Supplemental Material to

BOOTSTRAP INFERENCE FOR QUANTILE TREATMENT EFFECTS IN RANDOMIZED EXPERIMENTS WITH MATCHED PAIRS

By

Liang Jiang, Xiaobin Liu, Peter C.B. Phillips, and Yichong Zhang

August 2020

COWLES FOUNDATION DISCUSSION PAPER NO. 2249S

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.yale.edu/

Supplement to "Bootstrap Inference for Quantile Treatment Effects in Randomized Experiments with Matched Pairs"

Abstract

This paper gathers the supplementary material to the original paper. Sections A, B, C, and D contain the proofs of Theorems 3.1, 4.1, 4.2, and 4.3, respectively. Section E contains the proofs of all the lemmas.

Keywords: Bootstrap inference, matched pairs, quantile treatment effect, randomized control trials

JEL codes: C14, C21

A Proof of Theorem 3.1

Let $u = (u_0, u_1)' \in \Re^2$ and

$$L_n(u,\tau) = \sum_{i=1}^{2n} \left[\rho_\tau(Y_i - \dot{A}_i'\beta(\tau) - \dot{A}_i'u/\sqrt{n}) - \rho_\tau(Y_i - \dot{A}_i'\beta(\tau)) \right].$$

Then, by change of variables we have

$$\sqrt{n}(\hat{\beta}(\tau) - \beta(\tau)) = \arg\min_u L_n(u,\tau).$$

Note that $L_n(u,\tau)$ is convex in u for each τ and bounded in τ for each u. We divide the proof into three steps. In Step (1), we show that there exists

$$g_n(u,\tau) = -u'W_n(\tau) + \frac{u'Q(\tau)u}{2}$$

such that for each u,

$$\sup_{\tau \in \Upsilon} |L_n(u,\tau) - g_n(u,\tau)| \xrightarrow{p} 0;$$

and the maximum eigenvalue of $Q(\tau)$ is bounded from above and the minimum eigenvalue of $Q(\tau)$ is bounded away from 0, uniformly over $\tau \in \Upsilon$. In Step (2), we show $W_n(\tau)$ as a stochastic process over $\tau \in \Upsilon$ is tight. Then by Kato (2009, Theorem 2), we have

$$\sqrt{n}(\hat{\beta}(\tau) - \beta(\tau)) = [Q(\tau)]^{-1} W_n(\tau) + r_n(\tau),$$

where $\sup_{\tau \in \Upsilon} ||r_n(\tau)|| = o_p(1)$. Last, in Step (3), we establish weak convergence of $[Q(\tau)]^{-1} W_n(\tau)$ uniformly over $\tau \in \Upsilon$. The second element of the limiting process is $\mathcal{B}(\tau)$ stated in Theorem 3.1.

Step (1). By Knight's identity (Knight, 1998), we have

$$\begin{aligned} & L_n(u,\tau) \\ =& -\sum_{i=1}^{2n} \frac{u'}{\sqrt{n}} \dot{A}_i \left(\tau - 1\{Y_i \leq \dot{A}_i' \beta(\tau)\}\right) + \sum_{i=1}^{2n} \int_0^{\frac{\dot{A}_i' u}{\sqrt{n}}} \left(1\{Y_i - \dot{A}_i'\beta(\tau) \leq v\} - 1\{Y_i - \dot{A}_i'\beta(\tau) \leq 0\}\right) dv \\ \equiv& - u' W_n(\tau) + Q_n(u,\tau), \end{aligned}$$

where

$$W_n(\tau) = \sum_{i=1}^{2n} \frac{1}{\sqrt{n}} \dot{A}_i \left(\tau - 1\{Y_i \leq \dot{A}_i'\beta(\tau)\}\right)$$

and

$$\begin{aligned} Q_n(u,\tau) =& \sum_{i=1}^{2n} \int_0^{\frac{\dot{A}_i' u}{\sqrt{n}}} \left(1\{Y_i - \dot{A}_i'\beta(\tau) \leq v\} - 1\{Y_i - \dot{A}_i'\beta(\tau) \leq 0\}\right) dv \\ =& \sum_{i=1}^{2n} A_i \int_0^{\frac{u_0+u_1}{\sqrt{n}}} \left(1\{Y_i(1) - q_1(\tau) \leq v\} - 1\{Y_i(1) - q_1(\tau) \leq 0\}\right) dv \\ & + \sum_{i=1}^{2n} (1-A_i) \int_0^{\frac{u_0}{\sqrt{n}}} \left(1\{Y_i(0) - q_0(\tau) \leq v\} - 1\{Y_i(0) - q_0(\tau) \leq 0\}\right) dv \\ \equiv& Q_{n,1}(u,\tau) + Q_{n,0}(u,\tau). \end{aligned} \tag{A.1}$$

We first consider $Q_{n,1}(u,\tau)$. Let

$$H_n(X_i,\tau) = \mathbb{E}\left(\int_0^{\frac{u_0+u_1}{\sqrt{n}}} \left(1\{Y_i(1) - q_1(\tau) \leq v\} - 1\{Y_i(1) - q_1(\tau) \leq 0\}\right) dv | X_i \right). \tag{A.2}$$

Then,

$$Q_{n,1}(u,\tau) = \sum_{i=1}^{2n} \frac{H_n(X_i,\tau)}{2} + \sum_{i=1}^{2n} \left(A_i - \frac{1}{2}\right) H_n(X_i,\tau) + \sum_{i=1}^{2n} A_i \left[\int_0^{\frac{u_0+u_1}{\sqrt{n}}} \left(1\{Y_i(1) - q_1(\tau) \le v\} - 1\{Y_i(1) - q_1(\tau) \le 0\}\right) dv - H_n(X_i,\tau) \right]. \tag{A.3}$$

For the first term on the RHS of (A.3), we have, uniformly over $\tau \in \Upsilon$,

$$\sum_{i=1}^{2n} \frac{H_n(X_i,\tau)}{2} = \frac{1}{4n} \sum_{i=1}^{2n} f_1(q_1(\tau) + \tilde{v}|X_i)(u_0+u_1)^2 \xrightarrow{p} \frac{f_1(q_1(\tau))(u_0+u_1)^2}{2}, \tag{A.4}$$

where $\tilde{v}$ is between 0 and $|u_0+u_1|/\sqrt{n}$ and we use the fact that, due to Assumption 2,

$$\sup_{\tau \in \Upsilon} \frac{1}{2n} \sum_{i=1}^{2n} |f_1(q_1(\tau) + \tilde{v}|X_i) - f_1(q_1(\tau)|X_i)| \le \left(\frac{1}{2n} \sum_{i=1}^{2n} C(X_i) \right) \frac{|u_0+u_1|}{\sqrt{n}} \xrightarrow{p} 0.$$

Lemma E.2 shows

$$\sup_{\tau \in \Upsilon} \left| \sum_{i=1}^{2n} \left(A_i - \frac{1}{2}\right) H_n(X_i,\tau) \right| = o_p(1) \tag{A.5}$$

and

$$\sup_{\tau \in \Upsilon} \left| \sum_{i=1}^{2n} A_i \left[\int_0^{\frac{u_0+u_1}{\sqrt{n}}} \left(1\{Y_i(1) - q_1(\tau) \le v\} - 1\{Y_i(1) - q_1(\tau) \le 0\}\right) dv - H_n(X_i,\tau) \right] \right| = o_p(1). \tag{A.6}$$

Combining (A.3)–(A.6), we have

$$\sup_{\tau \in \Upsilon} \left| Q_{n,1}(u,\tau) - \frac{f_1(q_1(\tau))(u_0+u_1)^2}{2} \right| = o_p(1). \tag{A.7}$$

By a similar argument, we can show that

$$\sup_{\tau \in \Upsilon} \left| Q_{n,0}(u,\tau) - \frac{f_0(q_0(\tau))u_0^2}{2} \right| = o_p(1). \tag{A.8}$$

Combining (A.7) and (A.8), we have

$$Q_n(u,\tau) \xrightarrow{p} \frac{u'Q(\tau)u}{2},$$

where

$$Q(\tau) = \begin{pmatrix} f_1(q_1(\tau)) + f_0(q_0(\tau)) & f_1(q_1(\tau)) \\ f_1(q_1(\tau)) & f_1(q_1(\tau)) \end{pmatrix}. \tag{A.9}$$

Then,

$$\sup_{\tau\in\Upsilon} |L_n(u,\tau) - g_n(u,\tau)| = \sup_{\tau\in\Upsilon} \left| Q_n(u,\tau) - \frac{u'Q(\tau)u}{2} \right| = o_p(1).$$

Last, because $f_a(q_a(\tau))$ for $a = 0, 1$ is bounded and bounded away from zero uniformly over $\tau \in \Upsilon$, so are the eigenvalues of $Q(\tau)$ uniformly over $\tau \in \Upsilon$.

Step (2). Let $e_1 = (1,1)^T$, $e_0 = (1,0)^T$. Then,

$$\begin{aligned} W_n(\tau) =& \sum_{i=1}^{2n} \frac{e_1}{\sqrt{n}} A_i \left(\tau - 1\{Y_i(1) \le q_1(\tau)\}\right) + \sum_{i=1}^{2n} \frac{e_0}{\sqrt{n}} (1 - A_i) \left(\tau - 1\{Y_i(0) \le q_0(\tau)\}\right) \\ \equiv& e_1 W_{n,1}(\tau) + e_0 W_{n,0}(\tau). \end{aligned} \tag{A.10}$$

Recall $m_{1,\tau}(X_i) = \mathbb{E}(\tau - 1\{Y_i(1) \le q_1(\tau)\}|X_i)$. Denote

$$\eta_{i,1}(\tau) = \tau - 1\{Y_i(1) \le q_1(\tau)\} - m_{1,\tau}(X_i).$$

For $W_{n,1}(\tau)$, we have

$$W_{n,1}(\tau) = \sum_{i=1}^{2n} \frac{A_i}{\sqrt{n}} \eta_{i,1}(\tau) + \sum_{i=1}^{2n} \frac{1}{2\sqrt{n}} m_{1,\tau}(X_i) + R_1(\tau) \tag{A.11}$$

where

$$R_1(\tau) = \sum_{i=1}^{2n} \frac{(A_i - 1/2)}{\sqrt{n}} m_{1,\tau}(X_i).$$

By Lemma E.3, we have

$$\sup_{\tau\in\Upsilon} |R_1(\tau)| = o_p(1).$$

Next, we focus on the first two terms on the RHS of (A.11). Note $\{Y_i(1)\}_{i=1}^{2n}$ given $\{X_i\}_{i=1}^{2n}$ is an independent sequence that is also independent of $\{A_i\}_{i=1}^{2n}$. Let $\tilde{Y}_j(1)|\tilde{X}_j$ be distributed according to $Y_{i_j}(1)|X_{i_j}$ where i_j is the j-th smallest index in the set $\{i \in [2n] : A_i = 1\}$ and $\tilde{X}_j = X_{i_j}$. Then, by noticing that $\sum_{i=1}^{2n} A_i = n$, we have

$$\sum_{i=1}^{2n} \frac{A_i}{\sqrt{n}} \eta_{i,1}(\tau) | \{A_i, X_i\}_{i=1}^{2n} \stackrel{d}{=} \sum_{j=1}^{n} \frac{\tilde{\eta}_{j,1}(\tau)}{\sqrt{n}} \Bigg| \{\tilde{X}_j\}_{j=1}^{n}, \tag{A.12}$$

where $\tilde{\eta}_{j,1}(\tau) = \tau - 1\{\tilde{Y}_j(1) \leq q_1(\tau)\} - m_{1,\tau}(\tilde{X}_j)$, and given $\{\tilde{X}_j\}_{j=1}^n$, $\{\tilde{\eta}_{j,1}(\tau)\}_{j=1}^n$ is a sequence of independent random variables. Further denote the conditional distribution of $\tilde{Y}_j(1)$ given $\tilde{X}_j$ as $\mathbb{P}^{(j)}$ and $\Lambda_\tau(x) = F_1(q_1(\tau)|x)(1 - F_1(q_1(\tau)|x))$. Then,

$$\begin{aligned}
\frac{1}{n} \sum_{j=1}^{n} \mathbb{P}^{(j)} (\tilde{\eta}_{j,1}(\tau))^2 &= \frac{1}{n} \sum_{j=1}^{n} \Lambda_\tau(\tilde{X}_j) \\
&= \frac{1}{n} \sum_{i=1}^{2n} A_i \Lambda_\tau(X_i) \\
&= \frac{1}{2n} \sum_{i=1}^{2n} \Lambda_\tau(X_i) + \frac{1}{2n} \sum_{j=1}^{n} (A_{\pi(2j-1)} - A_{\pi(2j)}) \left[\Lambda_\tau(X_{\pi(2j-1)}) - \Lambda_\tau(X_{\pi(2j)})\right] \\
&\xrightarrow{p} \mathbb{E}\Lambda_\tau(X_i),
\end{aligned}$$

where the last convergence holds because

$$\frac{1}{2n} \sum_{i=1}^{2n} \Lambda_\tau(X_i) \xrightarrow{p} \mathbb{E}\Lambda_\tau(X_i),$$

and

$$\left| \frac{1}{2n} \sum_{j=1}^{n} (A_{\pi(2j-1)} - A_{\pi(2j)}) \left[\Lambda_\tau(X_{\pi(2j-1)}) - \Lambda_\tau(X_{\pi(2j)})\right] \right| \lesssim \frac{1}{2n} \sum_{j=1}^{n} ||X_{\pi(2j-1)} - X_{\pi(2j)}||_2 \xrightarrow{p} 0.$$

In addition, because $\tilde{\eta}_{j,1}(\tau)$ is bounded, the Lyapounov condition holds, i.e.,

$$\frac{1}{n^{3/2}} \sum_{i=1}^{n} \mathbb{P}^{(j)} |\tilde{\eta}_{j,1}(\tau)|^3 \xrightarrow{p} 0.$$

Therefore, by the triangular array CLT, for fixed τ, we have

$$\sum_{i=1}^{2n} \frac{A_i}{\sqrt{n}} \eta_{i,1}(\tau) | \{A_i, X_i\}_{i=1}^{2n} \stackrel{d}{=} \sum_{j=1}^{n} \frac{\tilde{\eta}_{j,1}(\tau)}{\sqrt{n}} \Bigg| \{\tilde{X}_j\}_{j=1}^{n} \rightsquigarrow \mathcal{N}(0, \mathbb{E}\Lambda_\tau(X_i)).$$

It is straightforward to extend the results to finite-dimensional convergence by the Cramér-Wold device. In particular, the covariance between $\sum_{i=1}^{2n} \frac{A_i}{\sqrt{n}} \eta_{i,1}(\tau)$ and $\sum_{i=1}^{2n} \frac{A_i}{\sqrt{n}} \eta_{i,1}(\tau')$ conditionally on

$\{X_i\}_{i=1}^{2n}$ converges to

$$\min(\tau,\tau') - \tau\tau' - \mathbb{E}m_{1,\tau}(X)m_{1,\tau'}(X).$$

Next, we show that the process $\{\sum_{i=1}^{2n} \frac{A_i}{\sqrt{n}}\eta_{i,1}(\tau) : \tau \in \Upsilon\}$ is stochastically equicontinuous. Denote $\overline{\mathbb{P}}f = \frac{1}{n}\sum_{j=1}^{n}\mathbb{P}^{(j)}f$ for a generic function f. Let

$$\mathcal{F}_1 = \{[\tau - 1\{Y \le q_1(\tau)\}] - \left[\tau' - 1\{Y \le q_1(\tau')\}\right] : \tau,\tau' \in \Upsilon, |\tau-\tau'| \le \varepsilon\}$$

which is a VC-class with a fixed VC-index and has an envelope $F_i = 2$. In addition,

$$\sigma_n^2 = \sup_{f\in\mathcal{F}_1} \overline{\mathbb{P}}f^2 \lesssim \sup_{\tilde{\tau}\in\Upsilon} \frac{1}{n}\sum_{i=1}^{n}\left[\varepsilon^2 + \frac{f_1(q_1(\tilde{\tau})|\tilde{X}_j)\varepsilon}{f_1(q_1(\tilde{\tau}))}\right] \lesssim \varepsilon \ a.s.$$

Then, by Lemma E.1,

$$\begin{aligned}\mathbb{E}\left[\sup_{\tau,\tau'\in\Upsilon,|\tau-\tau'|\le\varepsilon}\left|\sum_{j=1}^{n}\frac{\tilde{\eta}_{j,1}(\tau)-\tilde{\eta}_{j,1}(\tau')}{\sqrt{n}}\right|\middle|\{\tilde{X}_j\}_{j=1}^{n}\right] =&\mathbb{E}\left[\|\mathbb{P}_n - \overline{\mathbb{P}}\|_{\mathcal{F}_1}\middle|\{\tilde{X}_j\}_{j=1}^{n}\right] \\ \lesssim& \sqrt{\varepsilon\log(1/\varepsilon)} + \frac{\log(1/\varepsilon)}{\sqrt{n}} \ a.s.\end{aligned}$$

For any $\delta,\eta > 0$, we can find an $\varepsilon > 0$ such that

$$\begin{aligned}&\limsup_n \mathbb{P}\left(\sup_{\tau,\tau'\in\Upsilon,|\tau-\tau'|\le\varepsilon}\left|\sum_{i=1}^{2n}\frac{A_i}{\sqrt{n}}\left(\eta_{i,1}(\tau)-\eta_{i,1}(\tau')\right)\right| \ge \delta\right) \\ =&\limsup_n \mathbb{E}\mathbb{P}\left(\sup_{\tau,\tau'\in\Upsilon,|\tau-\tau'|\le\varepsilon}\left|\sum_{i=1}^{2n}\frac{A_i}{\sqrt{n}}\left(\eta_{i,1}(\tau)-\eta_{i,1}(\tau')\right)\right| \ge \delta \middle| \{A_i,X_i\}_{i=1}^{2n}\right) \\ \le&\limsup_n \mathbb{E}\frac{\mathbb{E}\left[\sup_{\tau,\tau'\in\Upsilon,|\tau-\tau'|\le\varepsilon}\left|\sum_{j=1}^{n}\frac{\tilde{\eta}_{j,1}(\tau)-\tilde{\eta}_{j,1}(\tau')}{\sqrt{n}}\right|\middle|\{\tilde{X}_j\}_{j=1}^{n}\right]}{\delta} \\ \lesssim&\limsup_n \frac{\sqrt{\varepsilon\log(1/\varepsilon)} + \frac{\log(1/\varepsilon)}{\sqrt{n}}}{\delta} \le \eta,\end{aligned}$$

where the last inequality holds because $\varepsilon\log(1/\varepsilon) \to 0$ as $\varepsilon \to 0$. This implies $\{\sum_{i=1}^{2n}\frac{A_i}{\sqrt{n}}\eta_{i,1}(\tau) : \tau \in \Upsilon\}$ is stochastically equicontinuous, and hence tight.

In addition, note $\{X_i\}_{i=1}^{2n}$ are i.i.d. and $\{m_{1,\tau}(x) : \tau \in \Upsilon\}$ is Donsker, then $\{\sum_{i=1}^{2n}\frac{1}{2\sqrt{n}}m_{1,\tau}(X_i) : \tau \in \Upsilon\}$ is tight. This leads to the desired result that $\{W_{n,1}(\tau) : \tau \in \Upsilon\}$ is tight. In the same manner, we can show that $\{W_{n,0}(\tau) : \tau \in \Upsilon\}$ is tight, which leads to tightness of $\{W_n(\tau) : \tau \in \Upsilon\}$.

