that β_0 is the only value $\beta_*\in\Theta$ such that $\mathbb{E}_{\epsilon,\beta}\left[(\beta-\beta_*)'\mathrm{Cov}\left(\Delta,X|T-\beta'X\right)\right]\ge0$ for all $\beta\in\Theta$. We start by assuming that, on the contrary, there exists $\beta_1\ne\beta_0$ in Θ such that

$$(\beta-\beta_0)'\psi_{1,\epsilon}(\beta)\ge0\qquad\text{and }(\beta-\beta_1)'\psi_{1,\epsilon}(\beta)\ge0\qquad\text{for all }\beta\in\Theta,$$

and we consider the point $\tilde\beta\in\Theta$ given by

$$\tilde\beta=\tfrac12\{\beta_0+\beta_1\}.$$

The existence of the point $\tilde\beta$ is ensured by the convexity of the set Θ. For this point, we have,

$$(\tilde\beta-\beta_0)'\psi_{1,\epsilon}(\tilde\beta)=-(\tilde\beta-\beta_1)'\psi_{1,\epsilon}(\tilde\beta),$$

which is not possible since both terms should be positive and $\psi_{1,\epsilon}(\tilde\beta)$ is not equal to zero (since, by the assumption that the covariance $\mathrm{Cov}(X,F_0(u+(\beta-\beta_0)'X)|T-\beta'X=u)$ is not identically zero for u in the region $A_{\epsilon,\beta}$, $\psi_{1,\epsilon}(\beta)$ is only zero at $\beta=\beta_0$.)

We now calculate the derivative of $\psi_{1,\epsilon}$ at $\beta=\beta_0$. We have,

$$\begin{aligned}\psi_{1,\epsilon}'(\beta)&=\frac{\partial}{\partial\beta}\int_{F_\beta^{-1}(\epsilon)\le t-\beta'x\le F_\beta^{-1}(1-\epsilon)}x\{\delta-F_\beta(t-\beta'x)\}dP_0(t,x,\delta)\\&=\frac{\partial}{\partial\beta}\int_{F_\beta^{-1}(\epsilon)\le t-\beta'x\le F_\beta^{-1}(1-\epsilon)}x\{F_0(t-\beta_0'x)-F_\beta(t-\beta'x)\}\,dG(t,x)\\&=\frac{\partial}{\partial\beta}\int_{u=F_\beta^{-1}(\epsilon)}^{F_\beta^{-1}(1-\epsilon)}\int x\left\{F_0(u+(\beta-\beta_0)'x)-F_\beta(u)\right\}f_{X|T-\beta'X}(x|u)f_{T-\beta'X}(u)\,dx\,du\end{aligned}$$

$$= \int_{u=F_\beta^{-1}(\epsilon)}^{F_\beta^{-1}(1-\epsilon)} \int \frac{\partial}{\partial\beta} \left\{x \left\{F_0(u+(\beta-\beta_0)'x) - F_\beta(u)\right\} f_{X|T-\beta'X}(x|u) f_{T-\beta'X}(u)\right\} \, dx\, du$$

$$+ \left\{\frac{\partial}{\partial\beta} F_\beta^{-1}(1-\epsilon)\right\} \int x \left\{F_0(F_\beta^{-1}(1-\epsilon)+(\beta-\beta_0)'x) - (1-\epsilon)\right\}$$
$$\cdot f_{X|T-\beta'X}(x|F_\beta^{-1}(1-\epsilon)) f_{T-\beta'X}(F_\beta^{-1}(1-\epsilon))\, dx$$

$$- \left\{\frac{\partial}{\partial\beta} F_\beta^{-1}(\epsilon)\right\} \int x \left\{F_0(F_\beta^{-1}(\epsilon)+(\beta-\beta_0)x) - \epsilon\right\}$$
$$\cdot f_{X|T-\beta'X}(x|F_\beta^{-1}(\epsilon)) f_{T-\beta'X}(F_\beta^{-1}(\epsilon))\, dx.$$

Note that if $\beta = \beta_0$, we get:

$$\psi'_{1,\epsilon}(\beta_0) =$$
$$\int_{F_0^{-1}(\epsilon)}^{F_0^{-1}(1-\epsilon)} \int \frac{\partial}{\partial\beta} \left\{x \left\{F_0(u+(\beta-\beta_0)'x) - F_\beta(u)\right\} f_{X|T-\beta'X}(x|u) f_{T-\beta'X}(u)\right\}\Big|_{\beta=\beta_0} du\, dx.$$

since the last two terms vanish because the integrands become zero in that case. Note that,

$$\frac{\partial}{\partial\beta} F_\beta(u) = \int y f_0(u+(\beta-\beta_0)'y) f_{X|T-\beta'X}(y|u)\, dy$$
$$+ \int F_0(u+(\beta-\beta_0)'y) \frac{\partial}{\partial\beta} f_{X|T-\beta'X}(y|u)\, dy,$$

implying that, at $\beta = \beta_0$,

$$\frac{\partial}{\partial\beta} F_\beta(u)\Big|_{\beta=\beta_0} = f_0(u)\mathbb{E}(X|T-\beta_0'X=u).$$

Since

$$\int_{u=F_0^{-1}(\epsilon)}^{F_0^{-1}(1-\epsilon)} \int x \left\{x - \mathbb{E}(X|T-\beta_0'X=u)\right\}' f_{X|T-\beta_0'X}(x|u)\, f_0(u) f_{T-\beta_0'X}(u)\, dx\, du$$
$$= \mathbb{E}_\epsilon\left[X\{X - \mathbb{E}(X|T-\beta_0'X)\}' f_0(T-\beta_0'X)\right],$$

(4.4) now follows. □

Proof of Theorem 4.1, Part 1 (Existence of a root). Consider the score function

$$\psi_{1,n}^{(\epsilon)}(\beta) = \int_{\hat F_{n,\beta}(t-\beta'x)\in[\epsilon,1-\epsilon]} x\left\{\delta - \hat F_{n,\beta}(t-\beta'x)\right\} d\mathbb{P}_n(t,x,\delta),$$

where $\hat F_{n,\beta}$ is the nonparametric maximum likelihood estimator (MLE) of the error distribution. According to the discussion in Section 4.1 we have to show that there exists a point $\hat\beta_n$ such that

$$\psi_{1,n}^{(\epsilon)}(\beta) = \int_{\hat F_{n,\beta}(t-\beta'x)\in[\epsilon,1-\epsilon]} x\left\{\delta - \hat F_{n,\beta}(t-\beta'x)\right\} d\mathbb{P}_n(t,x,\delta)$$

has a zero-crossing at $\beta = \hat\beta_n$. We have:

(S2.2)
$$\begin{aligned}\psi_{1,n}^{(\epsilon)}(\beta) &= \int_{\hat F_{n,\beta}(t-\beta' x)\in[\epsilon,1-\epsilon]} x\bigl\{\delta - F_\beta(t-\beta' x)\bigr\}\, d\mathbb{P}_n(t,x,\delta)\\ &\quad + \int_{\hat F_{n,\beta}(t-\beta' x)\in[\epsilon,1-\epsilon]} x\bigl\{F_\beta(t-\beta' x) - \hat F_{n,\beta}(t-\beta' x)\bigr\}\, d\mathbb{P}_n(t,x,\delta)\\ &= \int_{\hat F_{n,\beta}(t-\beta' x)\in[\epsilon,1-\epsilon]} x\bigl\{\delta - F_\beta(t-\beta' x)\bigr\}\, d\mathbb{P}_n(t,x,\delta)\\ &\quad + \int_{\hat F_{n,\beta}(t-\beta' x)\in[\epsilon,1-\epsilon]} x\bigl\{F_\beta(t-\beta' x) - \hat F_{n,\beta}(t-\beta' x)\bigr\}\, d(\mathbb{P}_n - P_0)(t,x,\delta)\\ &\quad + \int_{\hat F_{n,\beta}(t-\beta' x)\in[\epsilon,1-\epsilon]} x\bigl\{F_\beta(t-\beta' x) - \hat F_{n,\beta}(t-\beta' x)\bigr\} dP_0(t,x,\delta).\end{aligned}$$

Let $\mathcal{F}$ be the set of piecewise constant distribution functions with finitely many jumps (like the MLE $\hat F_{n,\hat\beta_n}$), and let, for $\beta \in \Theta$, $\mathcal{K}$ be the set of functions

$$\mathcal{K} = \left\{(t,x,\delta) \mapsto x\bigl\{\delta - F_\beta(t-\beta' x)\bigr\} 1_{[\epsilon,1-\epsilon]}\left(F(t-\beta' x)\right) : F \in \mathcal{F}, \beta \in \Theta\right\}. \tag{S2.3}$$

We add the function

$$(t,x,\delta) \mapsto x\bigl\{\delta - F_\beta(t-\beta' x)\bigr\} 1_{[\epsilon,1-\epsilon]}\left(F_\beta(t-\beta' x)\right)$$

to $\mathcal{K}$. We denote by $H_B(\zeta, \mathcal{K}, L_2(P_0))$ the bracketing ζ-entropy w.r.t. the L_2-distance d, defined by

$$d(k_1,k_2)^2 = \int \|k_1 - k_2\|^2\, dP_0, \qquad k_1, k_2 \in \mathcal{K}. \tag{S2.4}$$

Note that

$$x\bigl\{\delta - F_\beta(t-\beta' x)\bigr\} 1_{[\epsilon,1-\epsilon]}\left(F(t-\beta' x)\right) = f_{1,\beta}(t,x,\delta) f_{2,\beta}(t,x,\delta),$$

where

$$f_{1,\beta}(t,x,\delta) = x\bigl\{\delta - F_\beta(t-\beta' x)\bigr\},$$

and

$$f_{2,\beta}(t,x,\delta) = 1_{[\epsilon,1-\epsilon]}\left(F(t-\beta' x)\right).$$

Since t and x vary over a bounded region and, by (A4), F_β is of bounded variation, $f_{1,\beta}$ is of bounded variation. Moreover,

$$f_{2,\beta}(t,x,\delta) = 1_{[\epsilon,1-\epsilon]}\left(F(t-\beta' x)\right) = 1_{[\epsilon,1]}\left(F(t-\beta' x)\right) - 1_{(1-\epsilon,1]}\left(F(t-\beta' x)\right).$$

Since F is monotone, we have:

$$1_{[\epsilon,1]}\left(F(t-\beta' x)\right) - 1_{(1-\epsilon,1]}\left(F(t-\beta' x)\right) = 1_{[a_{\epsilon,F},M]}(t-\beta' x) - 1_{(b_{\epsilon,F},M]}(t-\beta' x) \tag{S2.5}$$

for points $a_{\epsilon,F} \le b_{\epsilon,F}$, where M is an upper bound for the values of $t-\beta'x$. Hence $f_{2,\beta}$ is also a function of uniformly bounded variation.

We therefore get, using Lemma S1.1

$$\sup_{\zeta>0} \zeta H_B\left(\zeta,\mathcal{K},L_2(P_0)\right) = O(1),$$

which implies:

$$\int_0^{\zeta} H_B\left(u,\mathcal{K},L_2(P_0)\right)^{1/2}\,du = O\left(\zeta^{1/2}\right), \qquad \zeta>0.$$

This implies

$$\begin{aligned}
&\int_{\hat F_{n,\beta}(t-\beta'x)\in[\epsilon,1-\epsilon]} x\bigl\{\delta - F_\beta(t-\beta'x)\bigr\}\,d\mathbb{P}_n(t,x,\delta)\\
&= \int_{\hat F_{n,\beta}(t-\beta'x)\in[\epsilon,1-\epsilon]} x\bigl\{\delta - F_\beta(t-\beta'x)\bigr\}\,dP_0(t,x,\delta)\\
&\qquad + \int_{\hat F_{n,\beta}(t-\beta'x)\in[\epsilon,1-\epsilon]} x\bigl\{\delta - F_\beta(t-\beta'x)\bigr\}\,d\bigl(\mathbb{P}_n-P_0\bigr)(t,x,\delta)\\
&= \int_{F_\beta(t-\beta'x)\in[\epsilon,1-\epsilon]} x\bigl\{\delta - F_\beta(t-\beta'x)\bigr\}\,dP_0(t,x,\delta)\\
&\qquad + \int_{F_\beta(t-\beta'x)\in[\epsilon,1-\epsilon]} x\bigl\{\delta - F_\beta(t-\beta'x)\bigr\}\,d\bigl(\mathbb{P}_n-P_0\bigr)(t,x,\delta) + o_p(1)\\
&= \psi_{1,\epsilon}(\beta) + o_p(1),
\end{aligned}$$

uniformly in β in Θ, by the convergence in probability (and almost surely) of $\hat F_{n,\beta}$ to F_β, where we use Lemma S0.1 for the second term on the right-hand side of the first equality to make the transition of the integration region $\hat F_{n,\beta}(t-\beta'x)\in[\epsilon,1-\epsilon]$ to $F_\beta(t-\beta'x)\in[\epsilon,1-\epsilon]$.

For the second term of (S2.2) we argue similarly, this time using the function class

$$\mathcal{K}' = \left\{(t,x,\delta)\mapsto x\bigl\{F_\beta(t-\beta'x) - F(t-\beta'x)\bigr\}1_{[\epsilon,1-\epsilon]}\left(F(t-\beta'x)\right) : F\in\mathcal{F},\,\beta\in\Theta\right\}. \tag{S2.6}$$

to which we add the function that is identically zero. This implies that these terms are $o_p(1)$. For the third term of (S2.2) we get by an application of the Cauchy-Schwarz inequality that, uniformly in β,

$$\begin{aligned}
&\int_{\hat F_{n,\beta}(t-\beta'x)\in[\epsilon,1-\epsilon]} x\bigl\{F_\beta(t-\beta'x) - \hat F_{n,\beta}(t-\beta'x)\bigr\}dP_0(t,x,\delta)\\
&\le \left(\int_{\hat F_{n,\beta}(t-\beta'x)\in[\epsilon,1-\epsilon]} x^2 dP_0(t,x,\delta)\int_{\hat F_{n,\beta}(t-\beta'x)\in[\epsilon,1-\epsilon]}\bigl\{F_\beta(t-\beta'x) - \hat F_{n,\beta}(t-\beta'x)\bigr\}^2 dP_0(t,x,\delta)\right)^{1/2}\\
&= O_p(n^{-1/3}).
\end{aligned}$$

The conclusion is that,

$$\psi_{1,n}^{(\epsilon)}(\beta) = \psi_{1,\epsilon}(\beta) + o_p(1), \tag{S2.7}$$

uniformly in $\beta \in \Theta$.

[Existence of $\hat\beta_n$:] Let $\psi_{1,\epsilon}$ be the population version of the statistic $\psi_{1,n}^{(\epsilon)}$ defined by,

$$\psi_{1,\epsilon}(\beta) = \int_{F_\beta(t-\beta' x)\in[\epsilon,1-\epsilon]} x\{\delta - F_\beta(t-\beta' x)\}\,dP_0(t,x,\delta). \tag{S2.8}$$

We have

$$\psi_{1,\epsilon}(\beta_0) = 0.$$

Furthermore,

$$\psi_{1,n}^{(\epsilon)}(\beta) = \psi_{1,\epsilon}'(\beta_0)(\beta-\beta_0) + R_n(\beta), \tag{S2.9}$$

where $R_n(\beta) = o_p(1) + o(\beta-\beta_0)$, and where the $o_p(1)$ term is uniform in $\beta\in\Theta$. Note that $\psi_{1,\epsilon}'(\beta_0)$ is by definition non-singular.

We now define, for $h>0$, the functions

$$\tilde R_{n,h}(\beta) = h^{-d}\int K_h(u_1-\beta_1)\dots K_h(u_d-\beta_d)\,R_n(u_1,\dots,u_d)\,du_1\dots du_d,$$

where d is the dimension of Θ and

$$K_h(x) = h^{-1}K(x/h), \qquad x\in\mathbb{R},$$

letting K be one of the usual smooth kernels with support $[-1,1]$, like the Triweight kernel that we used in the simulations.

Furthermore, we define:

$$\tilde\psi_{1,n,h}^{(\epsilon)}(\beta) = \psi_{1,\epsilon}'(\beta_0)(\beta-\beta_0) + \tilde R_{nh}(\beta).$$

Clearly:

$$\lim_{h\downarrow 0}\tilde\psi_{1,n,h}^{(\epsilon)}(\beta)) = \psi_{1,n}^{(\epsilon)}(\beta) \qquad \text{and} \qquad \lim_{h\downarrow 0}\tilde R_{nh}(\beta) = R_n(\beta)$$

for each continuity point β of $\psi_{1,n}^{(\epsilon)}$.