Step (3). Recall $m_{0,\tau}(X_i) = \mathbb{E}(\tau - 1\{Y_i(0) \le q_0(\tau)\}|X_i)$ and let $\eta_{i,0}(\tau) = \tau - 1\{Y_i(0) \le q_0(\tau)\} -$

$m_{0,\tau}(X_i)$. Then, based on the previous two steps, we have

$$\sqrt{n}(\hat{\beta}(\tau)-\beta(\tau)) = Q^{-1}\begin{pmatrix}1 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}\sum_{i=1}^{2n}\frac{A_i}{\sqrt{n}}\eta_{i,1}(\tau)+\sum_{i=1}^{2n}\frac{1}{2\sqrt{n}}m_{1,\tau}(X_i)\\ \sum_{i=1}^{2n}\frac{1-A_i}{\sqrt{n}}\eta_{i,0}(\tau)+\sum_{i=1}^{2n}\frac{1}{2\sqrt{n}}m_{0,\tau}(X_i)\end{pmatrix}+R(\tau) \tag{A.13}$$

where $\sup_{\tau\in\Upsilon}|R(\tau)| = o_p(1)$. In addition, we have already established the stochastic equicontinuity and finite-dimensional convergence of

$$\begin{pmatrix}\sum_{i=1}^{2n}\frac{A_i}{\sqrt{n}}\eta_{i,1}(\tau)+\sum_{i=1}^{2n}\frac{1}{2\sqrt{n}}m_{1,\tau}(X_i)\\ \sum_{i=1}^{2n}\frac{1-A_i}{\sqrt{n}}\eta_{i,0}(\tau)+\sum_{i=1}^{2n}\frac{1}{2\sqrt{n}}m_{0,\tau}(X_i)\end{pmatrix}.$$

Thus, in order to derive the weak limit of $\sqrt{n}(\hat{\beta}(\tau)-\beta(\tau))$ uniformly over $\tau\in\Upsilon$, it suffices to consider its covariance kernel. First, note that, by construction, $\sum_{i=1}^{2n}\frac{A_i}{\sqrt{n}}\eta_{i,1}(\tau)\perp\!\!\!\perp\sum_{i=1}^{2n}\frac{1-A_i}{\sqrt{n}}\eta_{i,0}(\tau')$ for any $(\tau,\tau')\in\Upsilon$. Second, note that $\sum_{i=1}^{2n}\frac{A_i}{\sqrt{n}}\eta_{i,1}(\tau)$ is asymptotically independent of $\sum_{i=1}^{2n}\frac{1}{2\sqrt{n}}m_{1,\tau'}(X_i)$. To see this, let $(s,t)\in\Re^2$, then

$$\begin{aligned}
&\mathbb{P}\left(\sum_{i=1}^{2n}\frac{A_i}{\sqrt{n}}\eta_{i,1}(\tau)\le t,\sum_{i=1}^{2n}\frac{1}{2\sqrt{n}}m_{1,\tau'}(X_i)\le s\right)\\
=&\mathbb{E}\left\{\mathbb{P}\left(\sum_{i=1}^{2n}\frac{A_i}{\sqrt{n}}\eta_{i,1}(\tau)\le t\middle|\{A_i,X_i\}_{i=1}^{2n}\right)1\left\{\sum_{i=1}^{2n}\frac{1}{2\sqrt{n}}m_{1,\tau'}(X_i)\le s\right\}\right\}\\
=&\Phi(t/\sqrt{\mathbb{E}\Lambda_\tau(X_i)})\mathbb{P}\left(\sum_{i=1}^{2n}\frac{1}{2\sqrt{n}}m_{1,\tau'}(X_i)\le s\right)\\
&+\mathbb{E}\Bigg\{\left[\mathbb{P}\left(\sum_{i=1}^{2n}\frac{A_i}{\sqrt{n}}\eta_{i,1}(\tau)\le t\middle|\{A_i,X_i\}_{i=1}^{2n}\right)-\Phi(t/\sqrt{\mathbb{E}\Lambda_\tau(X_i)})\right]1\left\{\sum_{i=1}^{2n}\frac{1}{2\sqrt{n}}m_{1,\tau'}(X_i)\le s\right\}\Bigg\}\\
\to&\Phi(t/\sqrt{\mathbb{E}\Lambda_\tau(X_i)})\Phi(s/\sqrt{\mathbb{E}m_{1,\tau}^2(X_i)/2}),
\end{aligned}$$

where the last convergence holds due to the fact that

$$\mathbb{P}\left(\sum_{i=1}^{2n}\frac{A_i}{\sqrt{n}}\eta_{i,1}(\tau)\le t\middle|\{A_i,X_i\}_{i=1}^{2n}\right)-\Phi(t/\sqrt{\mathbb{E}\Lambda_\tau(X_i)})\xrightarrow{p}0.$$

We can extend the independence result to multiple τ and τ', implying that the two stochastic processes

$$\left\{\sum_{i=1}^{2n}\frac{A_i}{\sqrt{n}}\eta_{i,1}(\tau):\tau\in\Upsilon\right\}\quad\text{and}\quad\left\{\sum_{i=1}^{2n}\frac{1}{2\sqrt{n}}m_{1,\tau}(X_i):\tau\in\Upsilon\right\}$$

are asymptotically independent. For the same reason, we can show

$$\left\{\left(\sum_{i=1}^{2n}\frac{A_i}{\sqrt{n}}\eta_{i,1}(\tau),\sum_{i=1}^{2n}\frac{1-A_i}{\sqrt{n}}\eta_{i,0}(\tau)\right):\tau\in\Upsilon\right\}\quad\text{and}\quad\left\{\left(\sum_{i=1}^{2n}\frac{1}{2\sqrt{n}}m_{1,\tau}(X_i),\sum_{i=1}^{2n}\frac{1}{2\sqrt{n}}m_{0,\tau}(X_i)\right):\tau\in\Upsilon\right\}$$

are asymptotically independent. Last, it is tedious but straightforward to show that, uniformly over $\tau\in\Upsilon$,

$$\left(\sum_{i=1}^{2n}\frac{A_i}{\sqrt{n}}\eta_{i,1}(\tau),\sum_{i=1}^{2n}\frac{1-A_i}{\sqrt{n}}\eta_{i,0}(\tau)\right)\rightsquigarrow\tilde{\mathcal{B}}_1(\tau),$$

and

$$\left(\sum_{i=1}^{2n}\frac{1}{2\sqrt{n}}m_{1,\tau}(X_i),\sum_{i=1}^{2n}\frac{1}{2\sqrt{n}}m_{0,\tau}(X_i)\right)\rightsquigarrow\tilde{\mathcal{B}}_2(\tau),$$

where $\tilde{\mathcal{B}}_1(\tau)$ and $\tilde{\mathcal{B}}_2(\tau)$ are two Gaussian processes with covariance kernels

$$\tilde{\Sigma}_1(\tau,\tau')=\begin{pmatrix}\mathbb{E}\left[\min(\tau,\tau')-\tau\tau'-\mathbb{E}m_{1,\tau}(X)m_{1,\tau'}(X)\right] & 0\\ 0 & \mathbb{E}\left[\min(\tau,\tau')-\tau\tau'-\mathbb{E}m_{0,\tau}(X)m_{0,\tau'}(X)\right]\end{pmatrix}\tag{A.14}$$

and

$$\tilde{\Sigma}_2(\tau,\tau')=\frac{1}{2}\begin{pmatrix}\mathbb{E}m_{1,\tau}(X)m_{1,\tau'}(X) & \mathbb{E}m_{1,\tau}(X)m_{0,\tau'}(X)\\ \mathbb{E}m_{1,\tau'}(X)m_{0,\tau}(X) & \mathbb{E}m_{0,\tau}(X)m_{0,\tau'}(X)\end{pmatrix},\quad\text{respectively.}\tag{A.15}$$

This implies $\sqrt{n}(\hat{\beta}(\tau)-\beta(\tau))\rightsquigarrow\tilde{\mathcal{B}}(\tau)$, where $\tilde{\mathcal{B}}(\tau)$ is a Gaussian process with covariance kernel

$$\tilde{\Sigma}(\tau,\tau')=Q^{-1}(\tau)\begin{pmatrix}1&1\\1&0\end{pmatrix}\left(\tilde{\Sigma}_1(\tau,\tau')+\tilde{\Sigma}_2(\tau,\tau')\right)\left[\begin{pmatrix}1&1\\1&0\end{pmatrix}Q^{-1}(\tau')\right]^T.$$

Focusing on the second element of $\hat{\beta}(\tau)$, we have

$$\sqrt{n}(\hat{q}(\tau)-q(\tau))\rightsquigarrow\mathcal{B}(\tau),$$

where $\mathcal{B}(\tau)$ is a Gaussian process with covariance kernel

$$\begin{aligned}\Sigma(\tau,\tau')=&\frac{\min(\tau,\tau')-\tau\tau'-\mathbb{E}m_{1,\tau}(X)m_{1,\tau'}(X)}{f_1(q_1(\tau))f_1(q_1(\tau'))}+\frac{\min(\tau,\tau')-\tau\tau'-\mathbb{E}m_{0,\tau}(X)m_{0,\tau'}(X)}{f_0(q_0(\tau))f_0(q_0(\tau'))}\\&+\frac{1}{2}\mathbb{E}\left(\frac{m_{1,\tau}(X)}{f_1(q_1(\tau))}-\frac{m_{0,\tau}(X)}{f_0(q_0(\tau))}\right)\left(\frac{m_{1,\tau'}(X)}{f_1(q_1(\tau'))}-\frac{m_{0,\tau'}(X)}{f_0(q_0(\tau'))}\right).\end{aligned}$$

B Proof of Theorem 4.1

Let $u = (u_0, u_1)' \in \Re^2$ and

$$L_n^w(u,\tau) = \sum_{i=1}^{2n} \xi_i \left[\rho_\tau(Y_i - \dot{A}_i'\beta(\tau) - \dot{A}_i'u/\sqrt{n}) - \rho_\tau(Y_i - \dot{A}_i'\beta(\tau)) \right].$$

Then, by change of variables we have

$$\sqrt{n}(\hat{\beta}^w(\tau) - \beta(\tau)) = \arg\min_u L_n^w(u,\tau).$$

Notice that $L_n^w(u,\tau)$ is convex in u for each τ and bounded in τ for each u. In the following, we divide the proof into three steps. In Step (1), we show that there exists

$$g_n^w(u,\tau) = -u'W_n^w(\tau) + \frac{u'Q(\tau)u}{2}$$

such that for each u,

$$\sup_{\tau \in \Upsilon} |L_n^w(u,\tau) - g_n^w(u,\tau)| \xrightarrow{p} 0$$

and $Q(\tau)$ is defined in the proof of Theorem 3.1. In Step (2), we show $W_n^w(\tau)$ as a stochastic process over $\tau \in \Upsilon$ is tight. Then by Kato (2009, Theorem 2), we have

$$\sqrt{n}(\hat{\beta}^w(\tau) - \beta(\tau)) = [Q(\tau)]^{-1} W_n^w(\tau) + r_n(\tau),$$

where $\sup_{\tau \in \Upsilon} ||r_n(\tau)||_2 = o_p(1)$. Last, in Step (3), we establish the weak convergence of

$$\sqrt{n}(\hat{\beta}^w(\tau) - \hat{\beta}(\tau))$$

conditionally on data.

Step (1). Similar to Step (1) in the previous section, we have

$$L_n^w(u,\tau) = -u'W_n^w(\tau) + Q_n^w(u,\tau),$$

where

$$W_n^w(\tau) = \sum_{i=1}^{2n} \frac{\xi_i}{\sqrt{n}} \dot{A}_i \left(\tau - 1\{Y_i \leq \dot{A}_i'\beta(\tau)\} \right)$$

and

$$
\begin{aligned}
Q_n^w(u,\tau) =& \sum_{i=1}^{2n} \xi_i \int_0^{\frac{\dot{A}_i' u}{\sqrt{n}}} \left(1\{Y_i - \dot{A}_i'\beta(\tau) \le v\} - 1\{Y_i - \dot{A}_i'\beta(\tau) \le 0\}\right) dv \\
=& \sum_{i=1}^{2n} \xi_i A_i \int_0^{\frac{u_0+u_1}{\sqrt{n}}} \left(1\{Y_i(1) - q_1(\tau) \le v\} - 1\{Y_i(1) - q_1(\tau) \le 0\}\right) dv \\
&+ \sum_{i=1}^{2n} \xi_i (1-A_i) \int_0^{\frac{u_0}{\sqrt{n}}} \left(1\{Y_i(0) - q_0(\tau) \le v\} - 1\{Y_i(0) - q_0(\tau) \le 0\}\right) dv \\
\equiv& Q_{n,1}^w(u,\tau) + Q_{n,0}^w(u,\tau). \qquad \text{(B.1)}
\end{aligned}
$$

We first consider $Q_{n,1}^w(u,\tau)$. Note

$$
\begin{aligned}
H_n(X_i,\tau) =& \mathbb{E}\xi_i \left(\int_0^{\frac{u_0+u_1}{\sqrt{n}}} \left(1\{Y_i(1) - q_1(\tau) \le v\} - 1\{Y_i(1) - q_1(\tau) \le 0\}\right) dv | X_i \right) \\
=& \mathbb{E} \left(\int_0^{\frac{u_0+u_1}{\sqrt{n}}} \left(1\{Y_i(1) - q_1(\tau) \le v\} - 1\{Y_i(1) - q_1(\tau) \le 0\}\right) dv | X_i \right). \qquad \text{(B.2)}
\end{aligned}
$$

Then,

$$
\begin{aligned}
Q_{n,1}^w(u,\tau) =& \sum_{i=1}^{2n} \frac{H_n(X_i,\tau)}{2} + \sum_{i=1}^{2n} \left(A_i - \frac{1}{2}\right) H_n(X_i,\tau) \\
&+ \sum_{i=1}^{2n} A_i \left[\xi_i \int_0^{\frac{u_0+u_1}{\sqrt{n}}} \left(1\{Y_i(1) - q_1(\tau) \le v\} - 1\{Y_i(1) - q_1(\tau) \le 0\}\right) dv - H_n(X_i,\tau) \right]. \qquad \text{(B.3)}
\end{aligned}
$$

By (A.4), we have, uniformly over $\tau \in \Upsilon$,

$$
\sum_{i=1}^{2n} \frac{H_n(X_i,\tau)}{2} \xrightarrow{p} \frac{f_1(q_1(\tau))(u_0+u_1)^2}{2}.
$$

In addition, (A.5) implies

$$
\sup_{\tau \in \Upsilon} \left| \sum_{i=1}^{2n} \left(A_i - \frac{1}{2}\right) H_n(X_i,\tau) \right| = o_p(1).
$$

Last, Lemma E.2 implies

$$
\sup_{\tau \in \Upsilon} \left| \sum_{i=1}^{2n} A_i \left[\xi_i \int_0^{\frac{u_0+u_1}{\sqrt{n}}} \left(1\{Y_i(1) - q_1(\tau) \le v\} - 1\{Y_i(1) - q_1(\tau) \le 0\}\right) dv - H_n(X_i,\tau) \right] \right| = o_p(1).
$$

Combining the above results, we have

$$\sup_{\tau \in \Upsilon} \left| Q^w_{n,1}(u,\tau) - \frac{f_1(q_1(\tau))(u_0+u_1)^2}{2} \right| = o_p(1). \tag{B.4}$$

By a similar argument, we can show that

$$\sup_{\tau \in \Upsilon} \left| Q^w_{n,0}(u,\tau) - \frac{f_0(q_0(\tau))u_0^2}{2} \right| = o_p(1). \tag{B.5}$$

Combining (B.4) and (B.5), we have

$$Q^w_n(u,\tau) \xrightarrow{p} \frac{u'Q(\tau)u}{2},$$

where $Q(\tau)$ is defined in (A.9). Then,

$$\sup_{\tau \in \Upsilon} |L^w_n(u,\tau) - g^w_n(u,\tau)| = \sup_{\tau \in \Upsilon} \left| Q^w_n(u,\tau) - \frac{u'Q(\tau)u}{2} \right| = o_p(1).$$

Step (2). We have

$$\begin{aligned} W^w_n(\tau) =& \sum_{i=1}^{2n} \frac{e_1}{\sqrt{n}} \xi_i A_i \left(\tau - 1\{Y_i(1) \le q_1(\tau)\}\right) + \sum_{i=1}^{2n} \frac{e_0}{\sqrt{n}} (1-A_i)\xi_i \left(\tau - 1\{Y_i(0) \le q_0(\tau)\}\right) \\ \equiv& e_1 W^w_{n,1}(\tau) + e_0 W^w_{n,0}(\tau). \end{aligned} \tag{B.6}$$

Recall $m_{1,\tau}(X_i) = \mathbb{E}(\tau - 1\{Y_i(1) \le q_1(\tau)\}|X_i)$, $e_1 = (1,1)^T$, and $e_0 = (1,0)^T$, and denote

$$\eta^w_{i,1}(\tau) = \xi_i(\tau - 1\{Y_i(1) \le q_1(\tau)\}) - m_{1,\tau}(X_i).$$

Then, for $W^w_{n,1}(\tau)$, we have

$$W^w_{n,1}(\tau) = \sum_{i=1}^{2n} \frac{A_i}{\sqrt{n}} \eta^w_{i,1}(\tau) + \sum_{i=1}^{2n} \frac{1}{2\sqrt{n}} m_{1,\tau}(X_i) + R_1(\tau), \tag{B.7}$$

where by Lemma E.3,

$$\sup_{\tau \in \Upsilon} |R_1(\tau)| = \sup_{\tau \in \Upsilon} \left| \sum_{i=1}^{2n} \frac{(A_i - 1/2)}{\sqrt{n}} m_{1,\tau}(X_i) \right| = o_p(1).$$

The second term on the RHS of (B.7) is stochastically equicontinuous and tight. Next, we focus on the first term. Similar to the argument in Step (2) in the previous section, we have

$$\sum_{i=1}^{2n} \frac{A_i}{\sqrt{n}} \eta_{i,1}^w(\tau) | \{A_i, X_i\}_{i=1}^{2n} \stackrel{d}{=} \sum_{j=1}^{n} \frac{\tilde{\eta}_{j,1}^w(\tau)}{\sqrt{n}} \bigg| \{\tilde{X}_j\}_{j=1}^n, \tag{B.8}$$

where $\tilde{\eta}_{j,1}^w(\tau) = \tilde{\xi}_j(\tau - 1\{\tilde{Y}_j(1) \leq q_1(\tau)\}) - m_{1,\tau}(\tilde{X}_j)$, $(\tilde{Y}_j(1), \tilde{X}_j)$ are as defined before, $\tilde{\xi}_j = \xi_{i_j}$, i_j is the j-th smallest index in the set $\{i \in [2n] : A_i = 1\}$, and given $\{\tilde{X}_j\}_{j=1}^n$, $\{\tilde{\eta}_{j,1}^w(\tau)\}_{j=1}^n$ is a sequence of independent random variables. Further, denote the conditional distribution of $(\tilde{\xi}_j, \tilde{Y}_j(1))$ given $\tilde{X}_j$ as $\mathbb{P}^{(j)}$. Then,

$$\frac{1}{n}\sum_{j=1}^{n} \mathbb{P}^{(j)}(\tilde{\eta}_{j,1}^w(\tau))^2 = \frac{1}{n}\sum_{j=1}^{n} \left\{ \mathbb{E}\left[(\tilde{\xi}_j^w)^2 (\tau - 1\{\tilde{Y}_j(1) \leq q_1(\tau)\})^2 | \tilde{X}_j \right] - m_{1,\tau}^2(\tilde{X}_j) \right\} \leq \overline{C} < \infty,$$

for some constant $\overline{C} > 0$. This implies that pointwise in $\tau \in \Upsilon$,

$$\sum_{i=1}^{2n} \frac{A_i}{\sqrt{n}} \eta_{i,1}^w(\tau) | \{A_i, X_i\}_{i=1}^{2n} \stackrel{d}{=} \sum_{j=1}^{n} \frac{\tilde{\eta}_{j,1}^w(\tau)}{\sqrt{n}} \bigg| \{\tilde{X}_j\}_{j=1}^n = O_p(1).$$

In addition, let

$$\mathcal{F}_2 = \{\xi\left[\tau - 1\{Y \leq q_1(\tau)\}\right] - \xi\left[\tau' - 1\{Y \leq q_1(\tau')\}\right] : \tau, \tau' \in \Upsilon, |\tau - \tau'| \leq \varepsilon\}$$

which is a VC-class with a fixed VC-index and has an envelope $F_i = 2\xi_i$. In addition, $||\max_{i \in [n]} F_i||_{\mathbb{P},2} \leq C\log(n)$ and

$$\sigma_n^2 = \sup_{f \in \mathcal{F}_2} \overline{\mathbb{P}} f^2 \lesssim \sup_{\tilde{\tau} \in \Upsilon} \frac{1}{n}\sum_{i=1}^{n} \left[\varepsilon^2 + \frac{f_1(q_1(\tilde{\tau})|\tilde{X}_j)\varepsilon}{f_1(q_1(\tilde{\tau}))} \right] \lesssim \varepsilon \ a.s.$$

Then, by Lemma E.1,

$$\begin{aligned}\mathbb{E}\left[\sup_{\tau,\tau' \in \Upsilon, |\tau-\tau'| \leq \varepsilon} \left| \sum_{j=1}^{n} \frac{\tilde{\eta}_{j,1}^w(\tau) - \tilde{\eta}_{j,1}^w(\tau')}{\sqrt{n}} \right| \bigg| \{\tilde{X}_j\}_{j=1}^n \right] =& \mathbb{E}\left[||\mathbb{P}_n - \overline{\mathbb{P}}||_{\mathcal{F}_2} \bigg| \{\tilde{X}_j\}_{j=1}^n \right] \\ \lesssim& \sqrt{\varepsilon \log(1/\varepsilon)} + \frac{\log(1/\varepsilon)\log(n)}{\sqrt{n}} \ a.s.\end{aligned}$$

The RHS of the above display vanishes as $n \to \infty$ followed by $\varepsilon \to 0$, which implies

$$\sum_{i=1}^{2n} \frac{A_i}{\sqrt{n}} \eta_{i,1}^w(\tau) | \{A_i, X_i\}_{i=1}^{2n} \stackrel{d}{=} \sum_{j=1}^{n} \frac{\tilde{\eta}_{j,1}^w(\tau)}{\sqrt{n}} \bigg| \{\tilde{X}_j\}_{j=1}^n \tag{B.9}$$

is stochastically equicontinuous. Therefore, $\sum_{i=1}^{2n} \frac{A_i}{\sqrt{n}} \eta^w_{i,1}(\tau) | \{A_i, X_i\}_{i=1}^{2n}$, and hence $W^w_{n,1}(\tau)$ is tight. Similarly, we can show $W^w_{n,0}(\tau)$ is tight.

Step (3). Based on the previous two steps, we have

$$\sqrt{n}(\hat{\beta}^w(\tau) - \beta(\tau)) = Q^{-1} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \sum_{i=1}^{2n} \frac{A_i}{\sqrt{n}} \eta^w_{i,1}(\tau) + \sum_{i=1}^{2n} \frac{1}{2\sqrt{n}} m_{1,\tau}(X_i) \\ \sum_{i=1}^{2n} \frac{1-A_i}{\sqrt{n}} \eta^w_{i,0}(\tau) + \sum_{i=1}^{2n} \frac{1}{2\sqrt{n}} m_{0,\tau}(X_i) \end{pmatrix} + R^w(\tau) \quad \text{(B.10)}$$

where $\sup_{\tau \in \Upsilon} ||R^w(\tau)||_2 = o_p(1)$ and $\sqrt{n}(\hat{\beta}^w(\tau) - \beta(\tau))$ is stochastically equicontinuous. Taking the difference between (A.13) and (B.10), we have

$$\sqrt{n}(\hat{\beta}^w(\tau) - \hat{\beta}(\tau)) = Q^{-1} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \sum_{i=1}^{2n} \frac{A_i}{\sqrt{n}} (\xi_i - 1)(\tau - 1\{Y_i(1) \le q_1(\tau)\}) \\ \sum_{i=1}^{2n} \frac{1-A_i}{\sqrt{n}} (\xi_i - 1)(\tau - 1\{Y_i(0) \le q_0(\tau)\}) \end{pmatrix} + R^*(\tau), \quad \text{(B.11)}$$

where $\sup_{\tau \in \Upsilon} |R^*(\tau)| = o_p(1)$. In addition, because both $\sqrt{n}(\hat{\beta}^w(\tau) - \beta(\tau))$ and $\sqrt{n}(\hat{\beta}(\tau) - \beta(\tau))$ are stochastically equicontinuous, so be $\sqrt{n}(\hat{\beta}^w(\tau) - \hat{\beta}(\tau))$. Then by Markov inequality, $\sqrt{n}(\hat{\beta}^w(\tau) - \hat{\beta}(\tau))$ is stochastically equicontinuous conditionally on data as well. In order to derive the limiting distribution of $\sqrt{n}(\hat{\beta}^w(\tau) - \hat{\beta}(\tau))$ conditionally on data, we only need to compute the covariance kernel. Note that

$$\begin{aligned} &\mathbb{E}\left[\begin{pmatrix} \sum_{i=1}^{2n} \frac{A_i}{\sqrt{n}} (\xi_i - 1)(\tau - 1\{Y_i(1) \le q_1(\tau)\}) \\ \sum_{i=1}^{2n} \frac{1-A_i}{\sqrt{n}} (\xi_i - 1)(\tau - 1\{Y_i(0) \le q_0(\tau)\}) \end{pmatrix} \begin{pmatrix} \sum_{i=1}^{2n} \frac{A_i}{\sqrt{n}} (\xi_i - 1)(\tau' - 1\{Y_i(1) \le q_1(\tau')\}) \\ \sum_{i=1}^{2n} \frac{1-A_i}{\sqrt{n}} (\xi_i - 1)(\tau' - 1\{Y_i(0) \le q_0(\tau')\}) \end{pmatrix}^T \middle| Data \right] \\ =& \frac{1}{n} \sum_{i=1}^{2n} \begin{pmatrix} A_i(\tau - 1\{Y_i(1) \le q_1(\tau)\})(\tau' - 1\{Y_i(1) \le q_1(\tau')\}) & 0 \\ 0 & (1-A_i)(\tau - 1\{Y_i(0) \le q_0(\tau)\})(\tau' - 1\{Y_i(0) \le q_0(\tau')\}) \end{pmatrix}. \end{aligned}$$

For the $(1,1)$ entry, we have

$$\begin{aligned} &\frac{1}{n} \sum_{i=1}^{2n} A_i(\tau - 1\{Y_i(1) \le q_1(\tau)\})(\tau' - 1\{Y_i(1) \le q_1(\tau')\}) \\ =& \frac{1}{n} \sum_{i=1}^{2n} A_i \eta_{1,i}(\tau) \eta_{1,i}(\tau') + \frac{1}{n} \sum_{i=1}^{2n} A_i \eta_{1,i}(\tau) m_{1,\tau'}(X_i) + \frac{1}{n} \sum_{i=1}^{2n} A_i \eta_{1,i}(\tau') m_{1,\tau}(X_i) + \frac{1}{n} \sum_{i=1}^{2n} A_i m_{1,\tau}(X_i) m_{1,\tau'}(X_i). \end{aligned}$$

Note that

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^{2n} A_i \eta_{1,i}(\tau) \eta_{1,i}(\tau') \stackrel{d}{=} & \frac{1}{n} \sum_{j=1}^{n} \tilde{\eta}_{1,j}(\tau) \tilde{\eta}_{1,j}(\tau') \\ & \xrightarrow{p} \lim_n \frac{1}{n} \sum_{j=1}^{n} (F_1(q_1(\min(\tau, \tau')) | \tilde{X}_j) - F_1(q_1(\tau) | \tilde{X}_j) F_1(q_1(\tau') | \tilde{X}_j)) \\ & = \min(\tau, \tau') - \mathbb{E} F_1(q_1(\tau) | X_i) F_1(q_1(\tau') | X_i). \end{aligned} \quad \text{(B.12)}$$

Lemma E.4 shows

$$\frac{1}{n}\sum_{i=1}^{2n} A_i\eta_{1,i}(\tau)m_{1,\tau'}(X_i) \xrightarrow{p} 0$$

and

$$\frac{1}{n}\sum_{i=1}^{2n} A_i\eta_{1,i}(\tau')m_{1,\tau}(X_i) \xrightarrow{p} 0.$$

Lemma E.6 implies

$$\frac{1}{n}\sum_{i=1}^{2n} A_i m_{1,\tau}(X_i)m_{1,\tau'}(X_i) \xrightarrow{p} \mathbb{E}m_{1,\tau}(X_i)m_{1,\tau'}(X_i).$$

This means

$$\frac{1}{n}\sum_{i=1}^{2n} A_i(\tau - 1\{Y_i(1) \leq q_1(\tau)\})(\tau' - 1\{Y_i(1) \leq q_1(\tau')\}) \xrightarrow{p} \min(\tau,\tau') - \tau\tau'.$$

For the same reason,

$$\frac{1}{n}\sum_{i=1}^{2n}(1 - A_i)(\tau - 1\{Y_i(0) \leq q_0(\tau)\})(\tau' - 1\{Y_i(0) \leq q_0(\tau')\} \xrightarrow{p} \min(\tau,\tau') - \tau\tau'.$$

Then, for the second element $\hat{\beta}_1^w(\tau)$ of $\hat{\beta}^w(\tau)$, conditional on the data, we have

$$\sqrt{n}(\hat{\beta}_1^w(\tau) - \hat{\beta}_1(\tau)) \rightsquigarrow \mathcal{B}^w(\tau),$$

where $\mathcal{B}^w(\tau)$ is a Gaussian process with covariance kernel

$$\Sigma^\dagger(\tau,\tau') = \frac{\min(\tau,\tau') - \tau\tau'}{f_1(q_1(\tau))f_1(q_1(\tau'))} + \frac{\min(\tau,\tau') - \tau\tau'}{f_0(q_0(\tau))f_0(q_0(\tau'))}.$$

C Proof of Theorem 4.2

Let $u \in \Re^2$ and

$$L_n^*(u,\tau) = \sum_{i=1}^{2n}\left[\rho_\tau(Y_i - \dot{A}_i'\beta(\tau) - \dot{A}_i'u/\sqrt{n}) - \rho_\tau(Y_i - \dot{A}_i'\beta(\tau))\right] - u^T\begin{pmatrix}1 & 1\\ 1 & 0\end{pmatrix}S_n^*(\tau).$$

Then,

$$\sqrt{n}\left(\hat{\beta}^*(\tau) - \beta(\tau)\right) = \arg\min_u L_n^*(u, \tau).$$

By the same argument as in the proof of Theorem 3.1, we have

$$L_n^*(u,\tau) = -u^T W_n(\tau) + Q_n(u,\tau) - u^T \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} S_n^*(\tau) = -u^T \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} (S_n(\tau) + S_n^*(\tau)) + Q_n(u,\tau).$$

Further note that $S_n^*(\tau) = \frac{1}{\sqrt{2}}\left(S_{n,1}^*(\tau) + S_{n,2}^*(\tau)\right)$. In the following, we divide the proof into three steps. In Step (1), we derive the weak limit of $S_{n,1}^*(\tau)$ given data. In Step (2), we derive the weak limit of $S_{n,2}^*(\tau)$. In Step (3), we derive the desired result of this theorem.

Step (1). Given the data, $S_{n,1}^*(\tau)$ is a Gaussian process with covariance kernel

$$\tilde{\Sigma}_1^*(\tau,\tau') = \begin{pmatrix} \tilde{\Sigma}_{1,1,1}^*(\tau,\tau') & \tilde{\Sigma}_{1,1,2}^*(\tau,\tau') \\ \tilde{\Sigma}_{1,2,1}^*(\tau,\tau') & \tilde{\Sigma}_{1,2,2}^*(\tau,\tau') \end{pmatrix}$$

where

$$\tilde{\Sigma}_{1,1,1}^*(\tau,\tau') = \frac{1}{n}\sum_{j=1}^n \left(\tau - 1\{Y_{(j,1)} \le \hat{q}_1(\tau)\}\right)\left(\tau' - 1\{Y_{(j,1)} \le \hat{q}_1(\tau')\}\right),$$

$$\tilde{\Sigma}_{1,1,2}^*(\tau,\tau') = \frac{1}{n}\sum_{j=1}^n \left(\tau - 1\{Y_{(j,1)} \le \hat{q}_1(\tau)\}\right)\left(\tau' - 1\{Y_{(j,0)} \le \hat{q}_0(\tau')\}\right),$$

$$\tilde{\Sigma}_{1,2,1}^*(\tau,\tau') = \frac{1}{n}\sum_{j=1}^n \left(\tau - 1\{Y_{(j,0)} \le \hat{q}_0(\tau)\}\right)\left(\tau' - 1\{Y_{(j,1)} \le \hat{q}_1(\tau')\}\right),$$

and

$$\tilde{\Sigma}_{1,2,2}^*(\tau,\tau') = \frac{1}{n}\sum_{j=1}^n \left(\tau - 1\{Y_{(j,0)} \le \hat{q}_0(\tau)\}\right)\left(\tau' - 1\{Y_{(j,0)} \le \hat{q}_0(\tau')\}\right).$$

Next, we derive the limit of $\tilde{\Sigma}_1^*(\tau,\tau')$ uniformly over $\tau,\tau' \in \Upsilon$. Recall $m_{1,\tau}(X_i,q) = \mathbb{E}\left(\tau - 1\{Y_i(1) \le q\}|X_i\right)$ and define $\eta_{1,i}(q,\tau) = (\tau - 1\{Y_i(1) \le q\}) - m_{1,\tau}(X_i,q)$. Then

$$\tilde{\Sigma}_{1,1,1}(\tau,\tau') = \frac{1}{n}\sum_{j=1}^n \eta_{1,(j,1)}(\hat{q}_1(\tau),\tau)\eta_{1,(j,1)}(\hat{q}_1(\tau'),\tau') + \frac{1}{n}\sum_{j=1}^n \eta_{1,(j,1)}(\hat{q}_1(\tau),\tau) m_{1,\tau'}(X_{(j,1)},\hat{q}_1(\tau'))$$

$$+\frac{1}{n}\sum_{j=1}^{n}\eta_{1,(j,1)}(\hat{q}_1(\tau'),\tau')m_{1,\tau}(X_{(j,1)},\hat{q}_1(\tau))+\frac{1}{n}\sum_{j=1}^{n}m_{1,\tau}(X_{(j,1)},\hat{q}_1(\tau))m_{1,\tau'}(X_{(j,1)},\hat{q}_1(\tau'))$$

$$=I(\tau,\tau')+II(\tau,\tau')+III(\tau,\tau')+IV(\tau,\tau'), \tag{C.1}$$

where we use the fact that $Y_{(j,1)}=Y_{(j,1)}(1)$ and $Y_{(j,0)}=Y_{(j,0)}(0)$. Given $\{A_i,X_i\}_{i=1}^{2n}$, $\{Y_{(j,1)}(1)\}_{j=1}^{n}$ is a sequence of independent random variables with probability measure $\Pi_{j=1}^{n}\mathbb{P}^{(j)}$, where $\mathbb{P}^{(j)}$ is the conditional probability of $Y(1)$ given X evaluated at $X=X_{(j,1)}$. Therefore,

$$I(\tau,\tau')=\overline{\mathbb{P}}\eta_{1,(j,1)}(\hat{q}_1(\tau),\tau)\eta_{1,(j,1)}(\hat{q}_1(\tau'),\tau')+\left(\mathbb{P}_n-\overline{\mathbb{P}}\right)\eta_{1,(j,1)}(\hat{q}_1(\tau),\tau)\eta_{1,(j,1)}(\hat{q}_1(\tau'),\tau'), \tag{C.2}$$

where $\overline{\mathbb{P}}\eta_{1,(j,1)}(\hat{q}_1(\tau),\tau)\eta_{1,(j,1)}(\hat{q}_1(\tau'),\tau')$ is interpreted as $\overline{\mathbb{P}}\eta_{1,(j,1)}(q,\tau)\eta_{1,(j,1)}(q',\tau')|_{q=\hat{q}_1(\tau),q'=\hat{q}(\tau')}$. In addition, by Theorem 3.1, for any $\varepsilon>0$, it is possible to find a sufficiently large constant L such that

$$\mathbb{P}(\sup_{\tau\in\Upsilon}|\hat{q}(\tau)-q(\tau)|\le L/\sqrt{n})\ge 1-\varepsilon. \tag{C.3}$$

Therefore, we have,

$$\begin{aligned}
&\overline{\mathbb{P}}\eta_{1,(j,1)}(\hat{q}_1(\tau),\tau)\eta_{1,(j,1)}(\hat{q}_1(\tau'),\tau')\\
=&\frac{1}{n}\sum_{j=1}^{n}\left[F_1(\min(\hat{q}_1(\tau),\hat{q}_1(\tau'))|X_{(j,1)})-F_1(\hat{q}_1(\tau)|X_{(j,1)})F_1(\hat{q}_1(\tau')|X_{(j,1)})\right]\\
=&\frac{1}{n}\sum_{j=1}^{n}\left[F_1(\min(q_1(\tau),q_1(\tau'))|X_{(j,1)})-F_1(q_1(\tau)|X_{(j,1)})F_1(q_1(\tau')|X_{(j,1)})\right]+R_I(\tau,\tau')\\
=&\frac{1}{n}\sum_{i=1}^{2n}A_i\left[F_1(\min(q_1(\tau),q_1(\tau'))|X_i)-F_1(q_1(\tau)|X_i)F_1(q_1(\tau')|X_i)\right]+R_I(\tau,\tau')\\
=&\frac{1}{2n}\sum_{i=1}^{2n}\left[F_1(\min(q_1(\tau),q_1(\tau'))|X_i)-F_1(q_1(\tau)|X_i)F_1(q_1(\tau')|X_i)\right]\\
+&\frac{1}{n}\sum_{i=1}^{2n}\left(A_i-\frac{1}{2}\right)\left[F_1(\min(q_1(\tau),q_1(\tau'))|X_i)-F_1(q_1(\tau)|X_i)F_1(q_1(\tau')|X_i)\right]+R_I(\tau,\tau'),
\end{aligned} \tag{C.4}$$

where $\sup_{\tau,\tau'\in\Upsilon}|R_I(\tau,\tau')|\xrightarrow{p}0$ due to (C.3) and Lipschitz continuity of $F_1(\cdot|X)$.

By the standard uniform convergence theorem (van der Vaart and Wellner (1996, Theorem 2.4.1)), uniformly over $\tau,\tau'\in\Upsilon$,

$$\frac{1}{2n}\sum_{i=1}^{2n}\left[F_1(\min(q_1(\tau),q_1(\tau'))|X_i)-F_1(q_1(\tau)|X_i)F_1(q_1(\tau')|X_i)\right]\xrightarrow{p}\min(\tau,\tau')-\tau\tau'-\mathbb{E}m_{1,\tau}(X)m_{1,\tau'}(X).$$

By the same argument in Lemma E.3,

$$\sup_{\tau,\tau'\in\Upsilon}\left|\frac{1}{n}\sum_{i=1}^{2n}\left(A_i-\frac{1}{2}\right)\left[F_1(\min(q_1(\tau),q_1(\tau'))|X_i)-F_1(q_1(\tau)|X_i)F_1(q_1(\tau')|X_i)\right]\right|\stackrel{p}{\longrightarrow}0$$

Therefore, uniformly over $\tau,\tau'\in\Upsilon$,

$$\overline{\mathbb{P}}\eta_{1,(j,1)}(\hat{q}_1(\tau),\tau)\eta_{1,(j,1)}(\hat{q}_1(\tau'),\tau')\stackrel{p}{\longrightarrow}\min(\tau,\tau')-\tau\tau'-\mathbb{E}m_{1,\tau}(X)m_{1,\tau'}(X).$$

To deal with the second term in (C.2), first denote

$$\mathcal{F}_3=\{(\tau-1\{Y\le q_1(\tau)+v\})\left(\tau'-1\{Y\le q_1(\tau')+v'\}\right):\tau,\tau'\in\Upsilon,|v|,|v'|\le L/\sqrt{n}\}.$$

Note $\mathcal{F}_3$ has an envelope $F=1$ and is nested by a VC-class of functions with a fixed VC-index. Then, by Lemma E.1,

$$\mathbb{E}\|\mathbb{P}_n-\overline{\mathbb{P}}\|_{\mathcal{F}_3}\lesssim 1/\sqrt{n}.$$

This implies, with probability greater than $1-\varepsilon$,

$$\sup_{\tau,\tau'\in\Upsilon}|\left(\mathbb{P}_n-\overline{\mathbb{P}}\right)\eta_{1,(j,1)}(\hat{q}_1(\tau),\tau)\eta_{1,(j,1)}(\hat{q}_1(\tau'),\tau')|\stackrel{p}{\longrightarrow}0. \qquad \text{(C.5)}$$

Since ε in (C.3) is arbitrary, we have, uniformly over $\tau,\tau'\in\Upsilon$,

$$I(\tau,\tau')\stackrel{p}{\longrightarrow}\min(\tau,\tau')-\tau\tau'-\mathbb{E}m_{1,\tau}(X)m_{1,\tau'}(X). \qquad \text{(C.6)}$$

By Lemma E.5, we have

$$\sup_{\tau,\tau'\in\Upsilon}|II(\tau,\tau')|=o_p(1)\quad\text{and}\quad\sup_{\tau,\tau'\in\Upsilon}|III(\tau,\tau')|=o_p(1).$$

For $IV(\tau,\tau')$, we note that

$$\begin{aligned}IV(\tau,\tau')=&\frac{1}{n}\sum_{j=1}^{n}m_{1,\tau}(X_{(j,1)})m_{1,\tau'}(X_{(j,1)})+R_{IV}(\tau,\tau')\\=&\frac{1}{n}\sum_{i=1}^{2n}A_im_{1,\tau}(X_i)m_{1,\tau'}(X_i)+R_{IV}(\tau,\tau')\\=&\frac{1}{2n}\sum_{i=1}^{2n}m_{1,\tau}(X_i)m_{1,\tau'}(X_i)+\frac{1}{n}\sum_{i=1}^{2n}\left(A_i-\frac{1}{2}\right)m_{1,\tau}(X_i)m_{1,\tau'}(X_i)+R_{IV}(\tau,\tau').\end{aligned} \qquad \text{(C.7)}$$

By the standard uniform convergence theorem (van der Vaart and Wellner (1996, Theorem 2.4.1)), uniformly over $\tau, \tau' \in \Upsilon$,

$$\frac{1}{2n}\sum_{i=1}^{2n} m_{1,\tau}(X_i)m_{1,\tau'}(X_i) \xrightarrow{p} \mathbb{E}m_{1,\tau}(X)m_{1,\tau'}(X).$$

Lemma E.6 further shows

$$\sup_{\tau,\tau'\in\Upsilon}|R_{IV}(\tau,\tau')| = o_p(1) \quad \text{and} \quad \sup_{\tau,\tau'\in\Upsilon}\left|\frac{1}{n}\sum_{i=1}^{2n}\left(A_i - \frac{1}{2}\right)m_{1,\tau}(X_i)m_{1,\tau'}(X_i)\right| = o_p(1).$$