We now reparametrize, defining

$$\gamma = \psi_{1,\epsilon}'(\beta_0)\beta, \qquad \gamma_0 = \psi_{1,\epsilon}'(\beta_0)\beta_0.$$

This gives:

$$\psi_{1,\epsilon}'(\beta_0)(\beta-\beta_0) + \tilde R_{nh}(\beta) = \gamma-\gamma_0 + \tilde R_{nh}\left(\psi_{1,\epsilon}'(\beta_0)^{-1}\gamma\right).$$

By (S2.9), the mapping

$$\gamma \mapsto \gamma_0 - R_n\left(\psi_{1,\epsilon}'(\beta_0)^{-1}\gamma\right),$$

maps, for each $\eta > 0$, the ball $B_\eta(\gamma_0) = \{\gamma : \|\gamma - \gamma_0\} \le \eta\}$ into $B_{\eta/2}(\gamma_0) = \{\gamma : \|\gamma - \gamma_0\} \le \eta/2\}$ for all large n, with probability tending to one, where $\|\cdot\|$ denotes the Euclidean norm, implying that the *continuous* map

$$\gamma \mapsto \gamma_0 - \tilde{R}_{nh}\left(\psi'_{1,\epsilon}(\beta_0)^{-1}\gamma\right),$$

maps $B_\eta(\gamma_0) = \{\gamma : \|\gamma - \gamma_0\} \le \eta\}$ into itself for all large n and small h. So for large n and small h there is, by Brouwer's fixed point theorem a point γ_{nh} such that

$$\gamma_{nh} = \gamma_0 - \tilde{R}_{nh}\left(\psi'_{1,\epsilon}(\beta_0)^{-1}\gamma_{nh}\right).$$

Defining $\beta_{nh} = \psi'_{1,\epsilon}(\beta_0)^{-1}\gamma_{nh}$, we get:

$$\text{(S2.10)} \qquad \tilde{\psi}^{(\epsilon)}_{1,n,h}(\beta_{nh}) = \psi'_{1,\epsilon}(\beta_0)(\beta_{nh} - \beta_0) + \tilde{R}_{nh}(\beta_{nh}) = 0.$$

By compactness, $(\beta_{n,1/k})_{k=1}^\infty$ must have a subsequence $(\beta_{n,1/k_i})$ with a limit $\tilde{\beta}_n$, as $i \to \infty$. We show that each component of $\psi^{(\epsilon)}_{1,n}$ has a crossing of zero at $\tilde{\beta}_n$.

Suppose that the jth component $\psi^{(\epsilon)}_{1,n,j}$ of $\psi^{(\epsilon)}_{1,n}$ does not have a crossing of zero at $\tilde{\beta}_n$. Then there must be an open ball $B_\delta(\tilde{\beta}_n) = \{\beta : \|\beta - \tilde{\beta}_n\| < \delta\}$ of $\tilde{\beta}_n$ such that $\psi^{(\epsilon)}_{1,n,j}$ has a constant sign in $B_\delta(\tilde{\beta}_n)$, say $\psi^{(\epsilon)}_{1,n,j}(\beta) > 0$ for $\beta \in B_\delta(\tilde{\beta}_n)$. Since $\psi^{(\epsilon)}_{1,n,j}$ only has finitely many values, this means that

$$\psi^{(\epsilon)}_{1,n,j}(\beta) \ge c > 0, \qquad \text{for all } \beta \in B_\delta(\tilde{\beta}_n),$$

for some $c > 0$. This means that the jth component $\tilde{\psi}^{(\epsilon)}_{1,n,h,j}$ of $\tilde{\psi}^{(\epsilon)}_{1,n,h}$ satisfies

$$\begin{aligned}
&\tilde{\psi}^{(\epsilon)}_{1,n,h,j}(\beta) = \psi'_{1,\epsilon}(\beta_0)(\beta - \beta_0) + \tilde{R}_{nh}(\beta)\\
&= h^{-d}\int\left\{\left[\psi'_{1,\epsilon}(\beta_0)(\beta - \beta_0)\right]_j + R_{nj}(u_1,\dots,u_d)\right\}K_h(u_1-\beta_1)\dots K_h(u_d-\beta_d)\,du_1\dots du_d\\
&\ge h^{-d}\int\left\{\left[\psi'_{1,\epsilon}(\beta_0)(u - \beta_0)\right]_j + R_{nj}(u_1,\dots,u_d)\right\}K_h(u_1-\beta_1)\dots K_h(u_d-\beta_d)\,du_1\dots du_d - c/2\\
&\ge c\,h^{-d}\int K_h(u_1-\beta_1)\dots K_h(u_d-\beta_d)\,du_1\dots du_d - c/2\\
&= c/2,
\end{aligned}$$

for $\beta \in B_{\delta/2}(\tilde{\beta}_n)$ and sufficiently small h, contradicting (S2.10), since β_{nh}, for $h = 1/k_i$, belongs to $B_{\delta/2}(\tilde{\beta}_n)$ for large k_i.

□

Proof of Theorem 4.1, Part 2 (Consistency). We assume that $\hat{\beta}_n$ is contained in the compact set Θ, and hence the sequence $(\hat{\beta}_n)$ has a subsequence $(\hat{\beta}_{n_k} = \hat{\beta}_{n_k}(\omega))$, converging to an element β_*. If $\hat{\beta}_{n_k} = \hat{\beta}_{n_k}(\omega) \longrightarrow \beta_*$, we get by Lemma 3.1,

$$\hat{F}_{n_k,\hat{\beta}_{n_k}}(t - \hat{\beta}'_{n_k}x) \longrightarrow F_{\beta_*}(t - \beta'_*x),$$

where F_β is defined in (3.2). In the limit we get therefore the relation

$$\lim_{k\to\infty}\int_{\hat F_{n_k,\hat\beta_{n_k}}(t-\hat\beta_{n_k}'x)\in[\epsilon,1-\epsilon]} x\bigl\{\delta - F_{n_k,\hat\beta_{n_k}}(t-\hat\beta_{n_k}'x)\bigr\}\,d\mathbb{P}_{n_k}(t,x,\delta)$$
$$=\int_{F_{\beta_*}(t-\beta_*'x)\in[\epsilon,1-\epsilon]} x\bigl\{F_0(t-\beta_0'x)-F_{\beta_*}(t-\beta_*'x)\bigr\}\,dG(t,x)=0,$$

using that, in the limit, the crossing of zero becomes a root of the continuous limiting function. Consider

$$\int_{F_{\beta_*}(t-\beta_*'x)\in[\epsilon,1-\epsilon]} x\bigl\{F_0(t-\beta_0'x)-F_{\beta_*}(t-\beta_*'x)\bigr\}\,dG(t,x)$$
$$=\int_{F_{\beta_*}(t-\beta_*'x)\in[\epsilon,1-\epsilon]} x\bigl\{F_0(t-\beta_*'x+(\beta_*-\beta_0)'x)-F_{\beta_*}(t-\beta_*'x)\bigr\}\,dG(t,x).$$

Since

$$F_{\beta_*}(t-\beta_*'x)=\int F_0(t-\beta_*'x+(\beta_*-\beta_0)'y)f_{X|T-\beta_*'X}(y|T-\beta_*'X=t-\beta_*'x)\,dy,$$

we get:

$$(\beta_*-\beta_0)'\int_{F_\beta(t-\beta'x)\in[\epsilon,1-\epsilon]} x\Bigl\{F_0(t-\beta_*'x+(\beta_*-\beta_0)'x)-F_{\beta_*}(t-\beta_*'x)\Bigr\}\,dG(t,x)$$
$$=\int_{F_{\beta_*}(t-\beta_*'x)\in[\epsilon,1-\epsilon]}(\beta_*-\beta_0)'x\Bigl\{F_0(t-\beta_*'x+(\beta_*-\beta_0)'x)$$
$$-\int F_0(t-\beta_*'x+(\beta_*-\beta_0)'y)f_{X|T-\beta_*'X}(y|T-\beta_*'X=t-\beta_*'x)\,dy\Bigr\}\,dG(t,x)$$
$$=\int_{F_{\beta_*}(u)\in[\epsilon,1-\epsilon]}\mathrm{Cov}\Bigl\{(\beta_*-\beta_0)'X,F_0(u+(\beta_*-\beta_0)'X)\;\Big|\;T-\beta_*'X=u\Bigr\}f_{T-\beta_*'X}(u)\,du$$
$$=0.$$

We first note that by Lemma 4.1 the integrand is positive for all $\beta_*\in\Theta$. Suppose that $\beta_*\neq\beta_0$, then this integral can only be zero if $\mathrm{Cov}((\beta_*-\beta_0)'X,F_0(u+(\beta_*-\beta_0)'X)|T-\beta_*'X=u)$ is zero for all u such that $F_{\beta_*}(u)\in[\epsilon,1-\epsilon]$, if $f_{T-\beta_*'X}(u)$ stays away from zero on this region (Assumptions (A3)), using continuity of the functions in the integrand (Assumptions (A5)) and the non negativity of the conditional covariance function (see also Remark 4.2). Since this is excluded by the condition that the covariance $\mathrm{Cov}(X,F_0(u+(\beta-\beta_0)'X)|T-\beta'X=u)$ is continuous in u and not identically zero for u in the region $\{u:\epsilon\le F_\beta(u)\le 1-\epsilon\}$, for each $\beta\in\Theta$, we must have: $\beta_*=\beta_0$. □

PROOF OF THEOREM 4.1, PART 3 (ASYMPTOTIC NORMALITY). Before working out the details, we give a kind of "road map" for the proof of Theorem 4.1, Part 3.

1. We define $\psi_{1,n}^{(\epsilon)}$ at $\hat\beta_n$ by putting

$$\psi_{1,n}^{(\epsilon)}(\hat\beta_n)=0. \tag{S2.11}$$

Note that, with this definition, $\psi_{1,n}^{(\epsilon)}(\hat\beta_n)$ is in dimension 1 just the convex combination of the left and right limit at $\hat\beta_n$:

$$\psi_{1,n}^{(\epsilon)}(\hat\beta_n) = \alpha\psi_{1,n}^{(\epsilon)}(\hat\beta_n-) + (1-\alpha)\psi_{1,n}^{(\epsilon)}(\hat\beta_n+) = 0, \tag{S2.12}$$

where we can choose $\alpha \in [0,1]$ in such a way that (S2.12) holds. In dimension d higher than one, we can also define $\psi_{1,n}^{(\epsilon)}$ at $\hat\beta_n$ by (S2.11) and use the representation of the components as a convex combination since we have a crossing of zero componentwise. Since the following asymptotic representations are also valid for one-sided limits as used in (S2.12) we can use Definition (S2.11) and assume $\psi_{1,n}^{(\epsilon)}(\hat\beta_n) = 0$.
We show:

$$\begin{aligned}
&\psi_{1,n}^{(\epsilon)}(\hat\beta_n)\\
&= \int_{F_0(t-\beta_0'x)\in[\epsilon,1-\epsilon]} \Big\{x - \phi_0(t-\beta_0'x)\Big\}\Big\{F_0(t-\beta_0'x) - F_{\hat\beta_n}(t-\hat\beta_n'x)\Big\}\,dP_0(t,x,\delta)\\
&\quad + \int_{F_0(t-\beta_0'x)\in[\epsilon,1-\epsilon]} \Big\{x - \phi_0(t-\beta_0'x)\Big\}\Big\{\delta - F_0(t-\beta_0'x)\Big\}\,d\big(\mathbb{P}_n - P_0\big)(t,x,\delta)\\
&\quad + o_p\left(n^{-1/2} + \hat\beta_n - \beta_0\right),
\end{aligned} \tag{S2.13}$$

where

$$\phi_0(u) = \phi_{\beta_0}(u),$$

and where ϕ_β is defined by:

$$\phi_\beta(u) = \mathbb{E}\left\{X | T - \beta'X = u\right\}. \tag{S2.14}$$

Since $\hat\beta_n \xrightarrow{p} \beta_0$ and

$$\begin{aligned}
&\int_{F_0(t-\beta_0'x)\in[\epsilon,1-\epsilon]} \Big\{x - \phi_0(t-\beta_0'x)\Big\}\Big\{F_0(t-\beta_0'x) - F_{\hat\beta_n}(t-\hat\beta_n'x)\Big\}\,dP_0(t,x,\delta)\\
&= \psi_{1,\epsilon}'(\beta_0)\left(\hat\beta_n - \beta_0\right) + o_p\left(\hat\beta_n - \beta_0\right),
\end{aligned}$$

this yields, using the invertibility of $\psi_{1,\epsilon}'(\beta_0)$,

$$\begin{aligned}
&\sqrt{n}(\hat\beta_n - \beta_0)\\
&= -\psi_{1,\epsilon}'(\beta_0)^{-1}\Bigg\{\sqrt{n}\int_{F_0(t-\beta_0'x)\in[\epsilon,1-\epsilon]} \Big\{x - \phi_{\beta_0}(t-\beta_0'x)\Big\}\Big\{\delta - F_0(t-\beta_0'x)\Big\}\\
&\qquad\qquad d\big(\mathbb{P}_n - P_0\big)(t,x,\delta)\Bigg\} + o_p\left(1 + \sqrt{n}(\hat\beta_n - \beta_0)\right).
\end{aligned}$$

As a consequence, the result of Theorem 4.1 follows, since

$$\begin{aligned}
&\sqrt{n}\int_{F_0(t-\beta_0'x)\in[\epsilon,1-\epsilon]} \Big\{x - \phi_0(t-\beta_0'x)\Big\}\Big\{\delta - F_0(t-\beta_0'x)\Big\}\,d\big(\mathbb{P}_n - P_0\big)(t,x,\delta)\\
&\xrightarrow{d} N(0,B).
\end{aligned}$$

2. To show that (S2.13) holds, we need entropy results for the functions $u \mapsto \hat{F}_{n,\beta}(u)$ and $u \mapsto \bar{\phi}_{\beta,\hat{F}_{n,\beta}}(u)$ (see (S2.15) below). We also have to deal with the simpler parametric functions F_β and ϕ_β, parametrized by the finite dimensional parameter β, which are the population equivalents of $\hat{F}_{n,\beta}$ and $\bar{\phi}_{\beta,\hat{F}_{n,\beta}}$.
3. The result will then follow from the properties of F_β and ϕ_β, together with the closeness of $\hat{F}_{n,\beta}$ to F_β and $\bar{\phi}_{\beta,\hat{F}_{n,\beta}}$ to ϕ_β, respectively, and the convergence of $\hat{\beta}_n$ to β_0.

Let $\bar{\phi}_{\hat{\beta}_n,\hat{F}_{n,\hat{\beta}_n}}$ be a (random) piecewise constant version of $\phi_{\hat{\beta}_n}$, where, for a piecewise constant distribution function F with finitely many jumps at $\tau_1 < \tau_2 < \dots$, the function $\bar{\phi}_{\beta,F}$ is defined in the following way.

$$\bar{\phi}_{\beta,F}(u) = \begin{cases} \phi_\beta(\tau_i), & \text{if } F_\beta(u) > F(\tau_i),\ u \in [\tau_i, \tau_{i+1}), \\ \phi_\beta(s), & \text{if } F_\beta(u) = F(s),\ \text{for some } s \in [\tau_i, \tau_{i+1}), \\ \phi_\beta(\tau_{i+1}), & \text{if } F_\beta(u) < F(\tau_i),\ u \in [\tau_i, \tau_{i+1}). \end{cases} \tag{S2.15}$$

We can write:

$$\begin{aligned}
\psi_{1,n}^{(\epsilon)}(\hat{\beta}_n) &= \int_{\hat{F}_{n,\hat{\beta}_n}(t-\hat{\beta}_n' x)\in[\epsilon,1-\epsilon]} x\bigl\{\delta - \hat{F}_{n,\hat{\beta}_n}(t-\hat{\beta}_n' x)\bigr\}\, d\mathbb{P}_n(t,x,\delta) \\
&= \int_{\hat{F}_{n,\hat{\beta}_n}(t-\hat{\beta}_n' x)\in[\epsilon,1-\epsilon]} \bigl\{x - \phi_{\hat{\beta}_n}(t-\hat{\beta}_n' x)\bigr\}\bigl\{\delta - \hat{F}_{n,\hat{\beta}_n}(t-\hat{\beta}_n' x)\bigr\}\, d\mathbb{P}_n(t,x,\delta) \\
&\quad + \int_{\hat{F}_{n,\hat{\beta}_n}(t-\hat{\beta}_n' x)\in[\epsilon,1-\epsilon]} \bigl\{\phi_{\hat{\beta}_n}(t-\hat{\beta}_n' x) - \bar{\phi}_{\hat{\beta}_n,\hat{F}_{n,\hat{\beta}_n}}(t-\hat{\beta}_n' x)\bigr\} \\
&\qquad\qquad \cdot \bigl\{\delta - \hat{F}_{n,\beta}(t-\hat{\beta}_n' x)\bigr\}\, d\mathbb{P}_n(t,x,\delta) \\
&= I + II,
\end{aligned}$$

using

$$\int_{\hat{F}_{n,\hat{\beta}_n}(t-\hat{\beta}_n' x)\in[\epsilon,1-\epsilon]} \bar{\phi}_{\hat{\beta}_n,\hat{F}_{n,\hat{\beta}_n}}(t-\hat{\beta}_n' x)\bigl\{\delta - \hat{F}_{n,\hat{\beta}_n}(t-\hat{\beta}_n' x)\bigr\}\, d\mathbb{P}_n(t,x,\delta) = 0,$$

by the definition of the MLE $\hat{F}_{n,\hat{\beta}_n}$ as the slope of the greatest convex minorant of the corresponding cusum diagram, based on the values of the Δ_i in the ordering of the $T_i - \hat{\beta}_n' X_i$ (see also Lemma A.5 on p.380 of [2]).