Combining the above results, we have, uniformly over $\tau, \tau' \in \Upsilon$,

$$\tilde{\Sigma}^*_{1,1,1}(\tau,\tau') \xrightarrow{p} \min(\tau,\tau') - \tau\tau'.$$

Now we turn to $\tilde{\Sigma}^*_{1,1,2}(\tau,\tau')$. Recall $m_{0,\tau}(X_i,q) = \mathbb{E}\left(\tau - 1\{Y_i(0) \leq q\}|X_i\right)$ and define $\eta_{0,i}(q,\tau) = (\tau - 1\{Y_i(0) \leq q\}) - m_{0,\tau}(X_i,q)$. Then,

$$\begin{aligned}
\tilde{\Sigma}^*_{1,1,2}(\tau,\tau') =& \frac{1}{n}\sum_{j=1}^{n}\eta_{1,(j,1)}(\hat{q}_1(\tau),\tau)\eta_{0,(j,0)}(\hat{q}_0(\tau'),\tau') + \frac{1}{n}\sum_{j=1}^{n}\eta_{1,(j,1)}(\hat{q}_1(\tau),\tau)m_{0,\tau'}(X_{(j,0)},\hat{q}_0(\tau')) \\
&+ \frac{1}{n}\sum_{j=1}^{n}\eta_{0,(j,0)}(\hat{q}_0(\tau'),\tau')m_{1,\tau}(X_{(j,1)},\hat{q}_1(\tau)) + \frac{1}{n}\sum_{j=1}^{n}m_{1,\tau}(X_{(j,1)},\hat{q}_1(\tau))m_{0,\tau'}(X_{(j,0)},\hat{q}_0(\tau')) \\
=& \widetilde{I}(\tau,\tau') + \widetilde{II}(\tau,\tau') + \widetilde{III}(\tau,\tau') + \widetilde{IV}(\tau,\tau').
\end{aligned}$$

We derive the uniform limit for each term on the RHS of the above display. First, note that

$$\widetilde{I}(\tau,\tau') = \overline{\mathbb{P}}\eta_{1,(j,1)}(\hat{q}_1(\tau),\tau)\eta_{0,(j,0)}(\hat{q}_0(\tau'),\tau') + (\mathbb{P}_n - \overline{\mathbb{P}})\eta_{1,(j,1)}(\hat{q}_1(\tau),\tau)\eta_{0,(j,0)}(\hat{q}_0(\tau'),\tau'). \quad \text{(C.8)}$$

Similar to (C.4), we have

$$\sup_{\tau,\tau'\in\Upsilon}\left|\overline{\mathbb{P}}\eta_{1,(j,1)}(\hat{q}_1(\tau),\tau)\eta_{0,(j,0)}(\hat{q}_0(\tau'),\tau') - \overline{\mathbb{P}}\eta_{1,(j,1)}(q_1(\tau),\tau)\eta_{0,(j,0)}(q_0(\tau'),\tau')\right| \xrightarrow{p} 0.$$

Furthermore, because $(j,1) \neq (j,0)$, conditionally on $\{A_i, X_i\}_{i=1}^{2n}$, $\eta_{1,(j,1)}(q_1(\tau),\tau) \perp\!\!\!\perp \eta_{1,(j,0)}(q_0(\tau),\tau)$,

$$\overline{\mathbb{P}}\eta_{1,(j,1)}(q_1(\tau),\tau)\eta_{0,(j,0)}(q_0(\tau'),\tau') = 0.$$

Similar to (C.5), we have

$$\sup_{\tau,\tau'\in\Upsilon}\left|(\mathbb{P}_n - \overline{\mathbb{P}})\eta_{1,(j,1)}(\hat{q}_1(\tau),\tau)\eta_{0,(j,0)}(\hat{q}_0(\tau'),\tau')\right| \xrightarrow{p} 0.$$

This implies that, uniformly over $\tau, \tau' \in \Upsilon$, $\widetilde{I}(\tau,\tau') \xrightarrow{p} 0$. By the same argument as in the proof of Lemma E.5, we can show that

$$\sup_{\tau,\tau'\in\Upsilon}\left|\widetilde{II}(\tau,\tau')\right| \xrightarrow{p} 0 \quad \text{and} \quad \sup_{\tau,\tau'\in\Upsilon}\left|\widetilde{III}(\tau,\tau')\right| \xrightarrow{p} 0.$$

Last, by the same argument in the proof of Lemma E.6, we can show that, uniformly over $\tau, \tau' \in \Upsilon$,

$$\begin{aligned}
\widetilde{IV}(\tau,\tau') &= \frac{1}{n}\sum_{j=1}^{n} m_{1,\tau}(X_{(j,1)})m_{0,\tau'}(X_{(j,0)}) + o_p(1) \\
&= \frac{1}{n}\sum_{j=1}^{n} m_{1,\tau}(X_{(j,1)})m_{0,\tau'}(X_{(j,1)}) + \frac{1}{n}\sum_{j=1}^{n} m_{1,\tau}(X_{(j,1)})[m_{0,\tau'}(X_{(j,0)}) - m_{0,\tau'}(X_{(j,1)})] + o_p(1) \\
&\xrightarrow{p} \mathbb{E}m_{1,\tau}(X)m_{0,\tau'}(X),
\end{aligned}$$

where the $o_p(1)$ holds uniformly over $\tau, \tau' \in \Upsilon$, and the last line holds because $m_{1,\tau}(x)$ is bounded and $m_{0,\tau}(x)$ is Lipschitz.

Combining the above results, we have uniformly over $\tau, \tau' \in \Upsilon$,

$$\tilde{\Sigma}^*_{1,1,2}(\tau,\tau') \xrightarrow{p} \mathbb{E}m_{1,\tau}(X)m_{0,\tau'}(X).$$

The limits of $\tilde{\Sigma}^*_{1,2,1}$ and $\tilde{\Sigma}^*_{1,2,2}$ can be derived similarly. To sum up, we have established that, uniformly over $\tau, \tau' \in \Upsilon$,

$$\tilde{\Sigma}^*_1(\tau,\tau') \xrightarrow{p} \begin{pmatrix} \min(\tau,\tau') - \tau\tau' & \mathbb{E}m_{1,\tau}(X_i)m_{0,\tau'}(X_i) \\ \mathbb{E}m_{0,\tau}(X_i)m_{1,\tau'}(X_i) & \min(\tau,\tau') - \tau\tau' \end{pmatrix}.$$

Lemma E.7 shows $S^*_{n,1}(\tau)$ is stochastically equicontinuous and tight. This concludes the proof of this step.

Step (2). Given the data, $S^*_{n,2}(\tau)$ is a Gaussian process with covariance kernel

$$\tilde{\Sigma}^*_2(\tau,\tau') = \begin{pmatrix} \tilde{\Sigma}^*_{2,1,1}(\tau,\tau') & \tilde{\Sigma}^*_{2,1,2}(\tau,\tau') \\ \tilde{\Sigma}^*_{2,2,1}(\tau,\tau') & \tilde{\Sigma}^*_{2,2,2}(\tau,\tau') \end{pmatrix}$$

where

$$\begin{aligned}
\tilde{\Sigma}^*_{2,1,1}(\tau,\tau') =& \frac{1}{n}\sum_{k=1}^{\lfloor n/2 \rfloor}\left[\left(\tau - 1\{Y_{(k,1)} \le \hat{q}_1(\tau)\}\right) - \left(\tau - 1\{Y_{(k,3)} \le \hat{q}_1(\tau)\}\right)\right] \\
&\times \left[\left(\tau' - 1\{Y_{(k,1)} \le \hat{q}_1(\tau')\}\right) - \left(\tau' - 1\{Y_{(k,3)} \le \hat{q}_1(\tau')\}\right)\right],
\end{aligned}$$

$$\tilde{\Sigma}^*_{2,1,2}(\tau,\tau') = \frac{1}{n}\sum_{k=1}^{\lfloor n/2 \rfloor} \left[\left(\tau - 1\{Y_{(k,1)} \le \hat{q}_1(\tau)\}\right) - \left(\tau - 1\{Y_{(k,3)} \le \hat{q}_1(\tau)\}\right)\right]$$
$$\times \left[\left(\tau' - 1\{Y_{(k,2)} \le \hat{q}_0(\tau')\}\right) - \left(\tau' - 1\{Y_{(k,4)} \le \hat{q}_0(\tau')\}\right)\right],$$

$$\tilde{\Sigma}^*_{2,2,1}(\tau,\tau') = \frac{1}{n}\sum_{k=1}^{\lfloor n/2 \rfloor} \left[\left(\tau - 1\{Y_{(k,2)} \le \hat{q}_0(\tau)\}\right) - \left(\tau - 1\{Y_{(k,4)} \le \hat{q}_0(\tau)\}\right)\right]$$
$$\times \left[\left(\tau' - 1\{Y_{(k,1)} \le \hat{q}_1(\tau')\}\right) - \left(\tau' - 1\{Y_{(k,3)} \le \hat{q}_1(\tau')\}\right)\right],$$

and

$$\tilde{\Sigma}^*_{2,2,2}(\tau,\tau') = \frac{1}{n}\sum_{k=1}^{\lfloor n/2 \rfloor} \left[\left(\tau - 1\{Y_{(k,2)} \le \hat{q}_0(\tau)\}\right) - \left(\tau - 1\{Y_{(k,4)} \le \hat{q}_0(\tau)\}\right)\right]$$
$$\times \left[\left(\tau' - 1\{Y_{(k,2)} \le \hat{q}_0(\tau')\}\right) - \left(\tau' - 1\{Y_{(k,4)} \le \hat{q}_0(\tau')\}\right)\right].$$

In the following, we derive the limit of $\tilde{\Sigma}^*_2(\tau,\tau')$. For $\tilde{\Sigma}^*_{2,1,1}(\tau,\tau')$, we have

$$\begin{aligned}
&\tilde{\Sigma}^*_{2,1,1}(\tau,\tau') \\
&= \frac{1}{n}\sum_{k=1}^{\lfloor n/2 \rfloor} \left[\eta_{1,(k,1)}(\hat{q}_1(\tau),\tau) - \eta_{1,(k,3)}(\hat{q}_1(\tau),\tau)\right]\left[\eta_{1,(k,1)}(\hat{q}_1(\tau'),\tau') - \eta_{1,(k,3)}(\hat{q}_1(\tau'),\tau')\right] \\
&\quad + \frac{1}{n}\sum_{k=1}^{\lfloor n/2 \rfloor} \left[\eta_{1,(k,1)}(\hat{q}_1(\tau),\tau) - \eta_{1,(k,3)}(\hat{q}_1(\tau),\tau)\right]\left[m_{1,\tau'}(X_{(k,1)},\hat{q}_1(\tau')) - m_{1,\tau'}(X_{(k,3)},\hat{q}_1(\tau'))\right] \\
&\quad + \frac{1}{n}\sum_{k=1}^{\lfloor n/2 \rfloor} \left[m_{1,\tau}(X_{(k,1)},\hat{q}_1(\tau)) - m_{1,\tau}(X_{(k,3)},\hat{q}_1(\tau))\right]\left[\eta_{1,(k,1)}(\hat{q}_1(\tau'),\tau') - \eta_{1,(k,3)}(\hat{q}_1(\tau'),\tau')\right] \\
&\quad + \frac{1}{n}\sum_{k=1}^{\lfloor n/2 \rfloor} \left[m_{1,\tau}(X_{(k,1)},\hat{q}_1(\tau)) - m_{1,\tau}(X_{(k,3)},\hat{q}_1(\tau))\right]\left[m_{1,\tau'}(X_{(k,1)},\hat{q}_1(\tau')) - m_{1,\tau'}(X_{(k,3)},\hat{q}_1(\tau'))\right] \\
&\equiv \widehat{I}(\tau,\tau') + \widehat{II}(\tau,\tau') + \widehat{III}(\tau,\tau') + \widehat{IV}(\tau,\tau').
\end{aligned}$$

Also note that

$$\begin{aligned}
&\widehat{I}(\tau,\tau') \\
&= \frac{1}{n}\sum_{k=1}^{\lfloor n/2 \rfloor} \left[\eta_{1,(k,1)}(\hat{q}_1(\tau),\tau)\eta_{1,(k,1)}(\hat{q}_1(\tau'),\tau') + \eta_{1,(k,3)}(\hat{q}_1(\tau),\tau)\eta_{1,(k,3)}(\hat{q}_1(\tau'),\tau')\right] \\
&\quad - \frac{1}{n}\sum_{k=1}^{\lfloor n/2 \rfloor} \eta_{1,(k,1)}(\hat{q}_1(\tau),\tau)\eta_{1,(k,3)}(\hat{q}_1(\tau'),\tau') - \frac{1}{n}\sum_{k=1}^{\lfloor n/2 \rfloor} \eta_{1,(k,1)}(\hat{q}_1(\tau'),\tau')\eta_{1,(k,3)}(\hat{q}_1(\tau),\tau)
\end{aligned}$$

$$= \frac{1}{n}\sum_{j=1}^{n}\eta_{1,(j,1)}(\hat{q}_1(\tau),\tau)\eta_{1,(j,1)}(\hat{q}_1(\tau'),\tau')$$
$$- \frac{1}{n}\sum_{k=1}^{\lfloor n/2 \rfloor}\eta_{1,(k,1)}(\hat{q}_1(\tau),\tau)\eta_{1,(k,3)}(\hat{q}_1(\tau'),\tau') - \frac{1}{n}\sum_{k=1}^{\lfloor n/2 \rfloor}\eta_{1,(k,1)}(\hat{q}_1(\tau'),\tau')\eta_{1,(k,3)}(\hat{q}_1(\tau),\tau).$$

The first term on the RHS of the above display is just $I(\tau,\tau')$ defined in Step (1), whose limit is established in (C.6). For the second and third terms, we note that $(k,1) \neq (k,3)$, which implies, given $\{X_i, A_i\}_{i=1}^{2n}$, $(\eta_{1,(k,1)}(\hat{q}_1(\tau),\tau), \eta_{1,(k,1)}(\hat{q}_1(\tau'),\tau')) \perp\!\!\!\perp (\eta_{1,(k,3)}(\hat{q}_1(\tau),\tau), \eta_{1,(k,3)}(\hat{q}_1(\tau'),\tau'))$. Then, by the same argument in (C.8) and the discussion below, we have

$$\sup_{\tau,\tau'\in\Upsilon}\left|\frac{1}{n}\sum_{k=1}^{\lfloor n/2 \rfloor}\eta_{1,(k,1)}(\hat{q}_1(\tau),\tau)\eta_{1,(k,3)}(\hat{q}_1(\tau'),\tau')\right| \xrightarrow{p} 0$$

and

$$\sup_{\tau,\tau'\in\Upsilon}\left|\frac{1}{n}\sum_{k=1}^{\lfloor n/2 \rfloor}\eta_{1,(k,1)}(\hat{q}_1(\tau'),\tau')\eta_{1,(k,3)}(\hat{q}_1(\tau),\tau)\right| \xrightarrow{p} 0.$$

This implies that, uniformly over $\tau,\tau' \in \Upsilon$,

$$\widehat{I}(\tau,\tau') \xrightarrow{p} \min(\tau,\tau') - \tau\tau' - \mathbb{E}m_{1,\tau}(X)m_{1,\tau'}(X).$$

By the same argument in the proof of Lemma E.5, we have

$$\sup_{\tau,\tau'\in\Upsilon}\left|\widehat{II}(\tau,\tau')\right| \xrightarrow{p} 0 \quad \text{and} \quad \sup_{\tau,\tau'\in\Upsilon}\left|\widehat{III}(\tau,\tau')\right| \xrightarrow{p} 0.$$

For $\widehat{IV}(\tau,\tau')$, we note $m_{1,\tau}(x,q)$ is Lipschitz in x by Assumption 2. Therefore, by Assumption 4, we have

$$\sup_{\tau,\tau'\in\Upsilon}\left|\widehat{IV}(\tau,\tau')\right| \lesssim \frac{1}{n}\sum_{k=1}^{\lfloor n/2 \rfloor}||X_{(k,1)} - X_{(k,3)}||_2^2 \xrightarrow{p} 0.$$

Combining the above results, we show that, uniformly over $\tau,\tau' \in \Upsilon$,

$$\tilde{\Sigma}^*_{2,1,1}(\tau,\tau') \xrightarrow{p} \min(\tau,\tau') - \tau\tau' - \mathbb{E}m_{1,\tau}(X)m_{1,\tau'}(X).$$

For $\tilde{\Sigma}^*_{2,1,2}(\tau,\tau')$, we have

$\tilde{\Sigma}^*_{2,1,1}(\tau,\tau')$

$$
\begin{aligned}
&= \frac{1}{n}\sum_{k=1}^{\lfloor n/2 \rfloor}\left[\eta_{1,(k,1)}(\hat{q}_1(\tau),\tau) - \eta_{1,(k,3)}(\hat{q}_1(\tau),\tau)\right]\left[\eta_{0,(k,2)}(\hat{q}_0(\tau'),\tau') - \eta_{0,(k,4)}(\hat{q}_0(\tau'),\tau')\right] \\
&\quad + \frac{1}{n}\sum_{k=1}^{\lfloor n/2 \rfloor}\left[\eta_{1,(k,1)}(\hat{q}_1(\tau),\tau) - \eta_{1,(k,3)}(\hat{q}_1(\tau),\tau)\right]\left[m_{0,\tau'}(X_{(k,2)},\hat{q}_0(\tau')) - m_{0,\tau'}(X_{(k,4)},\hat{q}_0(\tau'))\right] \\
&\quad + \frac{1}{n}\sum_{k=1}^{\lfloor n/2 \rfloor}\left[m_{1,\tau}(X_{(k,1)},\hat{q}_1(\tau)) - m_{1,\tau}(X_{(k,3)},\hat{q}_1(\tau))\right]\left[\eta_{0,(k,2)}(\hat{q}_0(\tau'),\tau') - \eta_{0,(k,4)}(\hat{q}_0(\tau'),\tau')\right] \\
&\quad + \frac{1}{n}\sum_{k=1}^{\lfloor n/2 \rfloor}\left[m_{1,\tau}(X_{(k,1)},\hat{q}_1(\tau)) - m_{1,\tau}(X_{(k,3)},\hat{q}_1(\tau))\right]\left[m_{0,\tau'}(X_{(k,2)},\hat{q}_0(\tau')) - m_{0,\tau'}(X_{(k,4)},\hat{q}_0(\tau'))\right] \\
&\equiv \overline{I}(\tau,\tau') + \overline{II}(\tau,\tau') + \overline{III}(\tau,\tau') + \overline{IV}(\tau,\tau').
\end{aligned}
$$

Because $(k,1),\cdots,(k,4)$ are distinctive,

$$
\left(\eta_{1,(k,1)}(q,\tau),\eta_{1,(k,3)}(q,\tau),\eta_{0,(k,2)}(q',\tau),\eta_{0,(k,4)}(q',\tau)\right)
$$

are mutually independent conditionally on $\{X_i,A_i\}_{i=1}^{2n}$. Then, by the same arguments as in (C.4) and (C.5), we have

$$
\sup_{\tau,\tau'\in\Upsilon}|\overline{I}(\tau,\tau')| \xrightarrow{p} 0.
$$

By the same argument as in the proof of Lemma E.5, we also have

$$
\sup_{\tau,\tau'\in\Upsilon}|\overline{II}(\tau,\tau')| \xrightarrow{p} 0 \quad \text{and} \quad \sup_{\tau,\tau'\in\Upsilon}|\overline{III}(\tau,\tau')| \xrightarrow{p} 0.
$$

Last, by Assumption 4, we have

$$
\begin{aligned}
\sup_{\tau,\tau'\in\Upsilon}|\overline{IV}(\tau,\tau')| &\lesssim \frac{1}{n}\sum_{k=1}^{\lfloor n/2 \rfloor}||X_{(k,1)} - X_{(k,3)}||_2||X_{(k,2)} - X_{(k,4)}||_2 \\
&\lesssim \frac{1}{n}\sum_{k=1}^{\lfloor n/2 \rfloor}||X_{(k,1)} - X_{(k,3)}||_2^2 + \frac{1}{n}\sum_{k=1}^{\lfloor n/2 \rfloor}||X_{(k,2)} - X_{(k,4)}||_2^2 \xrightarrow{p} 0.
\end{aligned}
$$

Combining the above results, we have

$$
\sup_{\tau,\tau'\in\Upsilon}|\tilde{\Sigma}^*_{2,1,2}(\tau,\tau')| \xrightarrow{p} 0.
$$

We can derive the limits of $\tilde{\Sigma}^*_{2,2,1}(\tau,\tau')$ and $\tilde{\Sigma}^*_{2,2,2}(\tau,\tau')$ in the same manner. To sum up, uniformly over $\tau, \tau' \in \Upsilon$, we have

$$\tilde{\Sigma}_2^* \xrightarrow{p} \begin{pmatrix} \min(\tau,\tau') - \tau\tau' - \mathbb{E}m_{1,\tau}(X_i)m_{1,\tau'}(X_i) & 0 \\ 0 & \min(\tau,\tau') - \tau\tau' - \mathbb{E}m_{0,\tau}(X_i)m_{0,\tau'}(X_i) \end{pmatrix}$$

The stochastic equicontinuity and tightness of $S_{n,2}^*(\tau)$ can be established similarly to $S_{n,1}^*(\tau)$.

Step (3). Because both $S_n(\tau)$ and $S_n^*(\tau)$ are stochastically equicontinuous and tight, we can apply Kato (2009, Theorem 2) and have

$$\sqrt{n}(\hat{\beta}^*(\tau) - \beta(\tau)) = Q^{-1} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} (S_n(\tau) + S_n^*(\tau)) + R^*(\tau), \tag{C.9}$$

where $\sup_{\tau\in\Upsilon} ||R^*(\tau)||_2 = o_p(1)$. Taking the difference between (C.9) and (A.13), we have

$$\sqrt{n}(\hat{\beta}^*(\tau) - \hat{\beta}(\tau)) = Q^{-1} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} S_n^*(\tau) + \tilde{R}^*(\tau),$$

where $\sup_{\tau\in\Upsilon} ||\tilde{R}^*(\tau)||_2 = o_p(1)$. In addition, given the data, $S_{n,1}^*(\tau)$ and $S_{n,2}^*(\tau)$ are independent. Steps (1) and (2) show that uniformly over $\tau \in \Upsilon$ and conditionally on data, $S_n^*(\tau) = \frac{S_{n,1}^*(\tau)+S_{n,2}^*(\tau)}{\sqrt{2}}$ converges to a Gaussian process with covariance kernel

$$\frac{1}{2}\left[\tilde{\Sigma}_1(\tau,\tau') + \tilde{\Sigma}_2(\tau,\tau')\right],$$

where $\tilde{\Sigma}_1(\tau,\tau')$ and $\tilde{\Sigma}_2(\tau,\tau')$ are defined in (A.14) and (A.15), respectively. The weak limit of $S_n^*(\tau)$ given data coincides with the weak limit of $S_n(\tau)$. This implies, given the data, that

$$\sqrt{n}(\hat{q}^*(\tau) - \hat{q}(\tau)) \rightsquigarrow \mathcal{B}(\tau),$$

where $\mathcal{B}(\tau)$ is the Gaussian process defined in Theorem 3.1. This concludes the proof.

D Proof of Theorem 4.3

We first focus on $\hat{q}_{ipw,1}^w(\tau)$. Let $u \in \Re$ and

$$\tilde{L}_n^w(u,\tau) = \sum_{i=1}^{2n} \frac{\xi_i A_i}{2\hat{A}_i}\left[\rho_\tau(Y_i - q_1(\tau) - u/\sqrt{n}) - \rho_\tau(Y_i - q_1(\tau))\right].$$

Then, by change of variables, we have

$$\sqrt{n}(\hat{q}_{ipw,1}^w(\tau) - q_1(\tau)) = \arg\min_u \tilde{L}_n^w(u,\tau).$$

Notice that $\tilde{L}_n^w(u,\tau)$ is convex in u for each τ and bounded in τ for each u. In the following, we divide the proof into three steps. In Step (1), we show that there exists

$$\tilde{g}_n^w(u,\tau) = -u'\widetilde{W}_{n,1}^w(\tau) + \frac{f_1(q_1(\tau))u^2}{2}$$

such that for each u,

$$\sup_{\tau\in\Upsilon} |\tilde{L}_n^w(u,\tau) - \tilde{g}_n^w(u,\tau)| \xrightarrow{p} 0.$$

In Step (2), we show $\widetilde{W}_{n,1}^w(\tau)$ as a stochastic process over $\tau \in \Upsilon$ is tight. Then by Kato (2009, Theorem 2), we have

$$\sqrt{n}(\hat{q}_{ipw,1}^w(\tau) - q_1(\tau)) = [f_1(q_1(\tau))]^{-1}\widetilde{W}_{n,1}^w(\tau) + \tilde{r}_{n,1}(\tau),$$

where $\sup_{\tau\in\Upsilon} |\tilde{r}_{n,1}(\tau)| = o_p(1)$. For the same reason, we can show

$$\sqrt{n}(\hat{q}_{ipw,0}^w(\tau) - q_0(\tau)) = [f_0(q_0(\tau))]^{-1}\widetilde{W}_{n,0}^w(\tau) + \tilde{r}_{n,0}(\tau),$$

for some $\widetilde{W}_{n,0}^w(\tau)$ to be specified later and $\sup_{\tau\in\Upsilon} |\tilde{r}_{n,0}(\tau)| = o_p(1)$. Last, in Step (3), we establish the weak convergence of

$$\sqrt{n}(\hat{q}_{ipw}^w(\tau) - \hat{q}(\tau))$$

conditionally on data.

Step (1). Similar to Step (1) in the previous section, we have

$$\tilde{L}_n^w(u,\tau) = -\widetilde{W}_{n,1}^w(\tau)u + \tilde{Q}_n^w(u,\tau),$$

where

$$\widetilde{W}_{n,1}^w(\tau) = \sum_{i=1}^{2n} \frac{\xi_i A_i}{2\sqrt{n}\hat{A}_i} \left(\tau - 1\{Y_i(1) \le q_1(\tau)\}\right),$$

and

$$\begin{aligned}\tilde{Q}_n^w(u,\tau) &= \sum_{i=1}^{2n} \frac{\xi_i A_i}{2\hat{A}_i} \int_0^{\frac{u}{\sqrt{n}}} \left(1\{Y_i(1) - q_1(\tau) \le v\} - 1\{Y_i(1) - q_1(\tau) \le 0\}\right) dv \\ &= \sum_{i=1}^{2n} \xi_i A_i \int_0^{\frac{u}{\sqrt{n}}} \left(1\{Y_i(1) - q_1(\tau) \le v\} - 1\{Y_i(1) - q_1(\tau) \le 0\}\right) dv\end{aligned}$$

$$+\sum_{i=1}^{2n}\frac{\xi_i A_i(1/2-\hat{A}_i)}{\hat{A}_i}\int_0^{\frac{u}{\sqrt{n}}}\left(1\{Y_i(1)-q_1(\tau)\le v\}-1\{Y_i(1)-q_1(\tau)\le 0\}\right)dv$$

$$\equiv \tilde{Q}^w_{n,1}(u,\tau)+\tilde{Q}^w_{n,2}(u,\tau). \tag{D.1}$$

Exactly the same as $Q^w_{n,1}(u,\tau)$ in Section B, we have

$$\sup_{\tau\in\Upsilon}\left|\tilde{Q}^w_{n,1}(u,\tau)-\frac{f_1(q_1(\tau))u^2}{2}\right|=o_p(1). \tag{D.2}$$

For $\tilde{Q}^w_{n,2}(u,\tau)$, we have, with probability approaching one,

$$\begin{aligned}|\tilde{Q}^w_{n,2}(u,\tau)| &\le \max_{i\in[2n]}|\hat{A}_i-1/2|\sum_{i=1}^{2n}\frac{\xi_i}{1/2-\max_{i\in[2n]}|\hat{A}_i-1/2|}1\{|Y_i(1)-q_1(\tau)|\le u/\sqrt{n}\}\frac{|u|}{\sqrt{n}}\\ &\le \max_{i\in[2n]}|\hat{A}_i-1/2|\sum_{i=1}^{2n}4\xi_i 1\{|Y_i(1)-q_1(\tau)|\le u/\sqrt{n}\}\frac{|u|}{\sqrt{n}},\end{aligned} \tag{D.3}$$

where the second inequality follows the fact that, w.p.a.1, $|\hat{A}_i-1/2|\le 1/4$ as proved in Lemma E.8. Because $\{\xi_i, Y_i(1)\}_{i\in[2n]}$ are i.i.d., by the usual maximal inequality, we can show that

$$\sup_{\tau\in\Upsilon}\left|\sum_{i=1}^{2n}4\xi_i 1\{|Y_i(1)-q_1(\tau)|\le u/\sqrt{n}\}\frac{|u|}{\sqrt{n}}-\mathbb{E}\sum_{i=1}^{2n}4\xi_i 1\{|Y_i(1)-q_1(\tau)|\le u/\sqrt{n}\}\frac{|u|}{\sqrt{n}}\right|=o_p(1). \tag{D.4}$$

In addition,

$$\mathbb{E}\sum_{i=1}^{2n}4\xi_i 1\{|Y_i(1)-q_1(\tau)|\le u/\sqrt{n}\}\frac{|u|}{\sqrt{n}}\lesssim\sqrt{n}u\left(F_1(q_1(\tau)+\frac{|u|}{\sqrt{n}})-F_1(q_1(\tau)-\frac{|u|}{\sqrt{n}})\right)\lesssim u^2. \tag{D.5}$$

Combining (D.3)–(D.5) with the fact that $\max_{i\in[2n]}|\hat{A}_i-1/2|=o_p(1)$ as proved in Lemma E.8, we have

$$\sup_{\tau\in\Upsilon}|\tilde{Q}^w_{n,2}(u,\tau)|=o_p(1).$$

This concludes the proof of Step (1).

Step (2). We have

$$\widetilde{W}^w_{n,1}(\tau)=\sum_{i=1}^{2n}\frac{\xi_i A_i}{\sqrt{n}}\left(\tau-1\{Y_i(1)\le q_1(\tau)\}\right)-\sum_{i=1}^{2n}\frac{2\xi_i A_i(\hat{A}_i-1/2)}{\sqrt{n}}\left(\tau-1\{Y_i(1)\le q_1(\tau)\}\right)$$

$$
\begin{aligned}
&+\sum_{i=1}^{2n} \frac{2\xi_i A_i (1/2-\hat{A}_i)^2}{\sqrt{n}\hat{A}_i} \left(\tau - 1\{Y_i(1) \le q_1(\tau)\}\right) \\
\equiv& \widetilde{W}^w_{n,1,1}(\tau) - \widetilde{W}^w_{n,1,2}(\tau) + \widetilde{W}^w_{n,1,3}(\tau). \qquad \text{(D.6)}
\end{aligned}
$$

First, $\widetilde{W}^w_{n,1,1}(\tau)$ is tight following the exact same argument as in Step (2) of Section B. Second, we have

$$
\begin{aligned}
\widetilde{W}^w_{n,1,2}(\tau) =& \sum_{i=1}^{2n} \frac{\xi_i m_{1,\tau}(X_i)(\hat{A}_i - 1/2)}{\sqrt{n}} + \sum_{i=1}^{2n} \frac{2\xi_i (A_i - 1/2) m_{1,\tau}(X_i)(\hat{A}_i - 1/2)}{\sqrt{n}} + \sum_{i=1}^{2n} \frac{2\xi_i A_i \eta_{1,i}(\tau)(\hat{A}_i - 1/2)}{\sqrt{n}} \\
\equiv& \; I(\tau) + II(\tau) + III(\tau).
\end{aligned}
$$

Lemma E.9 shows

$$
\sup_{\tau \in \Upsilon} \left| I(\tau) - \sum_{i=1}^{2n} \frac{\xi_i m_{1,\tau}(X_i)(A_i - 1/2)}{\sqrt{n}} \right| = o_p(1),
$$

$$
\sup_{\tau \in \Upsilon} |II(\tau)| = o_p(1), \quad \text{and} \quad \sup_{\tau \in \Upsilon} |III(\tau)| = o_p(1).
$$

Combining the above results, we have

$$
\sup_{\tau \in \Upsilon} \left| \widetilde{W}^w_{n,1,2}(\tau) - \sum_{i=1}^{2n} \frac{\xi_i m_{1,\tau}(X_i)(A_i - 1/2)}{\sqrt{n}} \right| = o_p(1). \qquad \text{(D.7)}
$$

Last, we have, w.p.a.1,

$$
\begin{aligned}
\sup_{\tau \in \Upsilon} |\widetilde{W}^w_{n,1,3}(\tau)| \le& \sum_{i=1}^{2n} \frac{2\xi_i}{\sqrt{n}(1/2 - \max_{i \in [2n]} |1/2 - \hat{A}_i|)} (1/2 - \hat{A}_i)^2 \\
\lesssim& \frac{4}{\sqrt{n}} \sum_{i=1} \xi_i (1/2 - \hat{A}_i)^2 = o_p(1), \qquad \text{(D.8)}
\end{aligned}
$$

where the first inequality holds because $\sup_{\tau \in \Upsilon} |\tau - 1\{Y_i(1) \le q_1(\tau)\}| \le 1$, the second inequality holds because $\max_i |1/2 - \hat{A}_i| \le 1/4$ w.p.a.1 as proved in Lemma E.8, and the last inequality holds due to Lemma E.8.

Combining (D.6)–(D.8), we have

$$
\widetilde{W}^w_{n,1}(\tau) = \sum_{i=1}^{2n} \frac{\xi_i A_i \eta_{1,i}(\tau)}{\sqrt{n}} + \sum_{i=1}^{2n} \frac{\xi_i m_{1,\tau}(X_i)}{2\sqrt{n}} + o_p(1),
$$

where the $o_p(1)$ term holds uniformly over $\tau \in \Upsilon$. By (B.9) and the argument above, we can show $\sum_{i=1}^{2n} \frac{\xi_i A_i \eta_{1,i}(\tau)}{\sqrt{n}}$ as a stochastic process over $\tau \in \Upsilon$ is stochastically equicontinuous and tight. Furthermore, $\{\xi_i, X_i\}_{i\in[2n]}$ is a sequence of i.i.d. random variables. Then, by the usual maximal inequality, we can show $\sum_{i=1}^{2n} \frac{\xi_i m_{1,\tau}(X_i)}{2\sqrt{n}}$ as a stochastic process over $\tau \in \Upsilon$ is stochastically equicontinuous and tight. This implies, $\widetilde{W}_{n,1}^{w}(\tau)$ as a stochastic process over $\tau \in \Upsilon$ is stochastically equicontinuous and tight, and thus, is stochastically equicontinuous conditionally on data by the Markov inequality. Therefore, we have

$$\sqrt{n}(\hat{q}_{ipw,1}^{w}(\tau) - q_1(\tau)) = \frac{1}{f_1(q_1(\tau))}\left(\sum_{i=1}^{2n} \frac{\xi_i A_i \eta_{1,i}(\tau)}{\sqrt{n}} + \sum_{i=1}^{2n} \frac{\xi_i m_{1,\tau}(X_i)}{2\sqrt{n}}\right) + \tilde{r}_{n,1}(\tau),$$

where $\sup_{\tau\in\Upsilon} |\tilde{r}_{n,1}(\tau)| = o_p(1)$. Similarly, we can show that

$$\sqrt{n}(\hat{q}_{ipw,0}^{w}(\tau) - q_0(\tau)) = \frac{1}{f_0(q_0(\tau))}\left(\sum_{i=1}^{2n} \frac{\xi_i (1-A_i) \eta_{0,i}(\tau)}{\sqrt{n}} + \sum_{i=1}^{2n} \frac{\xi_i m_{0,\tau}(X_i)}{2\sqrt{n}}\right) + \tilde{r}_{n,0}(\tau),$$

where $\sup_{\tau\in\Upsilon} |\tilde{r}_{n,1}(\tau)| = o_p(1)$.

Step (3). In the proof of Theorem 3.1, we established that

$$\begin{aligned}
&\sqrt{n}(\hat{q}(\tau) - q(\tau)) \\
=&\frac{1}{f_1(q_1(\tau))}\left(\sum_{i=1}^{2n} \frac{\xi_i A_i \eta_{1,i}(\tau)}{\sqrt{n}} + \sum_{i=1}^{2n} \frac{\xi_i m_{1,\tau}(X_i)}{2\sqrt{n}}\right) \\
&- \frac{1}{f_0(q_0(\tau))}\left(\sum_{i=1}^{2n} \frac{\xi_i (1-A_i) \eta_{0,i}(\tau)}{\sqrt{n}} + \sum_{i=1}^{2n} \frac{\xi_i m_{0,\tau}(X_i)}{2\sqrt{n}}\right) + r_b(\tau),
\end{aligned}$$

where $\sup_{\tau\in\Upsilon} |r_b(\tau)| = o_p(1)$. Then, we have

$$\begin{aligned}
\sqrt{n}(\hat{q}_{ipw}^{w}(\tau) - \hat{q}(\tau)) =&\frac{1}{f_1(q_1(\tau))}\left(\sum_{i=1}^{2n} \frac{(\xi_i - 1) A_i \eta_{1,i}(\tau)}{\sqrt{n}}\right) - \frac{1}{f_0(q_0(\tau))}\left(\sum_{i=1}^{2n} \frac{(\xi_i - 1)(1-A_i) \eta_{0,i}(\tau)}{\sqrt{n}}\right) \\
&+ \sum_{i=1}^{2n} \frac{(\xi_i - 1)}{2\sqrt{n}}\left(\frac{m_{1,\tau}(X_i)}{f_1(q_1(\tau))} - \frac{m_{0,\tau}(X_i)}{f_0(q_0(\tau))}\right) + \tilde{r}_b(r),
\end{aligned}$$

where $\sup_{\tau\in\Upsilon} |\tilde{r}_b(\tau)| = o_p(1)$. The conditional stochastic equicontinuity of the first three terms on the RHS of the above display has been established in Step (2). Here, we only need to determine the covariance kernel of $\sqrt{n}(\hat{q}_{ipw}^{w}(\tau) - \hat{q}(\tau))$ given data. Specifically, the covariance kernel is the limit of the display below:

$$\frac{1}{f_1(q_1(\tau))f_1(q_1(\tau'))}\sum_{i=1}^{2n} \frac{A_i \eta_{1,i}(\tau)\eta_{1,i}(\tau')}{n} + \frac{1}{f_0(q_0(\tau))f_0(q_0(\tau'))}\sum_{i=1}^{2n} \frac{(1-A_i) \eta_{0,i}(\tau)\eta_{0,i}(\tau')}{n}$$

$$
\begin{aligned}
&+\sum_{i=1}^{2n}\frac{1}{4n}\left(\frac{m_{1,\tau}(X_i)}{f_1(q_1(\tau))}-\frac{m_{0,\tau}(X_i)}{f_0(q_0(\tau))}\right)\left(\frac{m_{1,\tau'}(X_i)}{f_1(q_1(\tau'))}-\frac{m_{0,\tau'}(X_i)}{f_0(q_0(\tau'))}\right)\\
&+\frac{1}{2n}\sum_{i=1}^{2n}\frac{(1-A_i)\eta_{0,i}(\tau)}{f_0(q_0(\tau))}\left(\frac{m_{1,\tau'}(X_i)}{f_1(q_1(\tau'))}-\frac{m_{0,\tau'}(X_i)}{f_0(q_0(\tau'))}\right)+\frac{1}{2n}\sum_{i=1}^{2n}\frac{A_i\eta_{1,i}(\tau)}{f_1(q_1(\tau))}\left(\frac{m_{1,\tau'}(X_i)}{f_1(q_1(\tau'))}-\frac{m_{0,\tau'}(X_i)}{f_0(q_0(\tau'))}\right)\\
&+\frac{1}{2n}\sum_{i=1}^{2n}\frac{(1-A_i)\eta_{0,i}(\tau')}{f_0(q_0(\tau'))}\left(\frac{m_{1,\tau}(X_i)}{f_1(q_1(\tau))}-\frac{m_{0,\tau}(X_i)}{f_0(q_0(\tau))}\right)+\frac{1}{2n}\sum_{i=1}^{2n}\frac{A_i\eta_{1,i}(\tau')}{f_1(q_1(\tau'))}\left(\frac{m_{1,\tau}(X_i)}{f_1(q_1(\tau))}-\frac{m_{0,\tau}(X_i)}{f_0(q_0(\tau))}\right).
\end{aligned}
\tag{D.9}
$$

Note that (B.12) implies

$$
\begin{aligned}
\frac{1}{f_1(q_1(\tau))f_1(q_1(\tau'))}\sum_{i=1}^{2n}\frac{A_i\eta_{1,i}(\tau)\eta_{1,i}(\tau')}{n} &\xrightarrow{p}\frac{\min(\tau,\tau')-\mathbb{E}F_1(q_1(\tau)|X_i)F_1(q_1(\tau')|X_i)}{f_1(q_1(\tau))f_1(q_1(\tau'))}\\
&=\frac{\min(\tau,\tau')-\tau\tau'-\mathbb{E}m_{1,\tau}(X_i)m_{1,\tau'}(X_i)}{f_1(q_1(\tau))f_1(q_1(\tau'))}.
\end{aligned}
$$

Similarly,

$$
\frac{1}{f_0(q_0(\tau))f_0(q_0(\tau'))}\sum_{i=1}^{2n}\frac{(1-A_i)\eta_{0,i}(\tau)\eta_{0,i}(\tau')}{n}\xrightarrow{p}\frac{\min(\tau,\tau')-\tau\tau'-\mathbb{E}m_{0,\tau}(X_i)m_{0,\tau'}(X_i)}{f_0(q_0(\tau))f_0(q_0(\tau'))}.
$$

By the law of large numbers,

$$
\begin{aligned}
&\sum_{i=1}^{2n}\frac{1}{4n}\left(\frac{m_{1,\tau}(X_i)}{f_1(q_1(\tau))}-\frac{m_{0,\tau}(X_i)}{f_0(q_0(\tau))}\right)\left(\frac{m_{1,\tau'}(X_i)}{f_1(q_1(\tau'))}-\frac{m_{0,\tau'}(X_i)}{f_0(q_0(\tau'))}\right)\\
&\xrightarrow{p}\frac{1}{2}\mathbb{E}\left(\frac{m_{1,\tau}(X_i)}{f_1(q_1(\tau))}-\frac{m_{0,\tau}(X_i)}{f_0(q_0(\tau))}\right)\left(\frac{m_{1,\tau'}(X_i)}{f_1(q_1(\tau'))}-\frac{m_{0,\tau'}(X_i)}{f_0(q_0(\tau'))}\right).
\end{aligned}
$$

Last, by Lemma E.4, the last four terms on the RHS of (D.9) will vanish. Hence,

$$
\text{(D.9)}\xrightarrow{p}\Sigma(\tau,\tau'),
$$

where $\Sigma(\tau,\tau')$ is defined in Theorem 3.1. This concludes the proof.