Since the function $u \mapsto \phi_\beta(u)$ has a totally bounded derivative (as a consequence of (S2.14) and assumption (A5)), we can bound the Euclidean norm of the differences $\phi_\beta(u) - \bar{\phi}_{\beta,\hat{F}_{n,\beta}}(u)$ above by a constant times $|\hat{F}_{n,\beta}(u) - F_\beta(u)|$, for $u \in A_{\epsilon,\beta}$ (see (A2)), i.e.,

$$\|\phi_\beta(u) - \bar{\phi}_{\beta,\hat{F}_{n,\beta}}(u)\| \le K_\beta |\hat{F}_{n,\beta}(u) - F_\beta(u)|,$$

for some constant $K_\beta > 0$ where the constant K_β depends on β through f_β (see for this technique for example (10.64) in [3]). By Assumption (A2) we know that f_β is continuous for all $\beta \in \Theta$ such that we can find a constant $K > 0$ not depending on β, satisfying,

$$\|\phi_\beta(u) - \bar{\phi}_{\beta,\hat{F}_{n,\beta}}(u)\| \le K |\hat{F}_{n,\beta}(u) - F_\beta(u)|, \tag{S2.16}$$

uniformly in $\beta \in \Theta$. Note that we also need $f_\beta(u) > 0$ for applying this, which is ensured by (A2).

We have:

$$
\begin{aligned}
II &= \int_{\hat F_{n,\hat\beta_n}(t-\hat\beta_n' x)\in[\epsilon,1-\epsilon]} \left\{\phi_{\hat\beta_n}(t-\hat\beta_n'x) - \bar\phi_{\hat\beta_n,\hat F_{n,\hat\beta_n}}(t-\hat\beta_n'x)\right\} \\
&\qquad\qquad\qquad\qquad \cdot \left\{\delta - \hat F_{n,\hat\beta_n}(t-\hat\beta_n'x)\right\} d\mathbb{P}_n(t,x,\delta) \\
&= \int_{\hat F_{n,\hat\beta_n}(t-\hat\beta_n' x)\in[\epsilon,1-\epsilon]} \left\{\phi_{\hat\beta_n}(t-\hat\beta_n'x) - \bar\phi_{\hat\beta_n,\hat F_{n,\hat\beta_n}}(t-\hat\beta_n'x)\right\} \\
&\qquad\qquad\qquad\qquad \cdot \left\{\delta - \hat F_{n,\hat\beta_n}(t-\hat\beta_n'x)\right\} d(\mathbb{P}_n - P_0)(t,x,\delta) \\
&\quad + \int_{\hat F_{n,\hat\beta_n}(u)\in[\epsilon,1-\epsilon]} \left\{\phi_{\hat\beta_n}(u) - \bar\phi_{\hat\beta_n,\hat F_{n,\hat\beta_n}}(u)\right\}\left\{F_{\hat\beta_n}(u) - \hat F_{n,\hat\beta_n}(u)\right\} f_{T-\hat\beta_n' X}(u)\,du \\
&\quad + \int_{\hat F_{n,\hat\beta_n}(t-\hat\beta_n' x)\in[\epsilon,1-\epsilon]} \left\{\phi_{\hat\beta_n}(t-\hat\beta_n'x) - \bar\phi_{\hat\beta_n,\hat F_{n,\hat\beta_n}}(t-\hat\beta_n'x)\right\} \\
&\qquad\qquad\qquad\qquad \cdot \left\{F_0(t-\beta_0'x) - F_{\hat\beta_n}(t-\hat\beta_n'x)\right\} dP_0(t,x,\delta) \\
&= II_a + II_b + II_c.
\end{aligned}
$$

First consider II_a. Let $\mathcal{F}$ be the set of piecewise constant distribution functions with finitely many jumps (like the MLE $\hat F_{n,\hat\beta_n}$), and let $\mathcal{K}_1$ be the set of functions

$$
\begin{aligned}
\mathcal{K}_1 = \Big\{(t,x,\delta) \mapsto \big(\phi_\beta(t-\beta'x) - \bar\phi_{\beta,F}(t-\beta'x)\big)\big(\delta - F(t-\beta'x)\big) \\
\cdot 1_{[\epsilon,1-\epsilon]}(F(t-\beta'x)) : F\in\mathcal{F},\, \beta\in\Theta\Big\}\,,
\end{aligned}
\tag{S2.17}
$$

where $\bar\phi_{\beta,F}$ is again defined by (S2.15). We add the function which is identically zero to $\mathcal{K}_1$.

The functions $u \mapsto F(u)$, for $F \in \mathcal{F}$ and (as argued above) $u \mapsto \bar\phi_{\beta,F}(u)$ are bounded functions of uniformly bounded variation. Note that, for $F_1, F_2 \in \mathcal{F}$,

$$
\begin{aligned}
&F_1(t-\beta_1'x) - F_2(t-\beta_2'x) \\
&= F_1(t-\beta_1'x) - F_{\beta_1}(t-\beta_1'x) + F_{\beta_1}(t-\beta_1'x) - F_{\beta_2}(t-\beta_2'x) \\
&\qquad + F_{\beta_2}(t-\beta_2'x) - F_2(t-\beta_2'x),
\end{aligned}
$$

and that (see (3.2)):

$$
\begin{aligned}
&\left|F_{\beta_1}(t-\beta_1'x) - F_{\beta_2}(t-\beta_2'x)\right| \\
&= \Bigg|\int F_0(t-\beta_0'x + (\beta_1-\beta_0)'(y-x)) f_{X|T-\beta_1'X}(y|t-\beta_1'x)\,dy \\
&\qquad - \int F_0(t-\beta_0'x + (\beta_2-\beta_0)'(y-x)) f_{X|T-\beta_2'X}(y|t-\beta_2'x)\,dy\Bigg| \\
&= O\left(|\beta_1-\beta_2|\right),
\end{aligned}
$$

by (A2) and (A5).

For the indicator function $1_{[\epsilon,1]}\left(F(t-\beta'x)\right)$ we get, as in (S2.5), using the monotonicity of F,

$$
\begin{aligned}
&1_{[\epsilon,1]}\left(F(t-\beta'x)\right) \\
&= 1_{[\epsilon,1]}\left(F(t-\beta'x)\right) - 1_{(1-\epsilon,1]}\left(F(t-\beta'x)\right) = 1_{[a_{\epsilon,F},M]}(t-\beta'x) - 1_{(b_{\epsilon,F},M]}(t-\beta'x),
\end{aligned}
$$

for points $a_{\epsilon,F} \le b_{\epsilon,F}$, where M is an upper bound for the values of $t - \beta' x$, implying that the function

$$(t, x) \mapsto 1_{[\epsilon,1]}\left(F(t - \beta' x)\right),$$

is of uniformly bounded variation. So the functions in $\mathcal{K}_1$ are products of functions of uniformly bounded variation, and we therefore get, using Lemma S1.1

$$\sup_{\zeta>0} \zeta H_B\left(\zeta, \mathcal{K}_1, L_2(P_0)\right) = O(1),$$

which implies:

$$\int_0^{\zeta} H_B\left(u, \mathcal{K}_1, L_2(P_0)\right)^{1/2}\, du = O\left(\zeta^{1/2}\right), \qquad \zeta > 0.$$

Defining

$$k_{\beta,F}(t, x, \delta) = \left(\phi_\beta(t - \beta' x) - \bar\phi_{\beta,F}(t - \beta' x)\right)\left(\delta - F(t - \beta' x)\right) \cdot 1_{[\epsilon,1-\epsilon]}(F(t - \beta' x))$$

for $F \in \mathcal{F}$, we get, using (S2.16),

$$\begin{aligned}
&\left\{\int \left\|k_{\hat\beta_n,\hat F_{n,\hat\beta_n}}(t, x, \delta)\right\|^2 \, dP_0(t, x, \delta)\right\}^2 \\
&\le \int_{\hat F_{n,\hat\beta_n}(t-\hat\beta_n' x)\in[\epsilon,1-\epsilon]} \left\|\phi_{\hat\beta_n}(t - \hat\beta_n' x) - \bar\phi_{\hat\beta_n,\hat F_{n,\hat\beta_n}}(t - \hat\beta_n' x)\right\|^2 \, dP_0(t, x, \delta) \\
&\le K \int_{\hat F_{n,\hat\beta_n}(t-\hat\beta_n' x)\in[\epsilon,1-\epsilon]} \left\{\hat F_{n,\hat\beta_n}(t - \hat\beta_n' x) - F_{\hat\beta_n}(t - \hat\beta_n' x)\right\}^2 \, dP_0(t, x, \delta) \\
&\le K' \int_{\hat F_{n,\hat\beta_n}(u)\in[\epsilon,1-\epsilon]} \left\{\hat F_{n,\hat\beta_n}(u) - F_{\hat\beta_n}(u)\right\}^2 \, du \\
&\xrightarrow{p} 0,
\end{aligned}$$

for constants $K, K' > 0$. This implies

$$\sqrt{n} II_a = \sqrt{n} \int k_{\hat\beta_n,\hat F_{n,\hat\beta_n}}(t, x, \delta)\, d(\mathbb{P}_n - P_0)(t, x, \delta) = o_p(1), \tag{S2.18}$$

by an application of Lemma S0.1.

Using (S2.16), $\|F_{\hat\beta_n} - \hat F_{n,\hat\beta_n}\|_2 = O_p(n^{-1/3})$ and the Cauchy-Schwarz inequality on the second term we get,

$$II_b = O_p(n^{-2/3}).$$

The functions ϕ_β and F_β are of a simple parametric nature, since

$$\phi_\beta = \mathbb{E}(X | T - \beta' X),$$

and

$$F_\beta(u) = \int F_0(u + (\beta - \beta_0)'x) f_{X|T-\beta' X}(x|T - \beta' X = u)\, dx,$$

see (3.2). Moreover, since:

$$\begin{aligned} F_{\hat\beta_n}(u) &= F_0(u) + (\hat\beta_n - \beta_0)' \int x f_0(u) f_{X|T-\hat\beta_n' X}(x|u)\, dx + o_p(\hat\beta_n - \beta_0) \\ &= F_0(u) + (\hat\beta_n - \beta_0)' f_0(u) \mathbb{E}\{X|T - \hat\beta_n' X = u\} + o_p(\hat\beta_n - \beta_0), \end{aligned} \tag{S2.19}$$

and since the difference $\phi_{\hat\beta_n} - \bar\phi_{\hat\beta_n, \hat F_{n,\hat\beta_n}}$ converges to zero, we get for the third term II_c:

$$\begin{aligned} II_c &= \int_{\hat F_{n,\hat\beta_n}(t-\hat\beta_n' x)\in[\epsilon,1-\epsilon]} \left\{\phi_{\hat\beta_n}(t - \hat\beta_n' x) - \bar\phi_{\hat\beta_n, \hat F_{n,\hat\beta_n}}(t - \hat\beta_n' x)\right\} \\ &\qquad\qquad \cdot \left\{F_0(t - \beta_0' x) - F_{\hat\beta_n}(t - \hat\beta_n' x)\right\} dP_0(t, x, \delta) \\ &= o_p\left(\hat\beta_n - \beta_0\right). \end{aligned}$$

We therefore conclude,

$$\psi_{1,n}^{(\epsilon)}(\hat\beta_n) = I + o_p\left(n^{-1/2} + \hat\beta_n - \beta_0\right).$$

We now write,

$$\begin{aligned} I &= \int_{\hat F_{n,\hat\beta_n}(t-\hat\beta_n' x)\in[\epsilon,1-\epsilon]} \left\{x - \phi_{\hat\beta_n}(t - \hat\beta_n' x)\right\}\left\{\delta - \hat F_{n,\hat\beta_n}(t - \hat\beta_n' x)\right\} d\mathbb{P}_n(t, x, \delta) \\ &= \int_{\hat F_{n,\hat\beta_n}(t-\hat\beta_n' x)\in[\epsilon,1-\epsilon]} \left\{x - \phi_{\hat\beta_n}(t - \hat\beta_n' x)\right\}\left\{\delta - F_{\hat\beta_n}(t - \hat\beta_n' x)\right\} d\mathbb{P}_n(t, x, \delta) \\ &\quad + \int_{\hat F_{n,\hat\beta_n}(t-\hat\beta_n' x)\in[\epsilon,1-\epsilon]} \left\{x - \phi_{\hat\beta_n}(t - \hat\beta_n' x)\right\} \\ &\qquad\qquad \cdot \left\{F_{\hat\beta_n}(t - \hat\beta_n' x) - \hat F_{n,\hat\beta_n}(t - \hat\beta_n' x)\right\} d\mathbb{P}_n(t, x, \delta) \\ &= I_a + I_b. \end{aligned}$$

We get:

$$\begin{aligned} &I_a \\ &= \int_{\hat F_{n,\hat\beta_n}(t-\hat\beta_n' x)\in[\epsilon,1-\epsilon]} \left\{x - \phi_{\hat\beta_n}(t - \hat\beta_n' x)\right\}\left\{\delta - F_{\hat\beta_n}(t - \hat\beta_n' x)\right\} d\mathbb{P}_n(t, x, \delta) \\ &= \int_{\hat F_{n,\hat\beta_n}(t-\hat\beta_n' x)\in[\epsilon,1-\epsilon]} \left\{x - \phi_{\hat\beta_n}(t - \hat\beta_n' x)\right\}\left\{\delta - F_{\hat\beta_n}(t - \hat\beta_n' x)\right\} d(\mathbb{P}_n - P_0)(t, x, \delta) \\ &\quad + \int_{\hat F_{n,\hat\beta_n}(t-\hat\beta_n' x)\in[\epsilon,1-\epsilon]} \left\{x - \phi_{\hat\beta_n}(t - \hat\beta_n' x)\right\}\left\{F_0(t - \beta_0' x) - F_{\hat\beta_n}(t - \hat\beta_n' x)\right\} dP_0(t, x, \delta). \end{aligned}$$

For the second integral on the right-hand side we get:

$$\begin{aligned}
&\int_{\hat F_{n,\hat\beta_n}(t-\hat\beta_n'x)\in[\epsilon,1-\epsilon]}\Big\{x-\phi_{\hat\beta_n}(t-\hat\beta_n'x)\Big\}\Big\{F_0(t-\beta_0'x)-F_{\hat\beta_n}(t-\hat\beta_n'x)\Big\}\,dP_0(t,x,\delta)\\
&=\int_{\hat F_{n,\hat\beta_n}(t-\hat\beta_n'x)\in[\epsilon,1-\epsilon]}\Big\{x-\phi_{\hat\beta_n}(t-\hat\beta_n'x)\Big\}\Big\{F_0(t-\hat\beta_n'x)-F_{\hat\beta_n}(t-\hat\beta_n'x)\Big\}\,dP_0(t,x,\delta)\\
&\quad+\int_{\hat F_{n,\hat\beta_n}(t-\hat\beta_n'x)\in[\epsilon,1-\epsilon]}\Big\{x-\phi_{\hat\beta_n}(t-\hat\beta_n'x)\Big\}\Big\{F_0(t-\beta_0'x)-F_0(t-\hat\beta_n'x)\Big\}\,dP_0(t,x,\delta),
\end{aligned}$$

and next we get, using the definition of ϕ_β given in (S2.14), for the first integral on the right-hand side of the last display:

$$\begin{aligned}
&\int_{\hat F_{n,\hat\beta_n}(t-\hat\beta_n'x)\in[\epsilon,1-\epsilon]}\Big\{x-\phi_{\hat\beta_n}(t-\hat\beta_n'x)\Big\}\Big\{F_0(t-\hat\beta_n'x)-F_{\hat\beta_n}(t-\hat\beta_n'x)\Big\}\,dP_0(t,x,\delta)\\
&=\int_{\hat F_{n,\hat\beta_n}(u)\in[\epsilon,1-\epsilon]}\Big\{x-\phi_{\hat\beta_n}(u)\Big\}\Big\{F_0(u)-F_{\hat\beta_n}(u)\Big\}f_{T-\hat\beta_n'X}(u)\,f_{X|T-\hat\beta_n'X}(x|u)\,du\,dx\\
&=0.
\end{aligned}$$

Furthermore, we get by expanding $F_0(t-\beta'x)$ and by the continuity of $\beta\mapsto\phi_\beta(u)$ at $\beta=\beta_0$

$$\begin{aligned}
&\int_{\hat F_{n,\hat\beta_n}(t-\hat\beta_n'x)\in[\epsilon,1-\epsilon]}\Big\{x-\phi_{\hat\beta_n}(t-\hat\beta_n'x)\Big\}\Big\{F_0(t-\beta_0'x)-F_0(t-\hat\beta_n'x)\Big\}\,dP_0(t,x,\delta)\\
&=\int_{\hat F_{n,\hat\beta_n}(t-\hat\beta_n'x)\in[\epsilon,1-\epsilon]}\Big\{x-\phi_{\hat\beta_n}(t-\hat\beta_n'x)\Big\}(\hat\beta_n-\beta_0)'xf_0(t-\beta_0'x)\,dP_0(t,x,\delta)\\
&\qquad+o_p\left(\hat\beta_n-\beta_0\right)\\
&=\left\{\int_{\hat F_{n,\hat\beta_n}(t-\hat\beta_n'x)\in[\epsilon,1-\epsilon]}\Big\{x-\phi_0(t-\beta_0'x)\Big\}x'f_0(t-\beta_0'x)\,dP_0(t,x,\delta)\right\}(\hat\beta_n-\beta_0)\\
&\qquad+o_p\left(\hat\beta_n-\beta_0\right).
\end{aligned}$$

Finally we get from the consistency of $\hat F_{n,\hat\beta_n}$:

$$\begin{aligned}
&\left\{\int_{\hat F_{n,\hat\beta_n}(t-\hat\beta_n'x)\in[\epsilon,1-\epsilon]}\Big\{x-\phi_0(t-\beta_0'x)\Big\}x'f_0(t-\beta_0'x)\,dP_0(t,x,\delta)\right\}\left(\hat\beta_n-\beta_0\right)\\
&=\left\{\int_{F_0(t-\beta_0'x)\in[\epsilon,1-\epsilon]}\Big\{x-\phi_0(t-\beta_0'x)\Big\}x'f_0(t-\beta_0'x)\,dP_0(t,x,\delta)\right\}\left(\hat\beta_n-\beta_0\right)\\
&\qquad+o_p\left(\hat\beta_n-\beta_0\right)\\
&=\psi_{1,\epsilon}'(\beta_0)(\hat\beta_n-\beta_0)+o_p\left(\hat\beta_n-\beta_0\right).
\end{aligned}$$

So we obtain:

$$I_a = \int_{\hat F_{n,\hat\beta_n}(t-\hat\beta_n' x)\in[\epsilon,1-\epsilon]} \left\{x-\phi_{\hat\beta_n}(t-\hat\beta_n' x)\right\}\left\{\delta - F_{\hat\beta_n}(t-\hat\beta_n' x)\right\} d(\mathbb{P}_n - P_0)(t,x,\delta)$$
$$+ \psi_{1,\epsilon}'(\beta_0)(\hat\beta_n-\beta_0) + o_p\left(\hat\beta_n-\beta_0\right).$$

We now proceed again as before, and define $\mathcal{K}_1'$ to be the set of functions

$$\mathcal{K}_1' = \big\{(t,x,\delta)\mapsto \big(x-\phi_\beta(t-\beta' x)\big)\big(\delta - F_\beta(t-\beta' x)\big)1_{[\epsilon,1-\epsilon]}(F(t-\beta' x))$$
$$: F\in\mathcal{F},\, \beta\in\Theta\big\}.$$

We add the function

$$(t,x,\delta)\mapsto \big(x-\phi_0(t-\beta_0' x)\big)\big(\delta - F_0(t-\beta_0' x)\big)1_{[\epsilon,1-\epsilon]}(F_0(t-\beta_0' x))$$

to the set $\mathcal{K}_1'$. The distance d is defined by (S2.4) again, with $\mathcal{K}$ replaced by $\mathcal{K}_1'$. We therefore get, similarly as before, using Lemma S1.1,

$$\sup_{\zeta>0} \zeta H_B\left(\zeta,\mathcal{K}_1', L_2(P_0)\right) = O(1),$$

which implies:

$$\int_0^\zeta H_B\left(u,\mathcal{K}_1', L_2(P_0)\right)^{1/2}\,du = O\left(\zeta^{1/2}\right),\qquad \zeta>0.$$

Moreover, we get:

$$\big(x-\phi_{\hat\beta_n}(t-\hat\beta_n' x)\big)\big(\delta - F_{\hat\beta_n}(t-\hat\beta_n' x)\big)1_{[\epsilon,1-\epsilon]}\left(\hat F_{n,\hat\beta_n}(t-\hat\beta_n' x)\right)$$
$$- \big(x-\phi_0(t-\beta_0' x)\big)\big(\delta - F_0(t-\beta_0' x)\big)1_{[\epsilon,1-\epsilon]}\left(F_0(t-\beta_0' x)\right)$$
$$= \Big\{\big(x-\phi_{\hat\beta_n}(t-\hat\beta_n' x)\big)\big(\delta - F_{\hat\beta_n}(t-\hat\beta_n' x)\big)$$
$$- \big(x-\phi_0(t-\beta_0' x)\big)\big(\delta - F_0(t-\beta_0' x)\big)\Big\}1_{[\epsilon,1-\epsilon]}\left(\hat F_{n,\hat\beta_n}(t-\hat\beta_n' x)\right)$$
$$+ \big(x-\phi_0(t-\beta_0' x)\big)\big(\delta - F_0(t-\beta_0' x)\big)$$
$$\cdot\left\{1_{[\epsilon,1-\epsilon]}\left(F_0(t-\beta_0' x)\right) - 1_{[\epsilon,1-\epsilon]}\left(\hat F_{n,\hat\beta_n}(t-\hat\beta_n' x)\right)\right\}$$
$$= A_n(t,x,\delta) + B_n(t,x,\delta),$$

implying

$$\int\Big\{\big(x-\phi_{\hat\beta_n}(t-\hat\beta_n' x)\big)\big(\delta - F_{\hat\beta_n}(t-\hat\beta_n' x)\big)1_{[\epsilon,1-\epsilon]}\left(\hat F_{n,\hat\beta_n}(t-\hat\beta_n' x)\right)$$
$$- \big(x-\phi_0(t-\beta_0' x)\big)\big(\delta - F_0(t-\beta_0' x)\big)1_{[\epsilon,1-\epsilon]}\left(F_0(t-\beta_0' x)\right)\Big\}^2 dP_0(t,x,\delta)$$
$$\le 2\int \left\{A_n(t,x,\delta)^2 + B_n(t,x,\delta)^2\right\}\,dP_0(t,x,\delta) = o_p(1),$$

since the integrals w.r.t. A_n^2 and B_n^2 tends to zero using the consistency of $\hat\beta_n$ and $\hat F_{n,\hat\beta_n}$.

Hence we get from Lemma S0.1:

$$I_a = \int_{F_0(t-\beta_0'x)\in[\epsilon,1-\epsilon]} \left\{x - \phi_0(t-\beta_0'x)\right\}\left\{\delta - F_0(t-\beta_0'x)\right\} d(\mathbb{P}_n - P_0)(t,x,\delta)$$
$$+ \psi_{1,\epsilon}'(\beta_0)(\hat\beta_n - \beta_0) + o_p\left(\hat\beta_n - \beta_0\right) + o_p\left(n^{-1/2}\right).$$

This means that we get the conclusion

$$\int_{F_0(t-\beta_0'x)\in[\epsilon,1-\epsilon]} \left\{x - \phi_0(t-\beta_0'x)\right\}\left\{\delta - F_0(t-\beta_0'x)\right\} d\left(\mathbb{P}_n - P_0\right)(t,x,\delta)$$
$$= -\psi_{1,\epsilon}'(\beta_0)(\hat\beta_n - \beta_0) + o_p\left(\hat\beta_n - \beta_0\right) + o_p\left(n^{-1/2}\right), \tag{S2.20}$$

if we can show that I_b is negligible.

Since, by definition of ϕ_β given in (S2.14),

$$\int_{\hat F_{n,\hat\beta_n}(t-\hat\beta_n'x)\in[\epsilon,1-\epsilon]} \left\{x - \phi_{\hat\beta_n}(t-\hat\beta_n'x)\right\} f_{X|T-\hat\beta_n'X}(x|t-\hat\beta_n'x)\,dx = 0,$$

we have

$$I_b = \int_{\hat F_{n,\hat\beta_n}(t-\hat\beta_n'x)\in[\epsilon,1-\epsilon]} \left\{x - \phi_{\hat\beta_n}(t-\hat\beta_n'x)\right\}$$
$$\cdot \left\{F_{\hat\beta_n}(t-\hat\beta_n'x) - \hat F_{n,\hat\beta_n}(t-\hat\beta_n'x)\right\} d(\mathbb{P}_n - P_0)(t,x,\delta).$$

The negligibility of I_b now follows in the same way as (S2.18), using the parametric nature of the function ϕ_β and the entropy properties of the class of functions

$$u \mapsto \hat F_{n,\hat\beta_n}(u) - F_{\hat\beta_n}(u).$$

The conclusion now follows from (S2.20). $\square$

REMARK S2.1. Note that the proof above yields the representation

$$\hat\beta_n - \beta_0$$
$$\sim n^{-1}\psi_{1,\epsilon}'(\beta_0)^{-1}\sum_{i=1}^n (X_i - \mathbb{E}(X_i|T-\beta_0'X))\left\{\Delta_i - F_0(T_i - \beta_0'X_i\right\},$$

where $\psi_{1,\epsilon}'(\beta_0)$ is given by (4.4).

S3. Asymptotic behavior of the efficient estimate based on the MLE $\hat F_{n,\beta}$. In this section we prove the asymptotic efficiency of the score estimator defined in Section 4.2. The proof of existence of the root and the consistency proof for the score estimator is similar to the proof of existence and consistency of the first score estimator defined in Section 4.1, thus omitted.

S3.1. *Asymptotic normality of the efficient score estimator.*

PROOF OF THEOREM 4.2 (ASYMPTOTIC NORMALITY). Since the proof is very similar to the proof of Theorem 4.1, we only give the main steps of the proof. As in the proof of Theorem 4.1, we can define $\psi_{2,nh}^{(\epsilon)}$ at $\hat\beta_n$ by

$$\psi_{2,nh}^{(\epsilon)}(\hat\beta_n)=0,$$

and $\psi_{2,nh}^{(\epsilon)}(\hat\beta_n)$ is then a combination of one-sided limits at $\hat\beta_n$.

We prove that:

$$\begin{aligned}\psi_{2,nh}^{(\epsilon)}(\hat\beta_n)&\\ =\int_{F_0(t-\beta_0'x)\in[\epsilon,1-\epsilon]}&\frac{\{xf_0(t-\beta_0'x)-\varphi_{\beta_0}(t-\beta_0'x)\}\{\delta-F_0(t-\beta_0'x)\}}{F_0(t-\beta_0'x)\{1-F_0(t-\beta_0'x)\}}\,d\mathbb{P}_n(t,x,\delta)\\ &+\psi_{2,\epsilon}'(\beta_0)(\hat\beta_n-\beta_0)+o_p\left(n^{-1/2}+(\hat\beta_n-\beta_0)\right),\end{aligned}\tag{S3.1}$$

where φ_β is defined by

$$\varphi_\beta(t-\beta'x)=\mathbb{E}(X|T-\beta'X=t-\beta'x)f_\beta(t-\beta'x),\tag{S3.2}$$

and $\psi_{2,\epsilon}$ is defined by,

$$\begin{aligned}\psi_{2,\epsilon}(\beta)&\\ =\int_{F_\beta(t-\beta'x)\in[\epsilon,1-\epsilon]}&\frac{\Big\{xf_\beta(t-\beta'x)-\varphi_\beta(t-\beta'x)\Big\}\Big\{\delta-F_\beta(t-\beta'x)\Big\}}{F_\beta(t-\beta'x)\{1-F_\beta(t-\beta'x)\}}\,dP_0(t,x,\delta).\end{aligned}\tag{S3.3}$$

Straightforward calculations show that,

$$\begin{aligned}\psi_{2,\epsilon}'(\beta_0)&=\int_{F_0(t-\beta_0'x)\in[\epsilon,1-\epsilon]}\frac{\{xf_0(t-\beta_0'x)-\varphi_{\beta_0}(t-\beta_0'x)\}^2}{F_0(t-\beta_0'x)\{1-F_0(t-\beta_0'x)\}}\,dP_0(t,x,\delta)\\ &=\mathbb{E}_\epsilon\left\{\frac{f_0(T-\beta_0'X)^2\,\{X-\mathbb{E}(X|T-\beta_0'X)\}\{X-\mathbb{E}(X|T-\beta_0'X)\}'}{F_0(T-\beta_0'X)\{1-F_0(T-\beta_0'X)\}}\right\}\\ &=I_\epsilon(\beta_0).\end{aligned}$$

(See also the derivation of the derivative ψ_ϵ' for the first score equation in the proof of Theorem 4.1, Part 1). Since

$$\begin{aligned}&\sqrt{n}\int_{F_0(t-\beta_0'x)\in[\epsilon,1-\epsilon]}\frac{\{xf_0(t-\beta_0'x)-\varphi_{\beta_0}(t-\beta_0'x)\}\{\delta-F_0(t-\beta_0'x)\}}{F_0(t-\beta_0'x)\{1-F_0(t-\beta_0'x)\}}\,d\mathbb{P}_n(t,x,\delta)\\ &\stackrel{d}{\to}N(0,I_\epsilon(\beta_0)),\end{aligned}$$

(S3.1) implies, using the non-singularity of $\psi'_{2,\epsilon}(\beta_0)$ and the consistency of $\hat\beta_n$,

$$
\begin{aligned}
&\sqrt{n}(\hat\beta_n-\beta_0)\\
&=-\psi'_{2,\epsilon}(\beta_0)^{-1}\Big\{\sqrt{n}\int_{F_0(t-\beta_0'x)\in[\epsilon,1-\epsilon]}\frac{xf_0(t-\beta_0'x)-\varphi_{\beta_0}(t-\beta_0'x)}{F_0(t-\beta_0'x)\{1-F_0(t-\beta_0'x)\}}\\
&\qquad\qquad\cdot\left\{\delta-F_0(t-\beta_0'x)\right\}\,d\mathbb{P}_n(t,x,\delta)\Big\}\\
&\quad+o_p(1+\sqrt{n}(\hat\beta_n-\beta_0))\\
&\xrightarrow{d} N\left(0,I_\epsilon(\beta_0)^{-1}\right).
\end{aligned}
$$

Let, analogously to the start of the proof of Theorem 4.1, $\bar\varphi_{\hat\beta_n,\hat F_{n,\hat\beta_n}}$ be a (random) piecewise constant version of $\varphi_{\hat\beta_n}$, where, for a piecewise constant distribution function F with finitely many jumps at $\tau_1<\tau_2<\dots$, the function $\bar\varphi_{\beta,F}$ is defined in the following way.

$$
\bar\varphi_{\beta,F}(u)=\begin{cases}\varphi_\beta(\tau_i), & \text{if } F_\beta(u)>F(\tau_i),\ u\in[\tau_i,\tau_{i+1}),\\ \varphi_\beta(s), & \text{if } F_\beta(u)=F(s),\ \text{for some } s\in[\tau_i,\tau_{i+1}),\\ \varphi_\beta(\tau_{i+1}), & \text{if } F_\beta(u)<F(\tau_i),\ u\in[\tau_i,\tau_{i+1}).\end{cases}\tag{S3.4}
$$

We now have:

$$
\begin{aligned}
&\psi^{(\epsilon)}_{2,nh}(\beta)\\
&=\int_{\hat F_{n,\hat\beta_n}(t-\hat\beta_n'x)\in[\epsilon,1-\epsilon]} xf_{nh,\hat\beta_n}(t-\hat\beta_n'x)\frac{\delta-\hat F_{n,\hat\beta_n}(t-\hat\beta_n'x)}{\hat F_{n,\hat\beta_n}(t-\hat\beta_n'x)\{1-\hat F_{n,\hat\beta_n}(t-\hat\beta_n'x)\}}\,d\mathbb{P}_n(t,x,\delta)\\
&\quad=\int_{\hat F_{n,\hat\beta_n}(t-\hat\beta_n'x)\in[\epsilon,1-\epsilon]}\left\{xf_{nh,\hat\beta_n}(t-\hat\beta_n'x)-\varphi_{\hat\beta_n}(t-\hat\beta_n'x)\right\}\\
&\qquad\qquad\cdot\frac{\delta-\hat F_{n,\hat\beta_n}(t-\hat\beta_n'x)}{\hat F_{n,\hat\beta_n}(t-\hat\beta_n'x)\{1-\hat F_{n,\hat\beta_n}(t-\hat\beta_n'x)\}}\,d\mathbb{P}_n(t,x,\delta)\\
&\qquad+\int_{\hat F_{n,\hat\beta_n}(t-\hat\beta_n'x)\in[\epsilon,1-\epsilon]}\left\{\varphi_{\hat\beta_n}(t-\hat\beta_n'x)-\bar\varphi_{n,\hat F_{n,\hat\beta_n}}(t-\hat\beta_n'x)\right\}\\
&\qquad\qquad\cdot\frac{\delta-\hat F_{n,\hat\beta_n}(t-\hat\beta_n'x)}{\hat F_{n,\hat\beta_n}(t-\hat\beta_n'x)\{1-\hat F_{n,\hat\beta_n}(t-\hat\beta_n'x)\}}\,d\mathbb{P}_n(t,x,\delta)\\
&\quad=I+II.
\end{aligned}
$$

Let $\mathcal{F}$ be the set of piecewise constant distribution functions with finitely many jumps (like the MLE $\hat F_{n,\hat\beta_n}$), and let $\mathcal{K}_2$ be the set of functions

$$
\begin{aligned}
\mathcal{K}_2=\Big\{&(t,x,\delta)\mapsto\left\{\varphi_\beta(t-\beta'x)-\bar\varphi_{\beta,F}(t-\beta'x)\right\}\frac{\delta-F(t-\beta'x))}{F(t-\beta'x)\{1-F(t-\beta'x)\}}\\
&\cdot 1_{[\epsilon,1-\epsilon]}(F(t-\beta'x)):F\in\mathcal{F},\,\beta\in\Theta\Big\},
\end{aligned}\tag{S3.5}
$$

where $\bar\varphi_{\beta,F}$ is defined by (S3.4). We add the function which is identically zero to $\mathcal{K}_2$. As in the proof of Theorem 4.1, the functions are uniformly bounded and also of uniformly bounded variation, using conditions (A4) and (A5). For k_1 and k_2 in $\mathcal{K}_2$, we define

$$d(k_1,k_2)^2=\int \|k_1-k_2\|^2\,dP_0,\qquad k_1,k_2\in\mathcal{K}_2. \tag{S3.6}$$

For this distance, we therefore get, similarly as before, using Lemma S1.1,

$$\sup_{\zeta>0}\zeta H_B\left(\zeta,\mathcal{K}_2,L_2(P_0)\right)=O(1),$$

which implies:

$$\int_0^\zeta H_B\left(u,\mathcal{K}_2,L_2(P_0)\right)^{1/2}\,du=O\left(\zeta^{1/2}\right),\qquad \zeta>0.$$

Note that the indicator function keeps $F(t-\beta'x)$ away from zero and one, which is essential for getting the bounded variation property.