E Technical Lemmas

E.1 A Maximal Inequality with i.n.i.d. Random Variables

Although Chernozhukov, Chetverikov, and Kato (2014) derived their Corollary 5.1 for i.i.d. data, the result is still valid when the data are independent but not identically distributed (i.n.i.d.). In this section, we restate their corollary for i.n.i.d. data and provide a brief justification. The proof is due to Chernozhukov et al. (2014). We include this section purely for clarification purpose. Let $\{W_i\}_{i=1}^n$ be a sequence of i.n.i.d. random variables taking values in a measurable space $(S, \mathcal{S})$ with distributions $\Pi_{i=1}^n \mathbb{P}^{(i)}$. Let $\mathcal{F}$ be a generic class of measurable functions $S \mapsto \Re$ with envelope F. Further denote $\overline{\mathbb{P}} f = \frac{1}{n}\sum_{i=1}^n \mathbb{P}^{(i)} f$, $||f||_{\overline{\mathbb{P}},2} = \sqrt{\overline{\mathbb{P}} f^2}$ and $\mathbb{P}_n f$ is the usual empirical process $\mathbb{P}_n f = \frac{1}{n}\sum_{i=1}^n f(W_i)$, $\sigma^2 = \sup_{f \in \mathcal{F}} \overline{\mathbb{P}} f^2 \leq \overline{\mathbb{P}} F^2$, and $M = \max_{i \in [n]} F(W_i)$.

Lemma E.1. *Suppose $\overline{\mathbb{P}} F^2 < \infty$ and there exist constants $a \geq e$ and $v \geq 1$ such that*

$$\sup_Q N(\mathcal{F}, e_Q, \varepsilon ||F||_{Q,2}) \leq \left(\frac{a}{\varepsilon}\right)^v, \quad \forall \varepsilon \in (0,1], \tag{E.1}$$

where $e_Q(f,g) = ||f - g||_{Q,2}$ and the supremum is taken over all finitely discrete probability measures on $(S, \mathcal{S})$. Then,

$$\mathbb{E}||\sqrt{n}(\mathbb{P}_n - \overline{\mathbb{P}})||_{\mathcal{F}} \lesssim \sqrt{v\sigma^2 \log\left(\frac{a||F||_{\overline{\mathbb{P}},2}}{\sigma}\right)} + \frac{v||M||_2}{\sqrt{n}} \log\left(\frac{a||F||_{\overline{\mathbb{P}},2}}{\sigma}\right).$$

The proof of Lemma E.1 is exactly the same as that for Chernozhukov et al. (2014, Corollary 5.1) with $\mathbb{P}$ replaced by $\overline{\mathbb{P}}$. For brevity, we just highlight some key steps below.

Proof. Let $\{\varepsilon_i\}_{i=1}^n$ be a sequence of Rademacher random variables that is independent of $\{W_i\}_{i=1}^n$, $\sigma_n^2 = \sup_{f \in \mathcal{F}} \mathbb{P}_n f^2$, and $Z = \mathbb{E}\left[\left|\left|\frac{1}{\sqrt{n}}\sum_{i=1}^n \varepsilon_i f(W_i)\right|\right|_{\mathcal{F}}\right]$. Then, by van der Vaart and Wellner (1996, Lemma 2.3.1) or Ledoux and Talagrand (2013, Lemma 6.3),

$$\mathbb{E}||\sqrt{n}(\mathbb{P}_n - \overline{\mathbb{P}})||_{\mathcal{F}} \leq 2Z.$$

Note Ledoux and Talagrand (2013, Lemma 6.3) only requires $\{W_i\}_{i=1}^n$ to be independent. In addition, let the uniform entropy integral be

$$J(\delta) \equiv J(\delta, \mathcal{F}, F) = \int_0^\delta \sup_Q \sqrt{1 + \log N(\mathcal{F}, e_Q, \varepsilon||F||_{Q,2})} d\varepsilon \tag{E.2}$$

where $e_Q(f,g) = ||f - g||_{Q,2}$ and the supremum is taken over all finitely discrete probability measures on $(S, \mathcal{S})$. Then, we have

$$\begin{aligned} Z =& \mathbb{E}\mathbb{E}\left[\left|\left|\frac{1}{\sqrt{n}}\sum_{i=1}^n \varepsilon_i f(W_i)\right|\right|_{\mathcal{F}} | W_1, \cdots, W_n\right] \\ \lesssim& \mathbb{E}\left[||F||_{\mathbb{P}_n,2} J(\sigma_n / ||F||_{\mathbb{P}_n,2})\right] \\ \lesssim& ||F||_{\overline{\mathbb{P}},2} J(\sqrt{\mathbb{E}\sigma_n^2} / ||F||_{\overline{\mathbb{P}},2}), \end{aligned} \tag{E.3}$$

where the second inequality is due to the Jensen's inequality and the fact that $J(\sqrt{x/y})\sqrt{y}$ is concave in (x, y) as shown by Chernozhukov et al. (2014). To see the first inequality, note that by the Hoeffding's inequality,

$$\mathbb{P}\left(\left|\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\varepsilon_i f(W_i)\right| \geq t \middle| \{W_i\}_{i=1}^{n}\right) \lesssim \exp\left(-\frac{t^2/2}{\frac{1}{n}\sum_{i=1}^{n} f(W_i)^2}\right),$$

which implies the stochastic process $\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\varepsilon_i f(W_i)$ indexed by f is sub-Gaussian conditionally on $\{W_i\}_{i=1}^{n}$. Then, the first inequality in (E.3) follows van der Vaart and Wellner (1996, Corollary 2.2.8), where we let $\delta = \sigma_n/||F||_{\mathbb{P}_n,2}$ and σ_n can be viewed as the diameter of the class of functions $\mathcal{F}$. We also note that this is a conditional argument, which is still valid even when $\{W_i\}_{i=1}^{n}$ is i.n.i.d.

Next, we aim to bound $\mathbb{E}\sigma_n^2$. Recall $\sigma^2 = \sup_{f\in\mathcal{F}} \overline{\mathbb{P}} f^2$. We have, for i.n.i.d. $\{W_i\}_{i=1}^{n}$,

$$\begin{aligned}
\mathbb{E}\sigma_n^2 &\leq \sigma^2 + \mathbb{E}(||(\mathbb{P}_n - \overline{\mathbb{P}})f^2||_{\mathcal{F}}) \\
&\leq \sigma^2 + 2\mathbb{E}\left[\left\|\frac{1}{n}\sum_{i=1}^{n}\varepsilon_i f^2(W_i)\right\|_{\mathcal{F}}\right] \\
&\leq \sigma^2 + 8\mathbb{E}\left[M\left\|\frac{1}{n}\sum_{i=1}^{n}\varepsilon_i f(W_i)\right\|_{\mathcal{F}}\right] \\
&\leq \sigma^2 + 8||M||_{\mathbb{P},2}\{\mathbb{E}[||\mathbb{P}_n\varepsilon_i f(W_i)||_{\mathcal{F}}^2]\}^{1/2} \\
&\leq \sigma^2 + C||M||_{\mathbb{P},2}\{\mathbb{E}[||\mathbb{P}_n\varepsilon_i f(W_i)||_{\mathcal{F}}] + n^{-1}||M||_{\mathbb{P},2}\} \\
&= \sigma^2 + Cn^{-1/2}||M||_{\mathbb{P},2}Z + Cn^{-1}||M||_{\mathbb{P},2}^2,
\end{aligned} \tag{E.4}$$

where the first inequality is due to the triangle inequality, the second inequality is due to Ledoux and Talagrand (2013, Lemma 6.3), the third inequality is due to Ledoux and Talagrand (2013, Theorem 4.12), the fourth inequality is due to the Cauchy-Schwarz inequality, and the fifth inequality is due to Ledoux and Talagrand (2013, Lemma 6.8) with $q = 2$.

Given (E.4), Chernozhukov et al. (2014) then proved the results that, for $\delta = \sigma/||F||_{\overline{\mathbb{P}},2}$,

$$\mathbb{E}[\sqrt{n}||\mathbb{P}_n - \overline{\mathbb{P}}||_{\mathcal{F}}] \lesssim J(\delta, \mathcal{F}, F)||F||_{\overline{\mathbb{P}},2} + \frac{||M||_{\mathbb{P}_s,2}J^2(\delta, \mathcal{F}, F)}{\delta^2\sqrt{n}}. \tag{E.5}$$

In this step, they relied on the facts that $J(\delta) = J(\delta, \mathcal{F}, F)$ is concave in δ and $\delta \mapsto J(\delta)/\delta$ is nonincreasing. The desired result is a quick corollary of (E.5) by noticing that, under (E.1),

$$J(\delta) \leq \int_0^\delta \sqrt{1 + \nu\log\left(\frac{a}{\varepsilon}\right)}d\varepsilon \leq 2\sqrt{2\nu}\delta\sqrt{\log\left(\frac{a}{\delta}\right)}. \tag{E.6}$$

□

E.2 Technical Lemmas Used in the Proof of Theorem 3.1

Lemma E.2. *Recall* $H_n(X_i,\tau)$ *defined in* (A.2). *Under the assumptions in Theorem 3.1,*

$$\sup_{\tau\in\Upsilon}\left|\sum_{i=1}^{2n}\left(A_i-\frac{1}{2}\right)H_n(X_i,\tau)\right|=o_p(1),$$

and

$$\sup_{\tau\in\Upsilon}\left|\sum_{i=1}^{2n}A_i\left[\xi_i^*\int_0^{\frac{u_0+u_1}{\sqrt{n}}}\left(1\{Y_i(1)-q_1(\tau)\leq v\}-1\{Y_i(1)-q_1(\tau)\leq 0\}\right)dv-H_n(X_i,\tau)\right]\right|=o_p(1),$$

where either $\xi_i^*=1$ *or* $\xi_i^*=\xi_i$ *which satisfies Assumption 3.*

Proof. For the first result, we have

$$\begin{aligned}
&\sup_{\tau\in\Upsilon}\left|\sum_{i=1}^{2n}\left(A_i-\frac{1}{2}\right)H_n(X_i,\tau)\right|\\
&\leq\frac{1}{2}\sum_{j=1}^{n}\sup_{\tau\in\Upsilon}|H_n(X_{\pi(2j-1),\tau})-H_n(X_{\pi(2j),\tau})|\\
&\leq\sum_{j=1}^{n}\frac{1}{2}\int_0^{\frac{|u_0+u_1|}{\sqrt{n}}}\sup_{\tau\in\Upsilon}|f_1(q_1(\tau)+\tilde{v}_j|X_{\pi(2j-1)})-f_1(q_1(\tau)+\tilde{v}_j|X_{\pi(2j)})|vdv\\
&\lesssim\sum_{j=1}^{n}\int_0^{\frac{|u_0+u_1|}{\sqrt{n}}}||X_{\pi(2j-1)}-X_{\pi(2j)}||_2vdv\\
&\lesssim\frac{(u_0+u_1)^2}{n}\sum_{j=1}^{n}||X_{\pi(2j-1)}-X_{\pi(2j)}||_2\xrightarrow{p}0,
\end{aligned}$$

where the first inequality is due to the fact that for the j-th pair, $(A_{\pi(2j-1)}-1/2, A_{\pi(2j)}-1/2)$ is either $(1/2,-1/2)$ or $(-1/2,1/2)$, the second inequality is by standard Taylor expansion to the first order where $|\tilde{v}_j|\leq(|u_0+u_1|)/\sqrt{n}$, the third inequality is due to Assumption 2, and the last convergence is due to Assumption 1.

Let $(\tilde{\xi}_j^*,\tilde{Y}_j(1),\tilde{X}_j)=(\xi_{i_j}^*,Y_{i_j}(1),X_{i_j})$ where i_j is the j-th smallest index in the set $\{i\in[2n]: A_i=1\}$. Then, similar to (B.8), we have

$$\begin{aligned}
&\sup_{\tau\in\Upsilon}\left|\sum_{i=1}^{2n}A_i\left[\xi_i^*\int_0^{\frac{u_0+u_1}{\sqrt{n}}}\left(1\{Y_i(1)-q_1(\tau)\leq v\}-1\{Y_i(1)-q_1(\tau)\leq 0\}\right)dv-H_n(X_i,\tau)\right]\right|\Bigg|\{A_i,X_i\}_{i=1}^{2n}\\
&\stackrel{d}{=}||\mathbb{P}_n-\overline{\mathbb{P}}||_{\mathcal{F}_4}|\{\tilde{X}_j\}_{j=1}^{n},
\end{aligned}$$

where $\mathcal{F}_4=\{\tilde{\xi}^*\int_0^{(u_0+u_1)/\sqrt{n}}\left(1\{\tilde{Y}(1)\leq q_1(\tau)+v\}-1\{\tilde{Y}(1)\leq q_1(\tau)\}\right)dv:\tau\in\Upsilon\}$, $\mathbb{P}_nf$ is the usual empirical process, $\overline{\mathbb{P}}f = \frac{1}{n}\sum_{j=1}^n \mathbb{P}^{(j)}f$, and $\mathbb{P}^{(j)}$ denotes the probability measure of $(\tilde{\xi}_j^*, \tilde{Y}_j(1))$ given $\tilde{X}_j$. Note $\mathcal{F}_4$ is a VC-class with a fixed VC index, has an envelope $F_j = (|u_0+u_1|\tilde{\xi}_j^*)/\sqrt{n}$, $M = \max_{j\in[n]} F_j = (|u_0+u_1|\log(n))/\sqrt{n}$, and

$$\begin{aligned}
\sigma^2 = \sup_{f\in\mathcal{F}_4}\overline{\mathbb{P}}f^2 \leq & \sup_{\tau\in\Upsilon}\frac{1}{n}\sum_{j=1}^n\left[F_1\left(q_1(\tau)+\frac{|u_0+u_1|}{\sqrt{n}}\middle|\tilde{X}_j\right)-F_1\left(q_1(\tau)-\frac{|u_0+u_1|}{\sqrt{n}}\middle|\tilde{X}_j\right)\right]\frac{u^2}{n} \\
\leq & \frac{1}{n}\sum_{j=1}^n C(\tilde{X}_j)\frac{(u_0+u_1)^2}{n^{3/2}} \\
= & \frac{1}{n}\sum_{i=1}^{2n} A_i C(X_i)\frac{(u_0+u_1)^2}{n^{3/2}} \\
\leq & \left(\frac{1}{n}\sum_{i=1}^{2n} C(X_i)\right)\frac{(u_0+u_1)^2}{n^{3/2}}.
\end{aligned}$$

As $\left(\frac{1}{n}\sum_{i=1}^{2n} C(X_i)\right) \stackrel{a.s.}{\to} \mathbb{E}2C(X_i)$, we have $\left(\frac{1}{n}\sum_{i=1}^{2n} C(X_i)\right) \leq 3\mathbb{E}C(X_i)$ a.s. Given such a sequence $\{X_i\}_{i\geq 1}$, Lemma E.1 implies

$$\mathbb{E}\left[||\mathbb{P}_n - \overline{\mathbb{P}}||_{\mathcal{F}_4}|\{\tilde{X}_j\}_{i=1}^n\right] \lesssim \sqrt{\frac{3\mathbb{E}C(X_i)\log(n)}{n^{3/2}}} + \frac{\log^2(n)}{n} = o_{a.s.}(1).$$

This implies

$$\sup_{\tau\in\Upsilon}\left|\sum_{i=1}^{2n} A_i\left[\xi_i^*\int_0^{\frac{u_0+u_1}{\sqrt{n}}}\left(1\{Y_i(1)-q_1(\tau)\leq v\} - 1\{Y_i(1)-q_1(\tau)\leq 0\}\right)dv - H_n(X_i,\tau)\right]\right| = o_p(1).$$

□

Lemma E.3. *Under the assumptions in Theorem 3.1,*

$$\sup_{\tau\in\Upsilon}\left|\sum_{i=1}^{2n}\frac{(A_i-1/2)}{\sqrt{n}}m_{1,\tau}(X_i)\right| = o_p(1).$$

Proof. We have

$$\sup_{\tau\in\Upsilon}\left|\sum_{i=1}^{2n}\frac{(A_i-1/2)}{\sqrt{n}}m_{1,\tau}(X_i)\right| = \sup_{\tau\in\Upsilon}\left|\sum_{j=1}^{n}\frac{1}{2\sqrt{n}}(A_{\pi(2j-1)}-A_{\pi(2j)})(F_1(q_1(\tau)|X_{\pi(2j-1)}) - F_1(q_1(\tau)|X_{\pi(2j)}))\right|.$$

Note that

$$\mathcal{F}_5 = \{F_1(q_1(\tau)|X) - F_1(q_1(\tau)|X') : \tau\in\Upsilon\}$$

is a VC-class with a fixed VC-index and has an envelope $F = 2$. This implies (E.1) holds with some constants $a \geq e$ and $v \geq 1$. Then, as discussed in the (E.6), the uniform entropy integral $J(\delta)$ of $\mathcal{F}_5$ satisfies

$$J(\delta) \leq \int_0^\delta \sqrt{1 + \nu \log\left(\frac{a}{\varepsilon}\right)} d\varepsilon \leq 2\sqrt{2\nu}\delta\sqrt{\log\left(\frac{a}{\delta}\right)}.$$

In addition,

$$\sigma_n^2 = \sup_{\tau \in \Upsilon} \frac{1}{n} \sum_{j=1}^n (F_1(q_1(\tau)|X_{\pi(2j-1)}) - F_1(q_1(\tau)|X_{\pi(2j)}))^2 \lesssim \frac{1}{n} \sum_{j=1}^n \left\| X_{\pi(2j-1)} - X_{\pi(2j)} \right\|_2^2 \xrightarrow{p} 0.$$

We focus on the set $\mathcal{A}_n = \{\sigma_n^2 \leq \varepsilon\}$ for some arbitrary $\varepsilon > 0$ so that $\mathbb{P}(\mathcal{A}_n) \geq 1 - \varepsilon$ for n sufficiently large. Note that $\mathcal{A}_n$ belongs to the sigma field generated by $\{X_i\}_{i=1}^{2n}$. In addition, note that conditional on $\{X_i\}_{i=1}^{2n}$, $\{A_{\pi(2j-1)} - A_{\pi(2j)}\}_{j=1}^n$ is a sequence of i.i.d. Rademacher random variables. Then, following the same argument in (E.3)

$$\begin{aligned}
&\mathbb{E} \sup_{\tau \in \Upsilon} \left| \sum_{i=1}^{2n} \frac{(A_i - 1/2)}{\sqrt{n}} m_{1,\tau}(X_i) \right| 1\{\mathcal{A}_n\} \\
&= \mathbb{E}\left\{ \mathbb{E}\left[\left\| \frac{1}{2\sqrt{n}} \sum_{j=1}^n (A_{\pi(2j-1)} - A_{\pi(2j)}) f(X_{\pi(2j-1)}, X_{\pi(2j)}) \right\|_{\mathcal{F}_5} \middle| \{X_i\}_{i=1}^{2n} \right] 1\{\mathcal{A}_n\} \right\} \\
&\lesssim \mathbb{E} J(\sigma_n/2) 1\{\mathcal{A}_n\} \\
&\lesssim J(\varepsilon/2) \lesssim \sqrt{2\nu}\varepsilon \sqrt{\log\left(\frac{2a}{\varepsilon}\right)},
\end{aligned}$$

where the first inequality is due to van der Vaart and Wellner (1996, Corollary 2.2.8) and the fact that, by the Hoeffding's inequality, for any $f \in \mathcal{F}_5$,

$$\mathbb{P}\left(\left| \sum_{j=1}^n (A_{\pi(2j-1)} - A_{\pi(2j)}) f(X_{\pi(2j-1)}, X_{\pi(2j)}) \right| \geq x \middle| \{X_i\}_{i=1}^{2n} \right) \leq 2 \exp\left(-\frac{1}{2} \frac{x^2}{\sum_{j=1}^n f^2(X_{\pi(2j-1)}, X_{\pi(2j)})} \right).$$

As $\sqrt{2\nu}\varepsilon\sqrt{\log\left(\frac{2a}{\varepsilon}\right)} \to 0$ as $\varepsilon \to 0$, we derive the desired result by letting $n \to \infty$ followed by $\varepsilon \to 0$.