Following the same steps as in the proof of Theorem 4.1, we get:

$$II=o_p\left(n^{-1/2}+\hat\beta_n-\beta_0\right).$$

We now write,

$$\begin{aligned}
I&=\int_{\hat F_{n,\hat\beta_n}(t-\hat\beta_n'x)\in[\epsilon,1-\epsilon]}\left\{xf_{nh,\hat\beta_n}(t-\hat\beta_n'x)-\varphi_{\hat\beta_n}(t-\hat\beta_n'x)\right\}\\
&\qquad\qquad\cdot\frac{\delta-\hat F_{n,\hat\beta_n}(t-\hat\beta_n'x)}{\hat F_{n,\hat\beta_n}(t-\hat\beta_n'x)\{1-\hat F_{n,\hat\beta_n}(t-\hat\beta_n'x)\}}\,d\mathbb{P}_n(t,x,\delta)\\
&=\int_{\hat F_{n,\hat\beta_n}(t-\hat\beta_n'x)\in[\epsilon,1-\epsilon]}\left\{xf_{nh,\hat\beta_n}(t-\hat\beta_n'x)-\varphi_{\hat\beta_n}(t-\hat\beta_n'x)\right\}\\
&\qquad\qquad\cdot\frac{\delta-F_{\hat\beta_n}(t-\hat\beta_n'x)}{\hat F_{n,\hat\beta_n}(t-\hat\beta_n'x)\{1-\hat F_{n,\hat\beta_n}(t-\hat\beta_n'x)\}}\,d\mathbb{P}_n(t,x,\delta)\\
&\quad+\int_{\hat F_{n,\hat\beta_n}(t-\hat\beta_n'x)\in[\epsilon,1-\epsilon]}\left\{xf_{nh,\hat\beta_n}(t-\hat\beta_n'x)-\varphi_{\hat\beta_n}(t-\hat\beta_n'x)\right\}\\
&\qquad\qquad\cdot\frac{F_{\hat\beta_n}(t-\hat\beta_n'x)-\hat F_{n,\hat\beta_n}(t-\hat\beta_n'x)}{\hat F_{\hat\beta_n}(t-\hat\beta_n'x)\{1-\hat F_{n,\hat\beta_n}(t-\hat\beta_n'x)\}}\,dP_0(t,x,\delta)\\
&\quad+\int_{\hat F_{n,\hat\beta_n}(t-\hat\beta_n'x)\in[\epsilon,1-\epsilon]}\left\{xf_{nh,\hat\beta_n}(t-\hat\beta_n'x)-\varphi_{\hat\beta_n}(t-\hat\beta_n'x)\right\}\\
&\qquad\qquad\cdot\frac{F_{\hat\beta_n}(t-\hat\beta_n'x)-\hat F_{n,\hat\beta_n}(t-\hat\beta_n'x)}{\hat F_{\hat\beta_n}(t-\hat\beta_n'x)\{1-\hat F_{n,\hat\beta_n}(t-\hat\beta_n'x)\}}\,d(\mathbb{P}_n-P_0)(t,x,\delta)\\
&=I_a+I_b+I_c.
\end{aligned}$$

For the term I_b we get:

$$\int_{\hat F_{n,\hat\beta_n}(t-\hat\beta_n'x)\in[\epsilon,1-\epsilon]}\Bigl\{xf_{nh,\hat\beta_n}(t-\hat\beta_n'x)-\varphi_{\hat\beta_n}(t-\hat\beta_n'x)\Bigr\}$$

$$\cdot\frac{F_{\hat\beta_n}(t-\hat\beta_n'x)-\hat F_{n,\hat\beta_n}(t-\hat\beta_n'x)}{\hat F_{\hat\beta_n}(t-\hat\beta_n'x)\{1-\hat F_{n,\hat\beta_n}(t-\hat\beta_n'x)\}}\,dP_0(t,x,\delta)$$

$$=\int_{\hat F_{n,\hat\beta_n}(t-\hat\beta_n'x)\in[\epsilon,1-\epsilon]}\Bigl\{x-\mathbb{E}(X|T-\hat\beta_n'X=t-\hat\beta_n'x)\Bigr\}\,f_{nh,\hat\beta_n}(t-\hat\beta_n'x)$$

$$\cdot\frac{F_{\hat\beta_n}(t-\hat\beta_n'x)-\hat F_{n,\hat\beta_n}(t-\hat\beta_n'x)}{\hat F_{\hat\beta_n}(t-\hat\beta_n'x)\{1-\hat F_{n,\hat\beta_n}(t-\hat\beta_n'x)\}}\,dP_0(t,x,\delta)$$

$$+\int_{\hat F_{n,\hat\beta_n}(t-\hat\beta_n'x)\in[\epsilon,1-\epsilon]}\Bigl\{f_{nh,\hat\beta_n}(t-\hat\beta_n'x)-f_{\hat\beta_n}(t-\hat\beta_n'x)\Bigr\}\mathbb{E}(X|T-\hat\beta_n'X=t-\hat\beta_n'x)$$

$$\cdot\frac{F_{\hat\beta_n}(t-\hat\beta_n'x)-\hat F_{n,\hat\beta_n}(t-\hat\beta_n'x)}{\hat F_{\hat\beta_n}(t-\hat\beta_n'x)\{1-\hat F_{n,\hat\beta_n}(t-\hat\beta_n'x)\}}\,dP_0(t,x,\delta)$$

$$=\int_{\hat F_{n,\hat\beta_n}(t-\hat\beta_n'x)\in[\epsilon,1-\epsilon]}\Bigl\{f_{nh,\hat\beta_n}(t-\hat\beta_n'x)-f_{\hat\beta_n}(t-\hat\beta_n'x)\Bigr\}\mathbb{E}(X|T-\hat\beta_n'X=t-\hat\beta_n'x)$$

$$\cdot\frac{F_{\hat\beta_n}(t-\hat\beta_n'x)-\hat F_{n,\hat\beta_n}(t-\hat\beta_n'x)}{\hat F_{\hat\beta_n}(t-\hat\beta_n'x)\{1-\hat F_{n,\hat\beta_n}(t-\hat\beta_n'x)\}}\,dP_0(t,x,\delta)$$

$$=\int_{\hat F_{n,\hat\beta_n}(u)\in[\epsilon,1-\epsilon]}\Bigl\{f_{nh,\hat\beta_n}(u)-f_{\hat\beta_n}(u)\Bigr\}\mathbb{E}(X|T-\hat\beta_n'X=u)\frac{F_{\hat\beta_n}(u)-\hat F_{n,\hat\beta_n}(u)}{\hat F_{\hat\beta_n}(u)\{1-\hat F_{n,\hat\beta_n}u)\}}\,f_{T-\hat\beta_n'X}(u)\,du.$$

Furthermore,

$$=\int_{\hat F_{n,\hat\beta_n}(u)\in[\epsilon,1-\epsilon]}\Bigl\{f_{nh,\hat\beta_n}(u)-f_{\hat\beta_n}(u)\Bigr\}\mathbb{E}(X|T-\hat\beta_n'X=u)\frac{F_{\hat\beta_n}(u)-\hat F_{n,\hat\beta_n}(u)}{\hat F_{\hat\beta_n}(u)\{1-\hat F_{n,\hat\beta_n}u)\}}\,f_{T-\hat\beta_n'X}(u)\,du.$$

$$=h^{-2}\int_{\hat F_{n,\hat\beta_n}(u)\in[\epsilon,1-\epsilon]}\Bigl\{\int K'((u-v)/h)\hat F_{n,\hat\beta_n}(v)\,dv-f_{\hat\beta_n}(u)\Bigr\}\mathbb{E}(X|T-\hat\beta_n'X=u)$$

$$\cdot\frac{F_{\hat\beta_n}(u)-\hat F_{n,\hat\beta_n}(u)}{\hat F_{\hat\beta_n}(u)\{1-\hat F_{n,\hat\beta_n}(u)\}}\,f_{T-\hat\beta_n'X}(u)\,du$$

$$=h^{-2}\int_{\hat F_{n,\hat\beta_n}(u)\in[\epsilon,1-\epsilon]}\int K'((u-v)/h)\,\Bigl\{\hat F_{n,\hat\beta_n}(v)-F_{\hat\beta_n}(v)\Bigr\}\,dv\,\mathbb{E}(X|T-\hat\beta_n'X=u)$$

$$\cdot\frac{F_{\hat\beta_n}(u)-\hat F_{n,\hat\beta_n}(u)}{\hat F_{\hat\beta_n}(u)\{1-\hat F_{n,\hat\beta_n}(u)\}}\,f_{T-\hat\beta_n'X}(u)\,du$$

$$+\int_{\hat F_{n,\hat\beta_n}(u)\in[\epsilon,1-\epsilon]}\Bigl\{\int K_h(u-v)\,dF_{\hat\beta_n}(v)-f_{\hat\beta_n}(u)\Bigr\}\,\mathbb{E}(X|T-\hat\beta_n'X=u)$$

$$\cdot\frac{F_{\hat\beta_n}(u)-\hat F_{n,\hat\beta_n}(u)}{\hat F_{\hat\beta_n}(u)\{1-\hat F_{n,\hat\beta_n}(u)\}}\,f_{T-\hat\beta_n'X}(u)\,du.$$

The last term on the right-hand side has an upper bound of order $O_p(n^{-2/7-1/3}) = O_p(n^{-13/21}) = o_p(n^{-1/2})$, since

$$\left\{\int_{\hat F_{n,\hat\beta_n}(u)\in[\epsilon,1-\epsilon]}\left\{\int K_h(u-v)\,dF_{\hat\beta_n}(v)-f_{\hat\beta_n}(u)\right\}^2 du\right\}^{1/2} = O_p\left(n^{-2/7}\right),$$

and

$$\left\{\int_{\hat F_{n,\hat\beta_n}(u)\in[\epsilon,1-\epsilon]}\left\{\hat F_{n,\hat\beta_n}(u)-F_{\hat\beta_n}(u)\right\}^2 du\right\}^{1/2} = O_p\left(n^{-1/3}\right), \tag{S3.7}$$

using Lemma 3.1 for the last relation. We also use the Cauchy-Schwarz inequality.

The first term on the right is of order $O_p(n^{1/7-2/3}) = O_p(n^{-11/21}) = o_p(n^{-1/2})$ by (S3.7) and using

$$\begin{aligned}
&\left|h^{-2}\int_{\hat F_{n,\hat\beta_n}(u)\in[\epsilon,1-\epsilon]}\int K'((u-v)/h)\left\{\hat F_{n,\hat\beta_n}(v)-F_{\hat\beta_n}(v)\right\}dv\,\mathbb{E}(X|T-\hat\beta_n'X=u)\right.\\
&\qquad\qquad\left.\cdot\frac{F_{\hat\beta_n}(u)-\hat F_{n,\hat\beta_n}(u)}{\hat F_{\hat\beta_n}(u)\{1-\hat F_{n,\hat\beta_n}(u)\}}\,f_{T-\hat\beta_n'X}(u)\,du\right|\\
&= h^{-1}\left|\int_{\hat F_{n,\hat\beta_n}(u)\in[\epsilon,1-\epsilon]}\int K'(w)\left\{\hat F_{n,\hat\beta_n}(u-hw)-F_{\hat\beta_n}(u-hw)\right\}dw\,\mathbb{E}(X|T-\hat\beta_n'X=u)\right.\\
&\qquad\qquad\left.\cdot\frac{F_{\hat\beta_n}(u)-\hat F_{n,\hat\beta_n}(u)}{\hat F_{\hat\beta_n}(u)\{1-\hat F_{n,\hat\beta_n}(u)\}}\,f_{T-\hat\beta_n'X}(u)\,du\right|\\
&\le ch^{-1}\int_{\hat F_{n,\hat\beta_n}(u)\in[\epsilon/2,1-\epsilon/2]}\left\{\hat F_{n,\hat\beta_n}(u)-F_{\hat\beta_n}(u)\right\}^2 du,
\end{aligned}$$

for small h and a constant $c > 0$, where we first use Fubini's theorem and next the Cauchy-Schwarz inequality in the last inequality, together with

$$\begin{aligned}
&\int_{\hat F_{n,\hat\beta_n}(u)\in[\epsilon,1-\epsilon]}\left\{\hat F_{n,\hat\beta_n}(u-hw)-F_{\hat\beta_n}(u-hw)\right\}^2 du\\
&\le \int_{\hat F_{n,\hat\beta_n}(u)\in[\epsilon/2,1-\epsilon/2]}\left\{\hat F_{n,\hat\beta_n}(u)-F_{\hat\beta_n}(u)\right\}^2 du,
\end{aligned}$$

for small $h > 0$, together with $w \in [-1, 1]$. Finally we use Lemma 3.1 again.

For the term I_c we argue similarly as before using Lemma S0.1 that,

$$I_c = o_p\left(n^{-1/2}\right).$$

Finally,

$$I_a = \int_{\hat F_{n,\hat\beta_n}(t-\hat\beta_n' x)\in[\epsilon,1-\epsilon]} \left\{x f_{nh,\hat\beta_n}(t-\hat\beta_n' x) - \varphi_{\hat\beta_n}(t-\hat\beta_n' x)\right\} \cdot \frac{\delta - \hat F_{n,\hat\beta_n}(t-\hat\beta_n' x)}{\hat F_{n,\hat\beta_n}(t-\hat\beta_n' x)\{1-\hat F_{n,\hat\beta_n}(t-\hat\beta_n' x)\}}\, d\bigl(\mathbb{P}_n - P_0\bigr)(t,x,\delta)$$
$$+ \int_{\hat F_{n,\hat\beta_n}(t-\hat\beta_n' x)\in[\epsilon,1-\epsilon]} \left\{x f_{nh,\hat\beta_n}(t-\hat\beta_n' x) - \varphi_{\hat\beta_n}(t-\hat\beta_n' x)\right\} \cdot \frac{F_0(t-\beta_0' x) - \hat F_{n,\hat\beta_n}(t-\hat\beta_n' x)}{\hat F_{n,\hat\beta_n}(t-\hat\beta_n' x)\{1-\hat F_{n,\hat\beta_n}(t-\hat\beta_n' x)\}}\, dP_0(t,x,\delta).$$

This time we consider the class of functions

$$\mathcal{K}_2' = \left\{(t,x,\delta)\mapsto \bigl(xf(t-\beta' x) - \varphi_\beta(t-\beta' x)\bigr)\frac{\delta - F(t-\beta' x)}{F(t-\beta' x)\{1-F(t-\beta' x)\}} 1_{[\epsilon,1-\epsilon]}(F(t-\beta' x)) : F\in\mathcal{F},\ f\in\mathcal{F}',\ \beta\in\Theta\right\},$$

where $\mathcal{F}'$ is a class of uniformly bounded functions of uniformly bounded variation (which have the interpretation of estimates of F_β'), to which we add the function

$$(t,x,\delta)\mapsto \bigl(xf_0(t-\beta_0' x) - \varphi_{\beta_0}(t-\beta_0' x)\bigr)\frac{\delta - F_0(t-\beta_0' x)}{F_0(t-\beta_0' x)\{1-F_0(t-\beta_0' x)\}} 1_{[\epsilon,1-\epsilon]}(F_0(t-\beta_0' x)).$$

So we get, using Lemma S1.1,

$$\sup_{\zeta>0} \zeta H_B\left(\zeta, \mathcal{K}_2', L_2(P_0)\right) = O(1),$$

which implies:

$$\int_0^\zeta H_B\left(u,\mathcal{K}_2', L_2(P_0)\right)^{1/2} du = O\left(\zeta^{1/2}\right), \qquad \zeta>0.$$

As before, we now get:

$$\int_{\hat F_{n,\hat\beta_n}(t-\hat\beta_n' x)\in[\epsilon,1-\epsilon]} \left\{x f_{nh,\hat\beta_n}(t-\hat\beta_n' x) - \varphi_{\hat\beta_n}(t-\hat\beta_n' x)\right\} \cdot \frac{\delta - \hat F_{n,\hat\beta_n}(t-\hat\beta_n' x)}{\hat F_{n,\hat\beta_n}(t-\hat\beta_n' x)\{1-\hat F_{n,\hat\beta_n}(t-\hat\beta_n' x)\}}\, d\bigl(\mathbb{P}_n - P_0\bigr)(t,x,\delta)$$
$$= \int_{F_0(t-\beta_0' x)\in[\epsilon,1-\epsilon]} \left\{x f_0(t-\beta_0' x) - \varphi_{\beta_0}(t-\beta_0' x)\right\} \cdot \frac{\delta - F_0(t-\beta_0' x)}{F_0(t-\beta_0' x)\{1-F_0(t-\beta_0' x)\}}\, d\bigl(\mathbb{P}_n - P_0\bigr)(t,x,\delta)$$
$$+ o_p\left(n^{-1/2} + \hat\beta_n - \beta_0\right)$$

and

$$\int_{\hat F_{n,\hat\beta_n}(t-\hat\beta_n' x)\in[\epsilon,1-\epsilon]} \left\{x f_{nh,\hat\beta_n}(t-\hat\beta_n' x)-\varphi_{\hat\beta_n}(t-\hat\beta_n' x)\right\} \cdot \frac{F_0(t-\beta_0' x)-\hat F_{n,\hat\beta_n}(t-\hat\beta_n' x)}{\hat F_{n,\hat\beta_n}(t-\hat\beta_n' x)\{1-\hat F_{n,\hat\beta_n}(t-\hat\beta_n' x)\}}\, dP_0(t,x,\delta)$$
$$= \left\{\int_{F_0(t-\beta_0' x)\in[\epsilon,1-\epsilon]} \left\{x f_0(t-\beta_0' x)-\varphi_{\beta_0}(t-\beta_0' x)\right\} \cdot \frac{f_0(t-\beta_0' x)x'}{F_0(t-\beta_0' x)\{1-F_0(t-\beta_0' x)\}}\, dP_0(t,x,\delta)\right\}\left(\hat\beta_n-\beta_0\right) + o_p\left(n^{-1/2}+\hat\beta_n-\beta_0\right).$$

The result now follows. □

S4. Asymptotic behavior of the plug-in estimator. In this section we first sketch in Section S4.1 the proof of consistency of the plug-in estimator, denoted by $\hat\beta_n$. This is the second result stated in Theorem 4.3. The proof of existence of a root is similar to the proof of existence of a root of the simple score estimator defined in Section 4.1 and omitted. We next prove the asymptotic normality result of the plug-in estimator, which is the third result given in Theorem 4.3. The proof of Theorem 4.5 on the asymptotic representation of the plug-in estimator as a sum of i.i.d. random variables follows from the proof of 4.3. The asymptotic distribution of the estimator of the intercept, given in Theorem 5.1, is proved in Section S4.2.

Before we start the proofs, we give some auxiliary results on the L_2-distance between the plug-in estimate $F_{nh,\beta}$ and F_β and between the partial derivative of the plug-in estimate $\partial_\beta F_{nh,\beta}$ and $\partial_\beta F_\beta$ in Lemma S4.1. For simplicity, we derive the proof of Lemma S4.1 for the one-dimensional case and let $\Theta=[\beta_0-\eta,\beta_0+\eta]$ for some $\eta>0$. The higher-dimensional extension of the one-dimensional proof is straightforward. Next, we follow the arguments used to prove the asymptotic normality of the estimators defined in Theorem 4.1 and Theorem 4.2 and give a similar proof for the limiting distribution of the plug-in estimator.

LEMMA S4.1. *Let the conditions of Theorem 4.3 be satisfied and let $k=1$. Let the function F_β be defined by (3.2). Then we have, for the estimate $F_{nh,\beta}$, defined by (4.8),*

$$\int_{F_{nh,\beta}(t-\beta x)\in[\epsilon,1-\epsilon]} \{F_{nh,\beta}(t-\beta x)-F_\beta(t-\beta x)\}^2\, dG(t,x) = O_p\left(\frac{1}{nh}\right)+O_p\left(h^4\right), \tag{S4.1}$$

$$\int_{F_{nh,\beta}(t-\beta x)\in[\epsilon,1-\epsilon]} \{\partial_\beta F_{nh,\beta}(t-\beta x)-\partial_\beta F_\beta(t-\beta x)\}^2\, dG(t,x) = O_p\left(\frac{1}{nh^3}\right)+O_p\left(h^2\right) \tag{S4.2}$$

uniformly in $\beta\in[\beta_0-\eta,\beta_0+\eta]$. The results remain valid when dG in (S4.1) or (S4.2) is replaced by $d\mathbb{G}_n$.

PROOF OF LEMMA S4.1. We first prove the first part and show that (S4.1) holds. Recall that,

$$F_{nh,\beta}(t-\beta x)=\frac{g_{nh,1,\beta}(t-\beta x)}{g_{nh,\beta}(t-\beta x)}$$

where

$$g_{nh,1,\beta}(t-\beta x)=\int \delta K_h(t-\beta x-u+\beta y)\,d\mathbb{P}_n(u,y,\delta),$$

and

$$g_{nh,\beta}(t-\beta x)=\int K_h(t-\beta x-u+\beta y)\,d\mathbb{P}_n(u,y,\delta).$$

Moreover,

$$F_\beta(t-\beta x)=\int F_0(t-\beta_0 x+(\beta-\beta_0)(y-x))f_{X|T-\beta X}(y|t-\beta x)\,dy.$$

We first investigate the bias part.

$$\begin{aligned}
&\mathbb{E}g_{nh,1,\beta}(t-\beta x)=\int F_0(u-\beta_0 y)K_h(t-\beta x-u+\beta y)\,dG(u,y)\\
&=\int F_0(v+(\beta-\beta_0)y)K_h(t-\beta x-v)\,f_{T-\beta X}(v)\,f_{X|T-\beta X}(y|v)\,dy\,dv\\
&=\int F_0(t-\beta x+(\beta-\beta_0)y-hw)K(w)\,f_{T-\beta X}(t-\beta x-hw)\,f_{X|T-\beta X}(y|t-\beta x-hw)\,dy\,dw\\
&=f_{T-\beta X}(t-\beta x)\int F_0(t-\beta_0 x+(\beta-\beta_0)(y-x))f_{X|T-\beta X}(y|t-\beta x)\,dy+O\left(h^2\right),
\end{aligned}$$

uniformly in $\beta\in[\beta_0-\eta,\beta_0+\eta]$ and t,x varying over a finite interval, due to the assumptions of Theorem 4.3. In a similar way, we get

$$\mathbb{E}g_{nh,\beta}(t-\beta x)=f_{T-\beta X}(t-\beta x)+O\left(h^2\right),$$

uniformly in $\beta\in[\beta_0-\eta,\beta_0+\eta]$ and t,x varying over a finite interval. So we find:

$$\frac{\mathbb{E}g_{nh,1,\beta}(t-\beta x)}{\mathbb{E}g_{nh,\beta}(t-\beta x)}=F_\beta(t-\beta x)+O\left(h^2\right).$$

uniformly in $\beta\in[\beta_0-\eta,\beta_0+\eta]$ and t,x varying over a finite interval, such that $\mathbb{E}g_{nh,1,\beta}(t-\beta x)$ stays away from zero.

So we obtain

$$\begin{aligned}
&F_{nh,\beta}(t-\beta x)-F_\beta(t-\beta x)\\
&=\frac{g_{nh,1,\beta}(t-\beta x)-\mathbb{E}g_{nh,1,\beta}(t-\beta x)}{g_{nh,\beta}(t-\beta x)}+\mathbb{E}g_{nh,1,\beta}(t-\beta x)\frac{\mathbb{E}g_{nh,\beta}(t-\beta x)-g_{nh,\beta}(t-\beta x)}{g_{nh,\beta}(t-\beta x)\mathbb{E}g_{nh,\beta}(t-\beta x)}+O\left(h^2\right),
\end{aligned}$$

and

$$\{F_{nh,\beta}(t-\beta x)-F_\beta(t-\beta x)\}^2$$
$$\le 3\left\{\frac{g_{nh,1,\beta}(t-\beta x)-\mathbb{E}g_{nh,1,\beta}(t-\beta x)}{g_{nh,\beta}(t-\beta x)}\right\}^2+3\left\{\mathbb{E}g_{nh,1,\beta}(t-\beta x)\frac{\mathbb{E}g_{nh,\beta}(t-\beta x)-g_{nh,\beta}(t-\beta x)}{g_{nh,\beta}(t-\beta x)\mathbb{E}g_{nh,\beta}(t-\beta x)}\right\}^2$$
$$+O\left(h^4\right). \tag{S4.3}$$

uniformly in $\beta\in[\beta_0-\eta,\beta_0+\eta]$ and t,x varying over a finite interval, such that $\mathbb{E}g_{nh,1,\beta}(t-\beta x)$ stays away from zero.

Since $\eta>0$ is chosen in such a way that $a_1(\beta)=F_\beta^{-1}(\epsilon)>a$, $b_1(\beta)=F_\beta^{-1}(1-\epsilon)<b$, for each $\beta\in[\beta_0-\eta,\beta_0+\eta]$ and since $g_{nh,\beta}$ stays away from zero with probability tending to one if $\epsilon<F_{nh,\beta}(t-\beta x)<1-\epsilon$ we get

$$\int_{F_{nh,\beta}(t-\beta x)\in[\epsilon,1-\epsilon]}\left\{\frac{g_{nh,1,\beta}(t-\beta x)-\mathbb{E}g_{nh,1,\beta}(t-\beta x)}{g_{nh,\beta}(t-\beta x)}\right\}^2\,dG(t,x)$$
$$\lesssim\int_{F_{nh,\beta}(t-\beta x)\in[\epsilon,1-\epsilon]}\{g_{nh,1,\beta}(t-\beta x)-\mathbb{E}g_{nh,1,\beta}(t-\beta x)\}^2\,dG(t,x)$$

Furthermore

$$\mathbb{E}\{g_{nh,1,\beta}(t-\beta x)-\mathbb{E}g_{nh,1,\beta}(t-\beta x)\}^2=\mathbb{E}\left\{\int\delta K_h(t-\beta x-u+\beta y)\,d(\mathbb{P}_n-P_0)(u,y,\delta)\right\}^2$$
$$=O\left(\frac{1}{nh}\right),$$

uniformly for (t,x) in a bounded region, so we get

$$\mathbb{E}\int_{F_{nh,\beta}(t-\beta x)\in[\epsilon,1-\epsilon]}\{g_{nh,1,\beta}(t-\beta x)-\mathbb{E}g_{nh,1,\beta}(t-\beta x)\}^2\,dG(t,x)=O\left(\frac{1}{nh}\right).$$

Hence

$$\int_{F_{nh,\beta}(t-\beta x)\in[\epsilon,1-\epsilon]}\left\{\frac{g_{nh,1,\beta}(t-\beta x)-\mathbb{E}g_{nh,1,\beta}(t-\beta x)}{g_{nh,\beta}(t-\beta x)}\right\}^2\,dG(t,x)=O_p\left(\frac{1}{nh}\right).$$

The second term on the right-hand side of (S4.3) can be treated in a similar way. So we get (S4.1). This proves (S4.1).

We next replace dG in part (S4.1) by $d\mathbb{G}_n$ and we get

$$\int_{F_{nh,\beta}(t-\beta x)\in[\epsilon,1-\epsilon]}\left\{\frac{g_{nh,1,\beta}(t-\beta x)-\mathbb{E}g_{nh,1,\beta}(t-\beta x)}{g_{nh,\beta}(t-\beta x)}\right\}^2\,d\mathbb{G}_n(t,x)$$
$$\lesssim\int_{F_{nh,\beta}(t-\beta x)\in[\epsilon,1-\epsilon]}\{g_{nh,1,\beta}(t-\beta x)-\mathbb{E}g_{nh,1,\beta}(t-\beta x)\}^2\,d\mathbb{G}_n(t,x)$$
$$=\frac{1}{n}\sum_{i=1}^n\{g_{nh,1,\beta}(T_i-\beta X_i)-\mathbb{E}g_{nh,1,\beta}(T_i-\beta X_i)\}^2\,1_{\{\epsilon<F_{nh,\beta}(T_i-\beta X_i)<1-\epsilon\}}.$$

Moreover,

$$
\begin{aligned}
&\mathbb{E}\frac{1}{n}\sum_{i=1}^{n}\left\{g_{nh,1,\beta}(T_i-\beta X_i)-\mathbb{E}g_{nh,1,\beta}(T_i-\beta X_i)\right\}^2 1_{\{\epsilon<F_{nh,\beta}(T_i-\beta X_i)<1-\epsilon\}}\\
&=\mathbb{E}\left\{g_{nh,1,\beta}(T_1-\beta X_1)-\mathbb{E}g_{nh,1,\beta}(T_1-\beta X_1)\right\}^2 1_{\{\epsilon<F_{nh,\beta}(T_1-\beta X_1)<1-\epsilon\}}\\
&\lesssim \mathbb{E}\int_{\epsilon/2<F_\beta(t-\beta x)<1-\epsilon/2}\left\{g_{nh,1,\beta}(t-\beta x)-\mathbb{E}g_{nh,1,\beta}(t-\beta x)\right\}^2\,dG(t,x)\\
&=O\left(\frac{1}{nh}\right).
\end{aligned}
$$

This implies by the Markov inequality,

$$
\int_{F_{nh,\beta}(t-\beta x)\in[\epsilon,1-\epsilon]}\left\{\frac{g_{nh,1,\beta}(t-\beta x)-\mathbb{E}g_{nh,1,\beta}(t-\beta x)}{g_{nh,\beta}(t-\beta x)}\right\}^2\,d\mathbb{G}_n(t,x)=O_p\left(\frac{1}{nh}\right).
$$

The other term on the right-hand side of (S4.3) is treated similarly; and the result of (S4.1) also follows when we replace dG by $d\mathbb{G}_n$.

We next continue with the proof of (S4.2).