□

E.3 Technical Lemmas Used in the Proof of Theorem 4.1

Lemma E.4. *Suppose the assumptions in Theorem 4.1 hold, then*

$$\frac{1}{n}\sum_{i=1}^{2n} A_i \eta_{1,i}(\tau) m_{j,\tau'}(X_i) \xrightarrow{p} 0,$$

$$\frac{1}{n}\sum_{i=1}^{2n} A_i \eta_{1,i}(\tau) m_{0,\tau'}(X_i) \xrightarrow{p} 0,$$

$$\frac{1}{n}\sum_{i=1}^{2n} (1-A_i) \eta_{0,i}(\tau) m_{0,\tau'}(X_i) \xrightarrow{p} 0,$$

and

$$\frac{1}{n}\sum_{i=1}^{2n} (1-A_i) \eta_{0,i}(\tau) m_{1,\tau'}(X_i) \xrightarrow{p} 0.$$

Proof. We focus on the first statement. The rest can be proved in the same manner. Based on the notation in Section 4.2, we have

$$\frac{1}{n}\sum_{i=1}^{2n} A_i \eta_{1,i}(\tau) m_{1,\tau'}(X_i) = \frac{1}{n}\sum_{j=1}^{n} \eta_{1,(j,1)}(q_1(\tau),\tau) m_{1,\tau'}(X_{(j,1)}, q_1(\tau')).$$

where $\eta_{1,i}(q,\tau) = (\tau - 1\{Y_i(1) \le q\}) - m_{1,\tau}(X_i, q)$. Then, (E.7) implies the desired result.

□

E.4 Technical Lemmas Used in the Proof of Theorem 4.2

Lemma E.5. *Recall $II(\tau,\tau')$ and $III(\tau,\tau')$ defined in* (C.1). *Suppose the assumptions in Theorem 3.1 hold, then*

$$\sup_{\tau,\tau'\in\Upsilon} |II(\tau,\tau')| \xrightarrow{p} 0 \quad \textit{and} \quad \sup_{\tau,\tau'\in\Upsilon} |III(\tau,\tau')| \xrightarrow{p} 0.$$

Proof. We focus on bounding $II(\tau,\tau')$. The bound for $III(\tau,\tau')$ can be established similarly. By (C.3), we have, with probability greater than $1-\varepsilon$,

$$|II(\tau,\tau')| \le \sup_{\tau,\tau'\in\Upsilon, |v|,|v'|\le L/\sqrt{n}} \left| \frac{1}{n}\sum_{j=1}^{n} \eta_{1,(j,1)}(q_1(\tau)+v,\tau) m_{1,\tau'}(X_{(j,1)}, q_1(\tau')+v') \right|. \tag{E.7}$$

We aim to bound the RHS. Let $\{\varepsilon_j\}_{j=1}^n$ denote a sequence of i.i.d. Rademacher random variables that is independent of the data. Further denote the class of functions

$$\mathcal{F}_6 = \{\eta_{1,(j,1)}(q_1(\tau)+v,\tau)m_{1,\tau'}(X_{(j,1)},q_1(\tau')+v') : \tau,\tau' \in \Upsilon, |v|,|v'| \leq L/\sqrt{n}\}.$$

Note $\mathcal{F}_6$ has an envelope $F = 1$ and is nested by a VC-class of functions with a fixed VC-index. Then,

$$\begin{aligned}
&\mathbb{E}\left[\sup_{\tau,\tau'\in\Upsilon,|v|,|v'|\leq L/\sqrt{n}}\left|\frac{1}{n}\sum_{j=1}^n \eta_{1,(j,1)}(q_1(\tau)+v,\tau)m_{1,\tau'}(X_{(j,1)},q_1(\tau')+v')\right|\right] \\
&= \mathbb{E}\left\{\mathbb{E}\left[\sup_{\tau,\tau'\in\Upsilon,|v|,|v'|\leq L/\sqrt{n}}\left|\frac{1}{n}\sum_{j=1}^n \eta_{1,(j,1)}(q_1(\tau)+v,\tau)m_{1,\tau'}(X_{(j,1)},q_1(\tau')+v')\right| |\{X_i,A_i\}_{i=1}^{2n}\right]\right\} \\
&\lesssim \mathbb{E}\left\{\mathbb{E}\left[\sup_{\tau,\tau'\in\Upsilon,|v|,|v'|\leq L/\sqrt{n}}\left|\frac{1}{n}\sum_{j=1}^n \varepsilon_j\eta_{1,(j,1)}(q_1(\tau)+v,\tau)m_{1,\tau'}(X_{(j,1)},q_1(\tau')+v')\right| |\{X_i,A_i\}_{i=1}^{2n}\right]\right\} \\
&= \mathbb{E}\left\{\mathbb{E}\left[\sup_{\tau,\tau'\in\Upsilon,|v|,|v'|\leq L/\sqrt{n}}\left|\frac{1}{n}\sum_{j=1}^n \varepsilon_j\eta_{1,(j,1)}(q_1(\tau)+v,\tau)m_{1,\tau'}(X_{(j,1)},q_1(\tau')+v')\right| |\{X_i,A_i,Y_i(1)\}_{i=1}^{2n}\right]\right\} \\
&\leq \frac{||F||_{\mathbb{P},2}J(\sqrt{\mathbb{E}\sigma_n^2}/||F||_{\mathbb{P},2})}{\sqrt{n}} \lesssim \frac{1}{\sqrt{n}},
\end{aligned} \tag{E.8}$$

where the first equality is due to the law of iterated expectation, the first inequality is due to Ledoux and Talagrand (2013, Lemma 6.3) and the fact that $\{\eta_{1,(j,1)}(q_1(\tau)+v,\tau)\}_{j=1}^n$ is a sequence of independent and centered random variables given $\{X_i,A_i\}_{i=1}^{2n}$, the second inequality follows the same argument in (E.3) with $F = 2$,

$$\sigma_n^2 = \sup_{\tau,\tau'\in\Upsilon,|v|,|v'|\leq L/\sqrt{n}} \frac{1}{n}\sum_{j=1}^n \left[\eta_{1,(j,1)}(q_1(\tau)+v,\tau)m_{1,\tau'}(X_{(j,1)},q_1(\tau')+v')\right]^2 \leq 4,$$

and $J(\cdot)$ being the uniform entropy integral for the class of functions $\mathcal{F}_6$ defined in (E.2), and the last inequality holds because when $\mathcal{F}_6$ is nested by a VC-class, ε_i is bounded, and thus, has a sub-Gaussian tail, and $\delta = \sqrt{\mathbb{E}\sigma_n^2}/||F||_{\mathbb{P},2} \leq 1$, we have

$$J(\delta) \lesssim \delta \max(\sqrt{\log(1/\delta)}, 1) \lesssim 1,$$

as shown in (E.6). This implies, uniformly over $\tau,\tau' \in \Upsilon$,

$$II(\tau,\tau') \xrightarrow{p} 0.$$

□

Lemma E.6. *Recall $R_{IV}(\tau,\tau')$ defined in* (C.7)*. Suppose assumptions in Theorem 3.1 hold, then*

$$\sup_{\tau,\tau'\in\Upsilon}|R_{IV}(\tau,\tau')| = o_p(1) \quad \textit{and} \quad \sup_{\tau,\tau'\in\Upsilon}\left|\frac{1}{n}\sum_{i=1}^{2n}\left(A_i-\frac{1}{2}\right)m_{1,\tau}(X_i)m_{1,\tau'}(X_i)\right| = o_p(1).$$

Proof. Note

$$R_{IV}(\tau,\tau') = \frac{1}{n}\sum_{j=1}^{n}\left[m_{1,\tau}(X_{(j,1)})m_{1,\tau'}(X_{(j,1)}) - m_{1,\tau}(X_{(j,1)},\hat{q}_1(\tau))m_{1,\tau'}(X_{(j,1)},\hat{q}_1(\tau'))\right].$$

By (C.3) and the fact that $F_1(\cdot|X)$ is Lipschitz continuous, we have

$$\begin{aligned}&\sup_{\tau,\tau'\in\Upsilon}|R_{IV}(\tau,\tau')|\\ \leq &\sup_{\tau,\tau'\in\Upsilon}\frac{1}{n}\sum_{j=1}^{n}\left|m_{1,\tau}(X_{(j,1)})m_{1,\tau'}(X_{(j,1)}) - m_{1,\tau}(X_{(j,1)},\hat{q}_1(\tau))m_{1,\tau'}(X_{(j,1)},\hat{q}_1(\tau'))\right| \stackrel{p}{\longrightarrow} 0.\end{aligned}$$

By the same argument as in the proof of Lemma E.3, we have

$$\sup_{\tau,\tau'\in\Upsilon}\left|\frac{1}{n}\sum_{i=1}^{2n}\left(A_i-\frac{1}{2}\right)m_{1,\tau}(X_i)m_{1,\tau'}(X_i)\right| \stackrel{p}{\longrightarrow} 0.$$

□

Lemma E.7. *Recall $S^*_{n,1}(\tau)$ defined in* (4.5)*. Suppose assumptions in Theorem 3.1 hold. Then, $\{S^*_{n,1}(\tau):\tau\in\Upsilon\}$ is stochastically equicontinuous and tight.*

Proof. It suffices to show the two marginals of $S^*_{n,1}(\tau)$ are stochastically equicontinuous and tight. We focus on the first marginal

$$\left\{\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\eta_j(\tau - 1\{Y_{(j,1)}\leq\hat{q}_1(\tau)\}):\tau\in\Upsilon\right\}.$$

By (C.3), it suffices to establish the stochastic equicontinuity and tightness of

$$\left\{\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\eta_j(\tau - 1\{Y_{(j,1)}\leq q_1(\tau)+v/\sqrt{n}\}):\tau\in\Upsilon, |v|\leq L\right\}$$

for any fixed L. Let

$$\mathcal{F}_7 = \left\{ \begin{array}{c} (\tau - 1\{Y_{(j,1)} \leq q_1(\tau) + v/\sqrt{n}\}) - (\tau' - 1\{Y_{(j,1)} \leq q_1(\tau') + v'/\sqrt{n}\}) : \\ \tau, \tau' \in \Upsilon, |v|, |v'| \leq L, |\tau - \tau'| \leq \varepsilon, |v - v'| \leq \varepsilon \end{array} \right\},$$

which is nested by a VC-class with envelope 2. Then, by (E.2) and (E.6), the uniform entropy integral $J(\delta)$ of $\mathcal{F}_7$ satisfies

$$J(\delta) \lesssim \delta \max(1, \sqrt{\log(1/\delta)}).$$

By the calculation of $\tilde{\Sigma}^*_{1,1,1}(\tau, \tau')$ (with $\hat{q}_1(\tau)$ replaced by $q_1(\tau) + \frac{v}{\sqrt{n}}$) in Section C, we have, uniformly over $\tau, \tau' \in \Upsilon$, $v, v' \in [-L, L]$,

$$\begin{aligned} \sigma_n^2(\tau, \tau', v, v') =& \frac{1}{n}\sum_{j=1}^n \left[(\tau - 1\{Y_{(j,1)} \leq q_1(\tau) + v/\sqrt{n}\}) - (\tau' - 1\{Y_{(j,1)} \leq q_1(\tau') + v'/\sqrt{n}\})\right]^2 \\ \xrightarrow{p}& \tau(1-\tau) + \tau'(1-\tau') - 2(\min(\tau, \tau') - \tau\tau') = |\tau - \tau'| - (\tau - \tau')^2. \end{aligned} \tag{E.9}$$

Let $\mathcal{A}_n(\varepsilon) = 1\{\sup_{\tau,\tau' \in \Upsilon, v, v' \in [-L,L]} |\sigma_n^2(\tau, \tau', v, v') - \left(|\tau - \tau'| - (\tau - \tau')^2\right)| \leq \varepsilon\}$, which will occur with probability approaching one. Also by construction, conditionally on data, $\frac{1}{\sqrt{n}}\sum_{j=1}^n \eta_j(\tau - 1\{Y_{(j,1)} \leq q_1(\tau) + v/\sqrt{n}\})$ is a sub-Gaussian process. Then,

$$\begin{aligned} &\mathbb{E}\left[\sup \frac{1}{\sqrt{n}}\sum_{j=1}^n \eta_j(\tau - 1\{Y_{(j,1)} \leq q_1(\tau) + v/\sqrt{n}\}) - (\tau' - 1\{Y_{(j,1)} \leq q_1(\tau') + v'/\sqrt{n}\})\middle| Data\right] 1\{\mathcal{A}_n(\varepsilon)\} \\ &\lesssim J(\frac{\sup \sigma_n(\tau, \tau', v, v')}{2}) 1\{\mathcal{A}_n(\varepsilon)\} \\ &\lesssim J(\sqrt{\varepsilon}) \lesssim \sqrt{\varepsilon}\max(1, \sqrt{\log(1/\varepsilon)}), \end{aligned}$$

where the supremum is taken over $\tau, \tau' \in \Upsilon, |v|, |v'| \leq L, |\tau - \tau'| \leq \varepsilon, |v - v'| \leq \varepsilon$, the first inequality is due to (van der Vaart and Wellner, 1996, Corollary 2.2.8), and the second inequality is due to (E.9) and the definition of $\mathcal{A}_n$. Then, for any $t > 0$

$$\begin{aligned} &\mathbb{P}\left(\sup \frac{1}{\sqrt{n}}\sum_{j=1}^n \eta_j \left[(\tau - 1\{Y_{(j,1)} \leq q_1(\tau) + v/\sqrt{n}\}) - (\tau' - 1\{Y \leq q_1(\tau') + v'/\sqrt{n}\})\right] \geq t\right) \\ &\leq \mathbb{P}(\mathcal{A}_n^c(\varepsilon)) + \mathbb{P}\left(\sup \frac{1}{\sqrt{n}}\sum_{j=1}^n \eta_j \left[(\tau - 1\{Y \leq q_1(\tau) + v/\sqrt{n}\}) - (\tau' - 1\{Y \leq q_1(\tau') + v'/\sqrt{n}\})\right] \geq t, \mathcal{A}_n(\varepsilon)\right) \\ &\leq \mathbb{E}\left\{\frac{\mathbb{E}\left[\sup \frac{1}{\sqrt{n}}\sum_{j=1}^n \eta_j(\tau - 1\{Y_{(j,1)} \leq q_1(\tau) + v/\sqrt{n}\}) - (\tau' - 1\{Y_{(j,1)} \leq q_1(\tau') + v'/\sqrt{n}\})\middle| Data\right] 1\{\mathcal{A}_n(\varepsilon)\}}{t}\right\} \\ &+ \mathbb{P}(\mathcal{A}_n^c(\varepsilon)) \end{aligned}$$

$$\lesssim \mathbb{P}(\mathcal{A}_n^c(\varepsilon)) + \frac{\sqrt{\varepsilon}\max(1, \sqrt{\log(1/\varepsilon)})}{t},$$

where the supremum is taken over $\tau, \tau' \in \Upsilon, |v|, |v'| \le L, |\tau - \tau'| \le \varepsilon, |v - v'| \le \varepsilon$. Let $n \to \infty$ followed by $\varepsilon \to 0$, we have

$$\lim_{\varepsilon \to 0} \limsup_n \mathbb{P}\left(\sup \frac{1}{\sqrt{n}} \sum_{j=1}^n \eta_j \left[(\tau - 1\{Y_{(j,1)} \le q_1(\tau) + v/\sqrt{n}\}) - (\tau' - 1\{Y_{(j,1)} \le q_1(\tau') + v'/\sqrt{n}\}) \right] \ge t \right) = 0,$$

which implies $\left\{ \frac{1}{\sqrt{n}} \sum_{j=1}^n \eta_j (\tau - 1\{Y_{(j,1)} \le \hat{q}_1(\tau)\}) : \tau \in \Upsilon \right\}$ is stochastically equicontinuous. In addition, for any fixed τ,

$$\frac{1}{\sqrt{n}} \sum_{j=1}^n \eta_j (\tau - 1\{Y_{(j,1)} \le \hat{q}_1(\tau)\}) = O_p(1).$$

This implies it is also tight over $\tau \in \Upsilon$. □

E.5 Technical Lemmas Used in the Proof of Theorem 4.3

Lemma E.8. *Suppose the assumptions in Theorem 4.3 hold, then*

$$\max_{i \in [2n]} |\hat{A}_i - 1/2| = o_p(1)$$

and

$$\frac{1}{n} \sum_{i=1}^{2n} \xi_i (\hat{A}_i - 1/2)^2 = o_p(n^{-1/2}).$$

Proof. Let $\theta_0 = (0.5, 0, \cdots, 0)^T$ be a $K \times 1$ vector. Then,

$$\begin{aligned} ||\hat{\theta} - \theta_0||_2 &= \left\| \left[\frac{1}{n} \sum_{i=1}^{2n} \xi_i b(X_i) b(X_i)^T \right]^{-1} \left[\frac{1}{n} \sum_{i=1}^{2n} \xi_i b(X_i)(A_i - \frac{1}{2}) \right] \right\|_2 \\ &\lesssim \left\| \frac{1}{n} \sum_{i=1}^{2n} \xi_i b(X_i)(A_i - \frac{1}{2}) \right\|_2 \\ &\lesssim \sqrt{K} \left\| \frac{1}{n} \sum_{i=1}^{2n} \xi_i b(X_i)(A_i - \frac{1}{2}) \right\|_\infty . \end{aligned}$$

Next, we aim to bound $\left\| \frac{1}{n} \sum_{i=1}^{2n} \xi_i b(X_i)(A_i - \frac{1}{2}) \right\|_\infty$. Let $b_k(X)$ be the kth component of $b(X)$.

Then,

$$\begin{aligned}
&\max_{k\in[K]}\frac{1}{n}\sum_{j=1}^{n}(\xi_{\pi(2j-1)}b_k(X_{\pi(2j-1)})-\xi_{\pi(2j)}b_k(X_{\pi(2j)}))^2\\
&\lesssim \max_{k\in[K]}\frac{1}{n}\sum_{i=1}^{2n}\xi_i^2b_k^2(X_i)\\
&\lesssim \max_{k\in[K]}\mathbb{E}\xi_i^2b_k^2(X_i)+\max_{k\in[K]}\left|\frac{1}{n}\sum_{i=1}^{2n}\left[\xi_i^2b_k^2(X_i)-\mathbb{E}\xi_i^2b_k^2(X_i)\right]\right|.
\end{aligned}$$

The first term on the RHS of the above display is bounded by $\overline{C}$ based on Assumption 5. Let $\{\varepsilon_i\}_{i\in[2n]}$ be a sequence of i.i.d. Rademacher random variables. Then,

$$\mathbb{E}\max_{k\in[K]}\left|\frac{1}{n}\sum_{i=1}^{2n}\left[\xi_i^2b_k^2(X_i)-\mathbb{E}\xi_i^2b_k^2(X_i)\right]\right|\leq 2\mathbb{E}\max_{k\in[K]}\left|\frac{1}{n}\sum_{i=1}^{2n}\varepsilon_i\left[\xi_i^2b_k^2(X_i)-\mathbb{E}\xi_i^2b_k^2(X_i)\right]\right|.$$

By Hoeffding's inequality,

$$\mathbb{P}\left(\left|\frac{1}{\sqrt{2n}}\sum_{i=1}^{2n}\varepsilon_i\left[\xi_i^2b_k^2(X_i)-\mathbb{E}\xi_i^2b_k^2(X_i)\right]\right|\geq t\middle|\{\xi_i,X_i\}_{i\in[2n]}\right)\leq 2\exp(-\frac{t^2}{2\sigma_k^2}),$$

where $\sigma_k^2=\frac{1}{2n}\sum_{i=1}^{2n}\left[\xi_i^2b_k^2(X_i)-\mathbb{E}\xi_i^2b_k^2(X_i)\right]^2$. Then, by van der Vaart and Wellner (1996, Lemmas 2.2.1 and 2.2.2),

$$\mathbb{E}\left[\max_{k\in[K]}\left|\frac{1}{n}\sum_{i=1}^{2n}\varepsilon_i\xi_i^2b_k^2(X_i)\right|\middle|\{\xi_i,X_i\}_{i\in[2n]}\right]\lesssim\sqrt{\frac{\log(K)}{n}}\sqrt{\max_{k\in[K]}\sigma_k^2}.$$

Applying expectation on both sides and noticing that the square root function is concave, we have

$$\begin{aligned}
\mathbb{E}\max_{k\in[K]}\left|\frac{1}{n}\sum_{i=1}^{2n}\varepsilon_i\xi_i^2b_k^2(X_i)\right| &\lesssim\sqrt{\frac{\log(K)}{n}}\sqrt{\mathbb{E}\max_{k\in[K]}\sigma_k^2}\\
&\lesssim\sqrt{\frac{\log(K)}{n}}\sqrt{\sum_{k\in[K]}\mathbb{E}\sigma_k^2}\\
&\lesssim\sqrt{\frac{\log(K)}{n}}\zeta(K)\sqrt{K}=o(1).
\end{aligned}$$