We have:

$$
\partial_\beta F_{nh,\beta}(t-\beta x)=\frac{\int(y-x)\{\delta-F_{nh,\beta}(t-\beta x)\}K_h'(t-\beta x-u+\beta y)\,d\mathbb{P}_n(u,y,\delta)}{g_{nh,\beta}(t-\beta x)}\,. \tag{S4.4}
$$

We consider the numerator of (S4.4). It can be rewritten as

$$
\begin{aligned}
&\int(y-x)\{\delta-F_0(u-\beta_0 y)\}K_h'(t-\beta x-u+\beta y)\,d\mathbb{P}_n(u,y,\delta)\\
&\quad+\int(y-x)\{F_0(u-\beta_0 y)-F_\beta(t-\beta x)\}K_h'(t-\beta x-u+\beta y)\,d\mathbb{G}_n(u,y)\\
&\quad+\{F_\beta(t-\beta x)-F_{nh,\beta}(t-\beta x)\}\int(y-x)K_h'(t-\beta x-u+\beta y)\,d\mathbb{G}_n(u,y).
\end{aligned}
$$

The first term can be written as

$$
A_n(t,x,\beta)\stackrel{\text{def}}{=}\int(y-x)\{\delta-F_0(u-\beta_0 y)\}K_h'(t-\beta x-u+\beta y)\,d\bigl(\mathbb{P}_n-P_0\bigr)(u,y,\delta),
$$

and we have:

$$
\begin{aligned}
&\mathbb{E}\int_{F_{nh,\beta}(t-\beta x)\in[\epsilon,1-\epsilon]}A_n(t,x,\beta)^2\,dG(t,x)\le\mathbb{E}\int A_n(t,x,\beta)^2\,dG(t,x)\\
&\sim\frac{1}{nh^3}\int \mathrm{var}(X|v)F_0(v)\{1-F_0(v)\}f_{T-\beta X}(v)\,dv\int K'(u)^2\,du,\ n\to\infty.
\end{aligned}
$$

In the second term we must compare $F_0(u-\beta_0 y)$ with

$$
F_\beta(t-\beta x)=\int F_0(t-\beta_0 x+(\beta-\beta_0)(z-x))f_{X|T-\beta X}(z|t-\beta x)\,dz.
$$

We can write

$$
\begin{aligned}
&F_0(u-\beta_0 y)-F_\beta(t-\beta x)\\
&=\int\left\{F_0(u-\beta_0 y)-F_0(t-\beta_0 x+(\beta-\beta_0)(z-x))\right\}f_{X|T-\beta X}(z|t-\beta x)\,dz.
\end{aligned}
$$

So we find for the second term

$$
\begin{aligned}
B_n(t,x,\beta)&\stackrel{\text{def}}{=}\int (y-x)\left\{F_0(u-\beta_0 y)-F_\beta(t-\beta x)\right\}K_h'(t-\beta x-u+\beta y)\,d\mathbb{G}_n(u,y)\\
&=\int\int (y-x)\left\{F_0(u-\beta_0 y)-F_0(t-\beta_0 x+(\beta-\beta_0)(z-x))\right\}f_{X|T-\beta X}(z|t-\beta x)\,dz\\
&\qquad\qquad\cdot K_h'(t-\beta x-u+\beta y)\,d\mathbb{G}_n(u,y)\\
&=\int (y-x)\int\left\{F_0(u-\beta_0 y)-F_0(t-\beta_0 x+(\beta-\beta_0)(z-x))\right\}f_{X|T-\beta X}(z|t-\beta x)\,dz\\
&\qquad\qquad\cdot K_h'(t-\beta x-u+\beta y)\,dG(u,y)\\
&\quad+\int (y-x)\int\left\{F_0(u-\beta_0 y)-F_0(t-\beta_0 x+(\beta-\beta_0)(z-x))\right\}f_{X|T-\beta X}(z|t-\beta x)\,dz\\
&\qquad\qquad\cdot K_h'(t-\beta x-u+\beta y)\,d\big(\mathbb{G}_n-G\big)(u,y)\\
&=f_{T-\beta X}(t-\beta x)\partial_\beta F_\beta(t-\beta x)+O(h)+O_p\left(\frac{1}{nh^3}\right).
\end{aligned}
$$

where, using integration by parts, the last line follows by straightforward calculation. Since

$$
g_{nh,\beta}(t-\beta x)=f_{T-\beta X}(t-\beta x)+O_p(h^2),
$$

we get,

$$
\int_{F_{nh,\beta}(t-\beta x)\in[\epsilon,1-\epsilon]}\left\{\frac{B_n(t,x,\beta)}{g_{nh,\beta}(t-\beta x)}-\partial_\beta F_\beta(t-\beta x)\right\}^2 dG(t,x)=O_p\left(\frac{1}{nh^3}\right)+O_p\left(h^2\right).
$$

Finally, defining

$$
C_n(t,x,\beta)\stackrel{\text{def}}{=}\left\{F_\beta(t-\beta x)-F_{nh,\beta}(t-\beta x)\right\}\int (y-x)K_h'(t-\beta x-u+\beta y)\,d\mathbb{G}_n(u,y),
$$

we get, using,

$$
\begin{aligned}
&\int (y-x)K_h'(t-\beta x-u+\beta y)\,d\mathbb{G}_n(u,y)\\
&=\int (y-x)K_h'(t-\beta x-u+\beta y)\,dG(u,y)+\int (y-x)K_h'(t-\beta x-u+\beta y)\,d(\mathbb{G}_n-G)(u,y)\\
&=\int (y-x)K_h'(t-\beta x-v)f_{T-\beta X}(v)f_{X|T-\beta X}(y|v)\,dv\,dy+O_p\left(\frac{1}{nh^3}\right)\\
&=\int (y-x)K_h(t-\beta x-v)\frac{d}{dv}\left\{f_{T-\beta X}(v)f_{X|T-\beta X}(y|v)\right\}\,dv\,dy+O_p\left(\frac{1}{nh^3}\right)\\
&=O_p(1),
\end{aligned}
$$

and using the first part of Lemma S4.1 for the factor $F_\beta(t-\beta x)-F_{nh,\beta}(t-\beta x)$ that

$$\int_{F_{nh,\beta}(t-\beta x)\in[\epsilon,1-\epsilon]} C_n(t,x,\beta)^2\,dG(t,x)=O_p\left(\frac{1}{nh}\right)+O_p\left(h^4\right).$$

This proves (S4.2). The second part of the result, replacing dG by $d\mathbb{G}_n$ in (S4.2) is proved in the same way as the second part of (S4.1). □

S4.1. *Consistency and asymptotic normality of the plug-in estimator.* We first prove that $\hat\beta_n$ is a consistent estimate of β_0.

PROOF OF THEOREM 4.3, PART 1 (CONSISTENCY OF THE PLUG-IN ESTIMATOR). We assume that $\hat\beta_n$ is contained in the compact set Θ, and hence the sequence $(\hat\beta_n)$ has a subsequence $(\hat\beta_{n_k}=\hat\beta_{n_k}(\omega))$, converging to an element β_*. It is easily seen that, if $\hat\beta_{n_k}=\hat\beta_{n_k}(\omega)\longrightarrow\beta_*$, we get:

$$F_{n_kh,\hat\beta_{n_k}}(t-\hat\beta_{n_k}'x)\longrightarrow F_{\beta_*}(t-\beta_*'x)\stackrel{\text{def}}{=}\int F_0(t-\beta_*'x+(\beta_*-\beta_0)'y)f_{X|T-\beta_*'X}(y|t-\beta_*'x)\,dy.$$

In the limit we get therefore the relation

(S4.5)

$$\lim_{k\to\infty}-(\hat\beta_{n_k}-\beta_0)'\int_{F_{n_kh,\hat\beta_{n_k}}(t-\hat\beta_{n_k}'x)\in[\epsilon,1-\epsilon]}\frac{\{\delta-F_{n_kh,\hat\beta_{n_k}}(t-\hat\beta_{n_k}'x)\}\partial_\beta F_{n_kh,\beta}(t-\beta'x)\big|_{\beta=\hat\beta_{n_k}}}{F_{n_k,\hat\beta_{n_k}}(t-\hat\beta_{n_k}'x)\{1-F_{n_k,\hat\beta_{n_k}}(t-\hat\beta_{n_k}'x)\}}\,d\mathbb{P}_{n_k}(t,x,\delta)$$

(S4.6)

$$=-(\beta_*-\beta_0)'\int_{F_{\beta_*}(t-\beta_*'x)\in[\epsilon,1-\epsilon]}\frac{\{F_0(t-\beta_0'x)-F_{\beta_*}(t-\beta_*'x)\}\partial_\beta F_\beta(t-\beta'x)|_{\beta=\beta_*}}{F_{\beta_*}(t-\beta_*'x)\{1-F_{\beta_*}(t-\beta_*'x)\}}\,dG(t,x)=0,$$

which can only mean $\beta_*=\beta_0$ by condition (4.14).

□

We next continue with the proof of the asymptotic normality of the plug-in estimator.

PROOF OF THEOREM 4.3, PART 2 (ASYMPTOTIC NORMALITY OF THE PLUG-IN ESTIMATOR). To prove the asymptotic normality of the plug-in estimator, we follow the reasoning of the corresponding proofs of the simple score estimator and the efficient score estimator described in Section 4.1 and Section 4.2. We prove that,

(S4.7)

$$\begin{aligned}&\psi_{3,nh}^{(\epsilon)}(\hat\beta_n)\\&=\int_{F_0(t-\beta_0'x)\in[\epsilon,1-\epsilon]}\frac{\{E(X|T-\beta_0'X=t-\beta_0'x)-x\}\,f_0(t-\beta_0'x)\,\{\delta-F_0(t-\beta_0'x)\}}{F_0(t-\beta_0'x)\{1-F_0(t-\beta_0'x)\}}\,d\mathbb{P}_n(t,x,\delta)\\&\quad+\psi_{3,\epsilon}'(\beta_0)(\hat\beta_n-\beta_0)+o_p\left(n^{-1/2}+(\hat\beta_n-\beta_0)\right),\end{aligned}$$

where $\psi_{3,\epsilon}$ is defined by,

$$\psi_{3,\epsilon}(\beta) = \int_{F_\beta(t-\beta'x)\in[\epsilon,1-\epsilon]} \frac{\partial_\beta F_\beta(t-\beta x)\Big\{\delta - F_\beta(t-\beta' x)\Big\}}{F_\beta(t-\beta'x)\{1-F_\beta(t-\beta'x)\}}\, dP_0(t,x,\delta), \tag{S4.8}$$

and

$$\psi'_{3,\epsilon}(\beta_0) = -\mathbb{E}_\epsilon\left\{\frac{f_0(T-\beta_0'X)^2\;\{X-\mathbb{E}(X|T-\beta_0'X)\}\{X-\mathbb{E}(X|T-\beta_0'X)\}'}{F_0(T-\beta_0'X)\{1-F_0(T-\beta_0'X)\}}\right\} = -I_\epsilon(\beta_0),$$

which follows by straightforward calculations after noting that,

$$\begin{aligned}\partial_\beta F_\beta(t-\beta'x) &= \int (y-x) f_0(t-\beta_0'x+(\beta-\beta_0)'(y-x)) f_{X|T-\beta'X}(y|T-\beta'X=t-\beta'x)dy\\ &\quad + \int F_0(t-\beta_0'x+(\beta-\beta_0)'(y-x))\partial_\beta f_{X|T-\beta'X}(y|T-\beta'X=t-\beta'x)\,dG(t,x)\end{aligned}$$

is, at $\beta=\beta_0$ equal to

$$f_0(t-\beta_0'x)\mathbb{E}\left\{X-x|T-\beta_0'X=t-\beta_0'x\right\}.$$

We have,

$$\begin{aligned}&\psi_{3,nh}^{(\epsilon)}(\hat\beta_n)\\ &= \int_{F_{nh,\hat\beta_n}(t-\hat\beta_n'x)\in[\epsilon,1-\epsilon]} \partial_\beta F_{nh,\beta}(t-\beta'x)\mid_{\beta=\hat\beta_n} \frac{\delta - F_{nh,\hat\beta_n}(t-\hat\beta_n'x)}{F_{nh,\hat\beta_n}(t-\hat\beta_n'x)\{1-F_{nh,\hat\beta_n}(t-\hat\beta_n'x)\}}\, d\mathbb{P}_n(t,x,\delta)\\ &= \int_{F_{nh,\hat\beta_n}(t-\hat\beta_n'x)\in[\epsilon,1-\epsilon]} \partial_\beta F_\beta(t-\beta'x)\mid_{\beta=\hat\beta_n} \frac{\delta - F_{nh,\hat\beta_n}(t-\hat\beta_n'x)}{F_{nh,\hat\beta_n}(t-\hat\beta_n'x)\{1-F_{nh,\hat\beta_n}(t-\hat\beta_n'x)\}}\, d\mathbb{P}_n(t,x,\delta)\\ &\qquad + \int_{F_{nh,\hat\beta_n}(t-\hat\beta_n'x)\in[\epsilon,1-\epsilon]} \Big\{\partial_\beta F_{nh,\beta}(t-\beta'x)\mid_{\beta=\hat\beta_n} - \partial_\beta F_\beta(t-\beta'x)\mid_{\beta=\hat\beta_n}\Big\}\\ &\qquad\qquad \cdot \frac{\delta - F_{nh,\hat\beta_n}(t-\hat\beta_n'x)}{F_{nh,\hat\beta_n}(t-\hat\beta_n'x)\{1-F_{nh,\hat\beta_n}(t-\hat\beta_n'x)\}}\, d\mathbb{P}_n(t,x,\delta)\\ &\quad = I + II\end{aligned}$$

Let $\mathcal{F}$ be a class of functions with the property that

$$\int_{\epsilon/2<F_\beta(u)<1-\epsilon/2} f'(u)^2\,du \le M.$$

if $f\in\mathcal{F}$, for a fixed $M>0$. Using Proposition 5.1.9, p. 393 in [1], with $m=1$, $p=2$ and $h \asymp n^{-1/5}$, we may assume that the functions $u \to F_{nh,\beta}(u)$ and $u \to \partial_\beta F_{nh,\beta}(u)$ belong to $\mathcal{F}$. Since the plug-in estimates are monotonically increasing with probability tending to one we get that the function

$$(t,x) \mapsto 1_{[\epsilon,1-\epsilon]}\left(F_{nh,\beta}(t-\beta'x)\right),$$

can be written in the form

$$(t,x) \mapsto 1_{\left[a_{\epsilon,F_{nh,\beta}},b_{\epsilon,F_{nh,\beta}}\right]}(t-\beta'x) = 1_{\left[a_{\epsilon,F_{nh,\beta}},\infty\right)}(t-\beta'x) - 1_{\left(b_{\epsilon,F_{nh,\beta}},\infty\right)}(t-\beta'x),$$

for $a_{\epsilon,F_{nh,\beta}} \le b_{\epsilon,F_{nh,\beta}}$ for large n, with probability tending to one. The function is therefore of uniformly bounded variation for n sufficiently large (see also the proofs of Theorems 4.1 and 4.2). It now follows that the bracketing $\zeta-$entropy $H_B(\zeta,\mathcal{K}_3,L_2(P_0))$ for the class $\mathcal{K}_3$ of functions consisting of the function which is identically zero and the functions

$$\mathcal{K}_3 = \Bigg\{(t,x,\delta) \mapsto \left\{\partial_\beta F_{nh,\beta}(t-\beta'x) - \partial_\beta F_\beta(t-\beta'x)\right\}\frac{\delta - F(t-\beta'x)}{F(t-\beta'x)\{1-F(t-\beta'x)\}}$$

(S4.9)

$$\cdot\, 1_{[\epsilon,1-\epsilon]}(F_{nh,\beta}(t-\beta'x)) : F \in \mathcal{F},\ \beta \in \Theta\Bigg\},$$

w.r.t. the L_2-distance satisfies:

$$\sup_{\zeta>0} \zeta H_B\left(\zeta,\mathcal{K}_3,L_2(P_0)\right) = O(1),$$

which implies:

$$\int_0^\zeta H_B\left(u,\mathcal{K}_3,L_2(P_0)\right)^{1/2}\,du = O\left(\zeta^{1/2}\right), \qquad \zeta > 0.$$

Moreover, by Lemma S4.1 we also have,

$$\int_{F_{nh,\beta}(t-\beta'x)\in[\epsilon,1-\epsilon]}\Bigg\{\left\{\partial_\beta F_{nh,\beta}(t-\beta'x) - \partial_\beta F_\beta(t-\beta'x)\right\}$$
$$\cdot\,\frac{\delta - F_{nh,\beta}(t-\beta'x)}{F_{nh,\beta}(t-\beta'x)\{1-F_{nh,\beta}(t-\beta'x)\}}\Bigg\}^2 dP_0(t,x,\delta) \xrightarrow{p} 0.$$

This implies by an application of Lemma S0.1, that,

$$\int_{F_{nh,\hat\beta_n}(t-\hat\beta_n'x)\in[\epsilon,1-\epsilon]}\Bigg\{\left\{\partial_\beta F_{nh,\beta}(t-\beta'x)\,|_{\beta=\hat\beta_n} - \partial_\beta F_\beta(t-\beta'x)\,|_{\beta=\hat\beta_n}\right\}$$
$$\cdot\,\frac{\delta - F_{nh,\hat\beta_n}(t-\hat\beta_n'x)}{F_{nh,\hat\beta_n b}(t-\hat\beta_n'x)\{1-F_{nh,\hat\beta_n}(t-\hat\beta_n'x)\}}\Bigg\}\, d(\mathbb{P}_n - P_0)(t,x,\delta) = o_p\left(n^{-1/2}\right)$$

Furthermore, an application of the Cauchy-Schwarz inequality and Lemma S4.1 yield that

$$\sqrt{n}\int_{F_{nh,\hat\beta_n}(t-\beta x)\in[\epsilon,1-\epsilon]}\left\{\partial_\beta F_\beta(t-\beta'x)\,|_{\beta=\hat\beta_n} - \partial_\beta F_{nh,\beta}(t-\beta'x)\,|_{\beta=\hat\beta_n}\right\}$$
$$\cdot\left\{\frac{F_0(t-\beta_0'x) - F_{nh,\hat\beta_n}(t-\hat\beta_n'x)}{F_{nh,\hat\beta_n}(t-\hat\beta_n'x)\{1-F_{nh,\hat\beta_n}(t-\hat\beta_n'x)\}}\right\} dP_0(t,x,\delta)$$
$$= O_p\left(n^{-1/10}\right) + o_p\left(\sqrt{n}(\hat\beta_n - \beta_0)\right)$$