Therefore,

$$\max_{k\in[K]}\left|\frac{1}{n}\sum_{i=1}^{2n}\left[\xi_i^2b_k^2(X_i)-\mathbb{E}\xi_i^2b_k^2(X_i)\right]\right|=o_p(1)$$

and with probability approaching one,

$$\max_{k\in[K]}\frac{1}{n}\sum_{j=1}^{n}(\xi_{\pi(2j-1)}b_k(X_{\pi(2j-1)})-\xi_{\pi(2j)}b_k(X_{\pi(2j)}))^2\leq 2\overline{C}.$$

Let $I_n'=\{\max_{k\in[K]}\frac{1}{n}\sum_{j=1}^{n}(\xi_{\pi(2j-1)}b_k(X_{\pi(2j-1)})-\xi_{\pi(2j)}b_k(X_{\pi(2j)}))^2\leq 2\overline{C}\}$. For $t=\sqrt{\log(n)\overline{C}}$, we have

$$\begin{aligned}
&\mathbb{P}\left(\left\|\frac{1}{n}\sum_{i=1}^{2n}\xi_i b(X_i)(A_i-\frac{1}{2})\right\|_\infty\geq t/\sqrt{n},I_n'\right)\\
=&\mathbb{E}\mathbb{P}\left(\left\|\frac{1}{n}\sum_{i=1}^{2n}\xi_i b(X_i)(A_i-\frac{1}{2})\right\|_\infty\geq t/\sqrt{n}\middle|\{X_i,\xi_i\}_{i\in[2n]}\right)1\{I_n'\}\\
=&\mathbb{E}\mathbb{P}\left(\left\|\sum_{j=1}^{n}(A_{\pi(2j-1)}-A_{\pi(2j)})(\xi_{\pi(2j-1)}b(X_{\pi(2j-1)})-\xi_{\pi(2j)}b(X_{\pi(2j)}))\right\|_\infty\geq 2t\sqrt{n}\middle|\{X_i,\xi_i\}_{i\in[2n]}\right)1\{I_n'\}\\
\leq&\sum_{k=1}^{K}\mathbb{E}\mathbb{P}\left(\left|\sum_{j=1}^{n}(A_{\pi(2j-1)}-A_{\pi(2j)})(\xi_{\pi(2j-1)}b_k(X_{\pi(2j-1)})-\xi_{\pi(2j)}b_k(X_{\pi(2j)}))\right|\geq 2t\sqrt{n}\middle|\{X_i,\xi_i\}_{i\in[2n]}\right)1\{I_n'\}\\
\leq&\sum_{k=1}^{K}2\mathbb{E}\exp\left(\frac{-2t^2n}{\sum_{j=1}^{n}(\xi_{\pi(2j-1)}b_k(X_{\pi(2j-1)})-\xi_{\pi(2j)}b_k(X_{\pi(2j)}))^2}\right)1\{I_n'\}\\
\leq&2\exp\left(\log(K)-\frac{t^2}{\overline{\overline{C}}}\right)\to 0,
\end{aligned}$$

where the second last inequality is due to the Hoeffding's inequality and the fact that given $\{X_i,\xi_i\}_{i\in[2n]}$, $\{A_{\pi(2j-1)}-A_{\pi(2j)}\}_{j\in[n]}$ is i.i.d. sequence of Rademacher random variables.

This implies,

$$||\hat{\theta}-\theta_0||_2=O_p(\sqrt{\frac{K\log(n)}{n}}),$$

and thus

$$\max_{i\in[2n]}|\hat{A}_i-1/2|=\max_i|b(X_i)'(\hat{\theta}-\theta_0)|=O_p\left(\zeta(K)\sqrt{\frac{K\log(n)}{n}}\right)=o_p(1).$$

For the second result, we have

$$\frac{1}{n}\sum_{i=1}^{2n}\xi_i(\hat{A}_i-1/2)^2\leq\lambda_{\max}\left(\frac{1}{n}\sum_{i=1}^{2n}\xi_i b(X_i)b(X_i)'\right)||\hat{\theta}-\theta_0||_2^2=O_p\left(\frac{K\log(n)}{n}\right)=o_p(n^{-1/2}),$$

as $K^2\log^2(n)=o(n)$.

□

Lemma E.9. *Suppose assumptions in Theorem 4.3 hold, then*

$$\sup_{\tau\in\Upsilon}\left|\sum_{i=1}^{2n}\frac{\xi_i m_{1,\tau}(X_i)(\hat{A}_i-1/2)}{\sqrt{n}}-\sum_{i=1}^{2n}\frac{\xi_i m_{1,\tau}(X_i)(A_i-1/2)}{\sqrt{n}}\right|=o_p(1),$$

$$\sup_{\tau\in\Upsilon}\left|\sum_{i=1}^{2n}\frac{2\xi_i(A_i-1/2)m_{1,\tau}(X_i)(\hat{A}_i-1/2)}{\sqrt{n}}\right|=o_p(1),$$

and

$$\sup_{\tau\in\Upsilon}\left|\sum_{i=1}^{2n}\frac{2\xi_i A_i\eta_{1,i}(\tau)(\hat{A}_i-1/2)}{\sqrt{n}}\right|=o_p(1).$$

Proof. For the first result, note $m_{1,\tau}(X_i)=b(X_i)'\gamma_1(\tau)+B_\tau(X_i)$ such that $\sup_{x\in\text{Supp}(X),\tau\in\Upsilon}|B_\tau(x)|=o(1/\sqrt{n})$. Then,

$$\begin{aligned}&\sum_{i=1}^{2n}\frac{\xi_i m_{1,\tau}(X_i)(\hat{A}_i-1/2)}{\sqrt{n}}\\&=\sum_{i=1}^{2n}\frac{\xi_i m_{1,\tau}(X_i)b(X_i)'(\hat{\theta}-\theta_0)}{\sqrt{n}}\\&=\gamma_1'(\tau)\left[\sum_{i=1}^{2n}\frac{\xi_i b(X_i)b(X_i)'}{\sqrt{n}}\right](\hat{\theta}-\theta_0)+\sum_{i=1}^{2n}\frac{\xi_i B_\tau(X_i)b(X_i)'(\hat{\theta}-\theta_0)}{\sqrt{n}}\\&=\sum_{i=1}^{2n}\frac{\xi_i\gamma_1(\tau)'b(X_i)(A_i-1/2)}{\sqrt{n}}+\sum_{i=1}^{2n}\frac{\xi_i B_\tau(X_i)b(X_i)'(\hat{\theta}-\theta_0)}{\sqrt{n}}\\&=\sum_{i=1}^{2n}\frac{\xi_i m_{1,\tau}(X_i)(A_i-1/2)}{\sqrt{n}}-\sum_{i=1}^{2n}\frac{\xi_i B_\tau(X_i)(A_i-1/2)}{\sqrt{n}}+\sum_{i=1}^{2n}\frac{\xi_i B_\tau(X_i)b(X_i)'(\hat{\theta}-\theta_0)}{\sqrt{n}},\end{aligned}$$

where the third equality holds because

$$\hat{\theta}-\theta_0=\left[\sum_{i=1}^{2n}\frac{\xi_i b(X_i)b(X_i)'}{n}\right]^{-1}\left[\sum_{i=1}^{2n}\frac{\xi_i b(X_i)(A_i-1/2)}{n}\right].$$

Furthermore,

$$\sup_{\tau\in\Upsilon}\left|\sum_{i=1}^{2n}\frac{\xi_i B_\tau(X_i)(A_i-1/2)}{\sqrt{n}}\right|\leq o_p(1)\left(\frac{1}{2n}\sum_{i=1}^{2n}\xi_i\right)=o_p(1)$$

and

$$\sup_{\tau\in\Upsilon}\left|\sum_{i=1}^{2n}\frac{\xi_i B_\tau(X_i)b(X_i)'(\hat{\theta}-\theta_0)}{\sqrt{n}}\right| \leq \sum_{i=1}^{2n}\frac{\xi_i\zeta(K)||\hat{\theta}-\theta_0||_2}{\sqrt{n}}o_p(1/\sqrt{n})$$
$$= \left(\sum_{i=1}^{2n}\frac{\xi_i}{n}\right)o_p\left(\sqrt{\frac{K\zeta^2(K)\log(n)}{n}}\right) = o_p(1).$$

This leads to the first result.

For the second result, we have

$$\left|\sum_{i=1}^{2n}\frac{2\xi_i(A_i-1/2)m_{1,\tau}(X_i)(\hat{A}_i-1/2)}{\sqrt{n}}\right| \leq \left\|\sum_{i=1}^{2n}\frac{2\xi_i(A_i-1/2)m_{1,\tau}(X_i)b(X_i)}{\sqrt{n}}\right\|_2 ||\hat{\theta}-\theta_0||_2.$$

In addition,

$$\begin{aligned}
&\sup_{\tau\in\Upsilon}\left\|\sum_{i=1}^{2n}\frac{2\xi_i(A_i-1/2)m_{1,\tau}(X_i)b(X_i)}{\sqrt{n}}\right\|_2 \\
=& \sup_{\tau\in\Upsilon,\rho\in\Re^K,||\rho||_2=1}\sum_{i=1}^{2n}\frac{2\xi_i(A_i-1/2)m_{1,\tau}(X_i)b'(X_i)\rho}{\sqrt{n}} \\
=& \sup_{\tau\in\Upsilon,\rho\in\Re^K,||\rho||_2=1}\sum_{j=1}^{n}\frac{(A_{\pi(2j-1)}-A_{\pi(2j)})(\xi_{\pi(2j-1)}m_{1,\tau}(X_{\pi(2j-1)})b'(X_{\pi(2j-1)})-\xi_{\pi(2j)}m_{1,\tau}(X_{\pi(2j)})b'(X_{\pi(2j)}))\rho}{\sqrt{n}}.
\end{aligned}$$

Conditional on $\{X_i,\xi_i\}_{i\in[2n]}$, $\{(A_{\pi(2j-1)}-A_{\pi(2j)})\}_{j=1}^n$ is a sequence of i.i.d. Rademacher random variables. In addition, let

$$\mathcal{F}_8 = \{(\xi_{\pi(2j-1)}m_{1,\tau}(X_{\pi(2j-1)})b'(X_{\pi(2j-1)})-\xi_{\pi(2j)}m_{1,\tau}(X_{\pi(2j)})b'(X_{\pi(2j)}))\rho : \tau\in\Upsilon, \rho\in R^K, ||\rho||_2=1\}$$

with envelope $F_j = (\xi_{\pi(2j-1)}\zeta(K)+\xi_{\pi(2j)}\zeta(K))$. Then, w.p.a.1,

$$\mathbb{E}\frac{1}{n}\sum_{j=1}^n F_j^2 \leq \frac{1}{n}\sum_{i=1}^{2n}\mathbb{E}\xi_i^2\zeta^2(K) \leq \overline{C}\zeta^2(K).$$

In addition, for some constant $c>0$,

$$\sup_Q N(\mathcal{F}_8, e_Q, \varepsilon||F||_{Q,2}) \leq \left(\frac{a}{\varepsilon}\right)^{cK}, \quad \forall\varepsilon\in(0,1].$$

Let $\sigma_n^2 = \sup_{f\in\mathcal{F}_8}\mathbb{P}_n f^2$ and $\delta^2 = \frac{\sigma_n^2}{\frac{1}{n}\sum_{j=1}^n F_j^2} \leq 1$. Then, by van der Vaart and Wellner (1996, Corollary 2.2.8), (E.2) and (E.6), we have, w.p.a.1,

$$
\begin{aligned}
&\mathbb{E}\mathbb{E}\left[\sup_{\tau\in\Upsilon}\left\|\sum_{i=1}^{2n}\frac{2\xi_i(A_i-1/2)m_{1,\tau}(X_i)b(X_i)}{\sqrt{n}}\right\|_2\middle|\{X_i,\xi_i\}_{i\in[2n]}\right]\\
&\lesssim \mathbb{E}\int_0^{\sigma_n}\sqrt{1+\log(N(\mathcal{F}_8,e_{\mathbb{P}_n},\varepsilon))}d\varepsilon\\
&\lesssim \mathbb{E}\sqrt{\frac{1}{n}\sum_{j=1}^{n}F_j^2}\int_0^{\delta}\sqrt{1+\log\sup_{Q}N(\mathcal{F}_8,e_Q,\varepsilon||F||_{Q,2})}d\varepsilon\\
&\leq \left(\mathbb{E}\sqrt{\frac{1}{n}\sum_{j=1}^{n}F_j^2}\right)\sqrt{K}J(1)\\
&\leq \left(\sqrt{\mathbb{E}\frac{1}{n}\sum_{j=1}^{n}F_j^2}\right)\sqrt{K}J(1)\\
&\lesssim \sqrt{K}\zeta(K).
\end{aligned}
$$

This implies

$$
\sup_{\tau\in\Upsilon}\left\|\sum_{i=1}^{2n}\frac{2\xi_i(A_i-1/2)m_{1,\tau}(X_i)b(X_i)}{\sqrt{n}}\right\|_2 = O_p(\sqrt{K}\zeta(K))
$$

and

$$
\sup_{\tau\in\Upsilon}\left\|\sum_{i=1}^{2n}\frac{2\xi_i(A_i-1/2)m_{1,\tau}(X_i)(\hat{A}_i-1/2)}{\sqrt{n}}\right\|_2 = O_p\left(\sqrt{\frac{K^2\zeta^2(K)\log(n)}{n}}\right) = o_p(1).
$$

Last, for the third result, we have

$$
\begin{aligned}
\sup_{\tau\in\Upsilon}\left|\sum_{i=1}^{2n}\frac{2\xi_iA_i\eta_{1,i}(\tau)(\hat{A}_i-1/2)}{\sqrt{n}}\right| &\leq \sup_{\tau\in\Upsilon}\left\|\sum_{i=1}^{2n}\frac{2\xi_iA_i\eta_{1,i}(\tau)b(X_i)}{\sqrt{n}}\right\|_2 ||\hat{\theta}-\theta_0||_2\\
&\leq \sup_{\tau\in\Upsilon,\rho\in\Re^K,||\rho||_2=1}\left[\sum_{i=1}^{2n}\frac{2\xi_iA_i\eta_{1,i}(\tau)b'(X_i)\rho}{\sqrt{n}}\right]||\hat{\theta}-\theta_0||_2. \quad \text{(E.10)}
\end{aligned}
$$

Let $\{\tilde{\varepsilon}_j\}_{j\in[n]}$ and $\{\varepsilon_i\}_{i\in[2n]}$ be two sequences of i.i.d. Rademacher random variables that are independent of the data. By (A.12), we have

$$
\sum_{i=1}^{2n}\frac{2\xi_iA_i\eta_{1,i}(\tau)b'(X_i)\rho}{\sqrt{n}}\middle|\{A_i,X_i\}_{i\in[2n]} \stackrel{d}{=} \sum_{j=1}^{n}\frac{2\tilde{\xi}_j\tilde{\eta}_{1,j}(\tau)b'(\tilde{X}_j)\rho}{\sqrt{n}}\middle|\{\tilde{X}_j\}_{j\in[n]},
$$

and

$$\sum_{i=1}^{2n} \frac{2\varepsilon_i \xi_i A_i \eta_{1,i}(\tau) b'(X_i)\rho}{\sqrt{n}} \bigg| \{A_i, X_i\}_{i\in[2n]} \overset{d}{=} \sum_{j=1}^{n} \frac{2\tilde{\varepsilon}_j \tilde{\xi}_j \tilde{\eta}_{1,j}(\tau) b'(\tilde{X}_j)\rho}{\sqrt{n}} \bigg| \{\tilde{X}_j\}_{j\in[n]},$$

where conditionally on $\{\tilde{X}_j\}_{j\in[n]}$, $\{\tilde{\xi}_j \tilde{\eta}_{1,j}(\tau)\}_{j\in[n]}$ is a sequence of independent random variables. Then, by the same argument as in (E.8), we have

$$\begin{aligned}
&\mathbb{E} \sup_{\tau\in\Upsilon, \rho\in\Re^K, ||\rho||_2=1} \left[\sum_{i=1}^{2n} \frac{2\xi_i A_i \eta_{1,i}(\tau) b'(X_i)\rho}{\sqrt{n}}\right] \\
&= \mathbb{E}\left\{\mathbb{E} \sup_{\tau\in\Upsilon, \rho\in\Re^K, ||\rho||_2=1} \left[\sum_{i=1}^{2n} \frac{2\xi_i A_i \eta_{1,i}(\tau) b'(X_i)\rho}{\sqrt{n}} \bigg| \{X_i, A_i\}_{i\in[2n]}\right]\right\} \\
&= \mathbb{E}\left\{\mathbb{E} \sup_{\tau\in\Upsilon, \rho\in\Re^K, ||\rho||_2=1} \left[\sum_{j=1}^{n} \frac{2\tilde{\xi}_j \tilde{\eta}_{1,j}(\tau) b'(\tilde{X}_j)\rho}{\sqrt{n}} \bigg| \{\tilde{X}_j\}_{j\in[n]}\right]\right\} \\
&\lesssim \mathbb{E}\left\{\mathbb{E} \sup_{\tau\in\Upsilon, \rho\in\Re^K, ||\rho||_2=1} \left[\sum_{j=1}^{n} \frac{2\tilde{\varepsilon}_j \tilde{\xi}_j \tilde{\eta}_{1,j}(\tau) b'(\tilde{X}_j)\rho}{\sqrt{n}} \bigg| \{\tilde{X}_j\}_{j\in[n]}\right]\right\} \\
&= \mathbb{E}\left\{\mathbb{E} \sup_{\tau\in\Upsilon, \rho\in\Re^K, ||\rho||_2=1} \left[\sum_{i=1}^{2n} \frac{2\varepsilon_i \xi_i A_i \eta_{1,i}(\tau) b'(X_i)\rho}{\sqrt{n}} \bigg| \{X_i, A_i\}_{i\in[2n]}\right]\right\} \\
&= \mathbb{E}\left\{\mathbb{E} \sup_{\tau\in\Upsilon, \rho\in\Re^K, ||\rho||_2=1} \left[\sum_{i=1}^{2n} \frac{2\varepsilon_i \xi_i A_i \eta_{1,i}(\tau) b'(X_i)\rho}{\sqrt{n}} \bigg| \{X_i, A_i, Y_i(1)\}_{i\in[2n]}\right]\right\}.
\end{aligned}$$

Let

$$\mathcal{F}_9 = \{2\xi_i A_i \eta_{1,i}(\tau) b'(X_i)\rho : \tau \in \Upsilon, \rho \in \Re^K, ||\rho||_2 = 1\},$$

with envelope $F_i = 2\xi_i \zeta(K)$. In addition, for some constant $c > 0$,

$$\sup_Q N(\mathcal{F}_9, e_Q, \varepsilon||F||_{Q,2}) \leq \left(\frac{a}{\varepsilon}\right)^{cK}, \quad \forall \varepsilon \in (0, 1].$$

Then, following (E.3) and (E.6), we have

$$\mathbb{E}\left\{\mathbb{E} \sup_{\tau\in\Upsilon, \rho\in\Re^K, ||\rho||_2=1} \left[\sum_{i=1}^{2n} \frac{2\varepsilon_i \xi_i A_i \eta_{1,i}(\tau) b'(X_i)\rho}{\sqrt{n}} \bigg| \{X_i, A_i, Y_i(1)\}_{i\in[2n]}\right]\right\} \lesssim ||F||_{\mathbb{P},2}\sqrt{K}J(1) \lesssim \sqrt{K}\zeta(K).$$

This implies

$$\sup_{\tau\in\Upsilon,\rho\in\Re^K,||\rho||_2=1}\sum_{i=1}^{2n}\frac{2\xi_i A_i\eta_{1,i}(\tau)b'(X_i)\rho}{\sqrt{n}}=O_p(\sqrt{K}\zeta(K)).$$

Then, by (E.10) and Lemma E.8, we have

$$\sup_{\tau\in\Upsilon}\left|\sum_{i=1}^{2n}\frac{2\xi_i A_i\eta_{1,i}(\tau)(\hat{A}_i-1/2)}{\sqrt{n}}\right|=O_p\left(\sqrt{\frac{K^2\zeta^2(K)\log(n)}{n}}\right)=o_p(1).$$

□

References

Chernozhukov, V., D. Chetverikov, and K. Kato (2014). Gaussian approximation of suprema of empirical processes. *The Annals of Statistics 42*(4), 1564–1597.

Kato, K. (2009). Asymptotics for argmin processes: Convexity arguments. *Journal of Multivariate Analysis 100*(8), 1816–1829.

Knight, K. (1998). Limiting distributions for l1 regression estimators under general conditions. *The Annals of Statistics 26*(2), 755–770.

Ledoux, M. and M. Talagrand (2013). *Probability in Banach Spaces: Isoperimetry and Processes.* Springer Science & Business Media.

van der Vaart, A. and J. A. Wellner (1996). *Weak Convergence and Empirical Processes.* Springer, New York.