The conclusion is that

$$II = o_p\left(n^{-1/2} + (\hat\beta_n - \beta_0)\right)$$

We now write:

$$\begin{aligned}
I &= \int_{F_{nh,\hat\beta_n}(t-\hat\beta_n' x)\in[\epsilon,1-\epsilon]} \partial_\beta F_\beta(t-\beta' x)\,|_{\beta=\hat\beta_n} \\
&\qquad\qquad \cdot \frac{\delta - F_{nh,\hat\beta_n}(t-\hat\beta_n' x)}{F_{nh,\hat\beta_n}(t-\hat\beta_n' x)\{1-F_{nh,\hat\beta_n}(t-\hat\beta_n' x)\}}\, d\mathbb{P}_n(t,x,\delta) \\
&= \int_{F_{nh,\hat\beta_n}(t-\hat\beta_n' x)\in[\epsilon,1-\epsilon]} \partial_\beta F_\beta(t-\beta' x)\,|_{\beta=\hat\beta_n} \\
&\qquad\qquad \cdot \frac{\delta - F_{\hat\beta_n}(t-\hat\beta_n' x)}{F_{n,\hat\beta_n}(t-\hat\beta_n' x)\{1-F_{n,\hat\beta_n}(t-\hat\beta_n' x)\}}\, d\mathbb{P}_n(t,x,\delta) \\
&\quad + \int_{F_{nh,\hat\beta_n}(t-\hat\beta_n' x)\in[\epsilon,1-\epsilon]} \partial_\beta F_\beta(t-\beta' x)\,|_{\beta=\hat\beta_n} \\
&\qquad\qquad \cdot \frac{F_{\hat\beta_n}(t-\hat\beta_n' x) - F_{n,\hat\beta_n}(t-\hat\beta_n' x)}{\hat F_{\hat\beta_n}(t-\hat\beta_n' x)\{1-F_{n,\hat\beta_n}(t-\hat\beta_n' x)\}}\, d\mathbb{P}_n(t,x,\delta) \\
&= I_a + I_b.
\end{aligned}$$

We now get, using Lemma S4.1 and

$$\partial_\beta F_\beta(t-\beta' x)\,|_{\beta=\hat\beta_n} = \mathbb{E}(X-x|T-\hat\beta_n' X = t-\hat\beta_n' x) f_0(t-\hat\beta_n' x) + O_p\left(\hat\beta_n - \beta_0\right),$$

that $I_b = o_p\left(n^{-1/2} + \hat\beta_n - \beta_0\right)$. The result of Theorem 4.3 now follows by showing that,

$$\begin{aligned}
&\int_{F_{nh,\hat\beta_n}(t-\hat\beta_n' x)\in[\epsilon,1-\epsilon]} \partial_\beta F_\beta(t-\beta' x)\,|_{\beta=\hat\beta_n} \\
&\qquad\qquad \cdot \frac{\delta - F_{\hat\beta_n}(t-\hat\beta_n' x)}{F_{nh,\hat\beta_n}(t-\hat\beta_n' x)\{1-F_{nh,\hat\beta_n}(t-\hat\beta_n' x)\}}\, d(\mathbb{P}_n - P_0)(t,x,\delta) \\
&= \int_{F_0(t-\beta_0' x)\in[\epsilon,1-\epsilon]} \partial_\beta F_\beta(t-\beta' x)\,|_{\beta=\beta_0} \\
&\qquad\qquad \cdot \frac{\delta - F_0(t-\beta_0' x)}{F_0(t-\beta_0' x)\{1-F_0(t-\beta_0' x)\}}\, d(\mathbb{P}_n - P_0)(t,x,\delta)] \\
&\quad + o_p\left(n^{-1/2} + \hat\beta_n - \beta_0\right)
\end{aligned} \tag{S4.10}$$

and,

$$\begin{aligned}
&\int_{F_{nh,\hat\beta_n}(t-\hat\beta_n' x)\in[\epsilon,1-\epsilon]} \partial_\beta F_\beta(t-\beta' x)\mid_{\beta=\hat\beta_n} \\
&\qquad\qquad \cdot \frac{\delta - F_{\hat\beta_n}(t-\hat\beta_n' x)}{F_{nh,\hat\beta_n}(t-\hat\beta_n' x)\{1-F_{nh,\hat\beta_n}(t-\hat\beta_n' x)\}}\, dP_0(t,x,\delta) \qquad \text{(S4.11)}\\
&= \psi_{3,\epsilon}'(\beta_0)(\hat\beta_n-\beta_0) + o_p\left(n^{-1/2}+\hat\beta_n-\beta_0\right).
\end{aligned}$$

The proof of (S4.10) and (S4.11) is similar to the proof of the corresponding steps given in the proof of Theorem 4.1 and omitted. □

REMARK S4.1. It follows from the proof of Theorem 4.3 that

$$\begin{aligned}
&\sqrt{n} I_\epsilon(\beta_0)(\hat\beta_n-\beta_0)\\
&\quad = n^{-1/2}\sum_{i=1}^n f_0(T_i-\beta_0' X_i)\{\mathbb{E}(X_i|T_i-\beta_0' X_i)-X_i\}\\
&\qquad\qquad \cdot \frac{\Delta_i - F_0(T_i-\beta_0' X_i)}{F_0(T_i-\beta_0' X_i)\{1-F_0(T_i-\beta_0' X_i)\}} 1_{[\epsilon,1-\epsilon]}\left\{F_0(T_i-\beta_0' X_i)\right\} + o_p(1).
\end{aligned}$$

Therefore the result of Theorem 4.5 follows.

S4.2. *Estimation of the intercept.*

PROOF OF THEOREM 5.1. We will denote $dx_1\dots dx_k$ by dx. We have

$$\begin{aligned}
\hat\alpha_n - \alpha_0 &= \int u\, dF_{nh,\hat\beta_n}(u) - \int u\, dF_0(u) = \int \left\{F_0(u) - F_{nh,\hat\beta_n}(u)\right\} du\\
&= \int \frac{F_0(t-\hat\beta_n' x) - F_{nh,\hat\beta_n}(t-\hat\beta_n' x)}{f_{T-\hat\beta_n' X}(t-\hat\beta_n' x)}\, dG(t,x)\\
&= \int \frac{F_0(t-\hat\beta_n' x) - F_0(t-\beta_0' x)}{f_{T-\hat\beta_n' X}(t-\hat\beta_n' x)}\, dG(t,x) + \int \frac{F_0(t-\beta_0' x) - F_{nh,\hat\beta_n}(t-\hat\beta_n' x)}{f_{T-\hat\beta_n' X}(t-\hat\beta_n' x)}\, dG(t,x) \qquad \text{(S4.12)}
\end{aligned}$$

For the first term in the last expression we get

$$\begin{aligned}
&\int \frac{F_0(t-\hat\beta_n' x) - F_0(t-\beta_0' x)}{f_{T-\hat\beta_n' X}(t-\hat\beta_n' x)}\, dG(t,x)\\
&\quad = \int \{F_0(u) - F_0(u + x'(\hat\beta_n-\beta_0))\} f_{X|T-\hat\beta_n' X}(x|T-\hat\beta_n' X = u)\, du\, dx\\
&\quad \sim -\int x'(\hat\beta_n-\beta_0) f_0(u) f_{X|T-\beta_0' X}(x|T-\beta_0' X = u)\, du\, dx\\
&\quad \sim -\left\{\int \mathbb{E}\{X'|T-\beta_0' X = u\} f_0(u)\, du\right\}(\hat\beta_n-\beta_0)
\end{aligned}$$

This term, multiplied with $\sqrt{n}$, is asymptotically normal, with expectation zero and variance

$$\sigma_1^2 \stackrel{\text{def}}{=} a(\beta_0)' I_\epsilon(\beta_0)^{-1}\, a(\beta_0),$$

where $a(\beta_0)$ is the k-dimensional vector, defined by

$$a(\beta_0) = \int \mathbb{E}\{X|T - \beta_0'X = u\} f_0(u)\, du.$$

For the second term in (S4.12), we first note that,

$$F_{nh,\hat\beta_n}(t - \hat\beta_n'x) - F_0(t - \beta_0'x) = \frac{\int\{\delta - F_0(t-\beta_0'x)\}K_h(t - \hat\beta_n'x - u + \hat\beta_n'y)\, d\mathbb{P}_n(u,y,\delta)}{g_{nh,\hat\beta_n}(t - \hat\beta_n'x)}. \tag{S4.13}$$

We write (S4.13) as the sum of the integral over dP_0 and the integral over $d(\mathbb{P}_n - P_0)$ and show that the contribution of the dP_0 integral, evaluated in (S4.12) is negligible and that the contribution of the $d(\mathbb{P}_n - P_0)$ integral will yield an asymptotic normal distribution.

We have

$$\begin{aligned}
&\int\{\delta - F_0(t-\beta_0'x)\}K_h(t - \hat\beta_n'x - u + \hat\beta_n'y)\, dP_0(u,y,\delta)\\
&\quad= \int\{F_0(u - \beta_0'y) - F_0(t-\beta_0'x)\}K_h(t - \hat\beta_n'x - u + \hat\beta_n'y)\, dG(u,y)\\
&\quad= \int\{F_0(v + (\hat\beta_n - \beta_0)y) - F_0(t-\beta_0'x)\}\}K_h(t - \hat\beta_n'x - v)\\
&\qquad\qquad\qquad\qquad\cdot f_{T-\hat\beta_n'X}(v) f_{X|T-\hat\beta_n'X}(y|T - \hat\beta_n'X = v)\, dv\, dy\\
&\quad= f_{T-\hat\beta_n'X}(t - \hat\beta_n'x)\int\{F_0(t - \hat\beta_n'x + (\hat\beta_n - \beta_0)y) - F_0(t-\beta_0'x)\}\\
&\qquad\qquad\qquad\qquad\cdot f_{X|T-\hat\beta_n'X}(y|T - \hat\beta_n'X = t - \hat\beta_n'x)\, dy + O_p\left(h^2\right)\\
&\quad= f_{T-\hat\beta_n'X}(t - \hat\beta_n'x) f_0(t - \beta_0'x)\big(\hat\beta_n - \beta_0\big)'\mathbb{E}\{X - x|T - \hat\beta_n'X = t - \hat\beta_n'x\}\\
&\qquad\qquad\qquad\qquad + O_p\left(h^2\right) + o_p\big(\|\hat\beta_n - \beta_0\|\big),
\end{aligned}$$

where $\|x\|$ is the euclidean norm of the vector x. Hence we get

$$\begin{aligned}
&\int \frac{\int\{\delta - F_0(t-\beta_0'x)\}K_h(t - \hat\beta_n'x - u + \hat\beta_n'y)\, dP_0(u,y,\delta)}{g_{nh,\hat\beta_n}(t - \hat\beta_n'x) f_{T-\hat\beta_n'X}(t - \hat\beta_n'x)}\, dG(t,x)\\
&\quad= (\hat\beta_n - \beta_0)'\int \frac{f_0(t - \beta_0'x)E\{X - x|T - \beta_0'X = t - \beta_0'x\}}{g_{nh}(t - \hat\beta_n'x)}\, dG(t,x) + O_p\left(h^2\right) + o_p(\hat\beta_n - \beta_0)\\
&\quad= (\hat\beta_n - \beta_0)'\int f_0(v)\mathbb{E}\{X - x|T - \beta_0'X = v\} f_{X|T-\beta_0'X}(x|T - \beta_0'X = v)\, dx\, dv\\
&\qquad\qquad\qquad\qquad + O_p\left(h^2\right) + o_p\big(\|\hat\beta_n - \beta_0\|\big)\\
&\quad= O_p\left(h^2\right) + o_p\big(\|\hat\beta_n - \beta_0\|\big),
\end{aligned}$$

which is $o_p(n^{-1/2})$ if $h \ll n^{-1/4}$.

Finally,

$$\begin{aligned}
&\sqrt{n}\int\frac{\int\{\delta-F_0(t-\beta_0'x)\}K_h(t-\hat\beta_n'x-u+\hat\beta_n'y)\,d\big(\mathbb{P}_n-P_0\big)(u,y,\delta)}{g_{nh,\hat\beta_n}(t-\hat\beta_n'x)f_{T-\hat\beta_n'X}(t-\hat\beta_n'x)}\,dG(t,x)\\
&\quad=\sqrt{n}\iint\frac{\{\delta-F_0(t-\beta_0'x)\}K_h(t-\hat\beta_n'x-u+\hat\beta_n'y)}{g_{nh,\hat\beta_n}(t-\hat\beta_n'x)f_{T-\hat\beta_n'X}(t-\hat\beta_n'x)}\,dG(t,x)\,d\big(\mathbb{P}_n-P_0\big)(u,y,\delta)\\
&\quad=\sqrt{n}\int\frac{\{\delta-F_0(u-\beta_0'y)\}}{f_{T-\beta_0'X}(u-\beta_0'y)}\,d\big(\mathbb{P}_n-P_0\big)(u,y,\delta)+O_p\left(h^2\right)+O_p\big(\|\hat\beta_n-\beta_0\|\big)
\end{aligned}$$

is asymptotically normal, with expectation zero and variance

$$\int\frac{F_0(v)\{1-F_0(v)\}}{f_{T-\beta_0'X}(v)}\,dv, \tag{S4.14}$$

if $h\ll n^{-1/4}$.

Both terms in the representation on the right of (S4.12) are, apart from a negligible contribution, sums of independent variables with expectation zero. By Theorem 4.5 we have

$$\begin{aligned}
&\sqrt{n}(\hat\beta_n-\beta_0)\\
&\quad=n^{-1/2}I_\epsilon(\beta_0)^{-1}\sum_{i=1}^n f_0(T_i-\beta_0X_i)\{\mathbb{E}(X_i|T_i-\beta_0'X_i)-X_i\}\\
&\qquad\qquad\cdot\frac{\Delta_i-F_0(T_i-\beta_0'X_i)}{F_0(T_i-\beta_0'X_i)\{1-F_0(T_i-\beta_0'X_i)\}}1_{[\epsilon,1-\epsilon]}\left\{F_0(T_i-\beta_0'X_i)\right\}+o_p(1).
\end{aligned}$$

and the second term of (S4.12) has the representation

$$n^{-1/2}\sum_{i=1}^n\frac{\Delta_i-F_0(T_i-\beta_0'X_i)}{f_{T-\beta_0'X}(T_i-\beta_0'X_i)}.$$

By the independence of the summands with indices $i\ne j$, the only contribution to the covariance of the two terms in the representation can come from summands with the same index. But,

$$\begin{aligned}
&\mathbb{E}\left\{\frac{f_0(T_i-\beta_0'X_i)\{\mathbb{E}(X_i|T_i-\beta_0'X_i)-X_i\}\{\Delta_i-F_0(T_i-\beta_0'X_i)\}^2}{F_0(T_i-\beta_0'X_i)\{1-F_0(T_i-\beta_0'X_i)\}f_{T-\beta_0'X}(T_i-\beta_0'X_i)}1_{[\epsilon,1-\epsilon]}\{F_0(T_i-\beta_0'X_i)\}\right\}.\\
&=\int_{F_0(u-\beta_0'y)\in[\epsilon,1-\epsilon]}\frac{f_0(u-\beta_0'y)\{\mathbb{E}(X|T-\beta_0'X=u-\beta_0'y)-y\}\{\delta-F_0(u-\beta_0'y)\}^2}{F_0(u-\beta_0'y)\{1-F_0(u-\beta_0'y)\}f_{T-\beta_0'X}(u-\beta_0'y)}dP_0(u,y,\delta)\\
&=\iint_{F_0(v)\in[\epsilon,1-\epsilon]}\frac{f_0(v)\{\mathbb{E}(X|T-\beta_0'X=v)-y\}F_0(v)\{1-F_0(v)\}}{F_0(v)\{1-F_0(v)\}}f_{X|T-\beta_0'X}(y|v)dvdy\\
&=\int_{F_0(v)\in[\epsilon,1-\epsilon]}\int\{\mathbb{E}(X|T-\beta_0'X=v)-y\}f_{X|T-\beta_0'X}(y|v)dy\frac{f_0(v)F_0(v)\{1-F_0(v)\}}{F_0(v)\{1-F_0(v)\}}dv\\
&=0
\end{aligned}$$

So the covariance is zero and Theorem 5.1 follows.

□

References.

[1] GINÉ, E. and NICKL, R. (2015). Mathematical Foundations of Infinite-Dimensional Statistical Models. Cambridge Series in Statistical and Probabilistic Mathematics.

[2] GROENEBOOM, P., JONGBLOED, G. and WITTE, B. I. (2010). Maximum smoothed likelihood estimation and smoothed maximum likelihood estimation in the current status model. *Ann. Statist.* **38** 352–387.

[3] GROENEBOOM, P. and JONGBLOED, G. (2014). *Nonparametric Estimation under Shape Constraints*. Cambridge Univ. Press, Cambridge.

[4] VAN DE GEER, S. A. (2000). *Applications of empirical process theory. Cambridge Series in Statistical and Probabilistic Mathematics* **6**. Cambridge University Press, Cambridge. MR1739079 (2001h:62002)

DELFT UNIVERSITY OF TECHNOLOGY, MEKELWEG 4, 2628 CD DELFT, THE NETHERLANDS.
E-MAIL: P.Groeneboom@tudelft.nl

HASSELT UNIVERSITY, I-BIOSTAT, AGORALAAN, B3590
E-MAIL: kim.hendrickx@uhasselt.be