LOCALLY EFFICIENT ESTIMATORS FOR PARTIALLY LINEAR SINGLE INDEX QUANTILE REGRESSION

Ruimiao Luo, Seungchul Baek and Yanyuan Ma

Shanxi University, University of Maryland Baltimore County and Pennsylvania State University

Supplementary Materials

The supplementary materials provide the proofs of Proposition 1, Theorem 1, Theorem 2 and Theorem 3 and additional figures.

S1 Proof of Proposition 1

Proof of Proposition 1: The score function for $\boldsymbol{\gamma}$ and $\boldsymbol{\beta}_{-1}$ denoted by $\mathbf{S}_{\boldsymbol{\gamma}}(y, \mathbf{x}, \mathbf{z})$ and $\mathbf{S}_{\boldsymbol{\beta}}(y, \mathbf{x}, \mathbf{z})$ is obtained as

$$\mathbf{S}_{\boldsymbol{\gamma}}(y, \mathbf{x}, \mathbf{z}) = -\frac{f'_\epsilon(\epsilon, \mathbf{x}, \mathbf{z})}{f_\epsilon(\epsilon, \mathbf{x}, \mathbf{z})}\mathbf{x}, \quad \mathbf{S}_{\boldsymbol{\beta}}(y, \mathbf{x}, \mathbf{z}) = -\frac{f'_\epsilon(\epsilon, \mathbf{x}, \mathbf{z})}{f_\epsilon(\epsilon, \mathbf{x}, \mathbf{z})} g'(\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta})\mathbf{z}_{-1}.$$

The nuisance tangent space for the three nuisance parameters $f_\epsilon(\epsilon, \mathbf{x}, \mathbf{z})$, $f_{\mathbf{x},\mathbf{z}}(\mathbf{x}, \mathbf{z})$ and $g(\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta})$ can be verified to be

$$\begin{aligned}
\Lambda \;=\; & (\mathbf{f}_1(\epsilon, \mathbf{x}, \mathbf{z}) + \mathbf{f}_2(\mathbf{x}, \mathbf{z}) + \mathbf{f}_3(\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}) f'_\epsilon(\epsilon, \mathbf{x}, \mathbf{z})/f_\epsilon(\epsilon, \mathbf{x}, \mathbf{z}) : \forall \mathbf{f}_1(\epsilon, \mathbf{x}, \mathbf{z}) \\
& \quad \text{s.t. } E\{\mathbf{f}_1(\epsilon, \mathbf{x}, \mathbf{z}) \mid \mathbf{x}, \mathbf{z}\} = \mathbf{0}, E[\{I(\epsilon < 0) - \tau\}\mathbf{f}_1(\epsilon, \mathbf{x}, \mathbf{z}) \mid \mathbf{x}, \mathbf{z}] = 0, \\
& \forall \mathbf{f}_2(\mathbf{x}, \mathbf{z}) \text{ s.t. } E\{\mathbf{f}_2(\mathbf{x}, \mathbf{z})\} = \mathbf{0}, \forall \mathbf{f}_3(\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta})),
\end{aligned}$$

and the orthogonal complement of Λ is

$$\Lambda^{\perp} = \left(\frac{f_\epsilon(0,\mathbf{x},\mathbf{z})}{\tau(1-\tau)}\{I(\epsilon<0)-\tau\}[\mathbf{a}(\mathbf{x},\mathbf{z}) - \frac{E\{f_\epsilon^2(0,\mathbf{x},\mathbf{z})\mathbf{a}(\mathbf{x},\mathbf{z})|\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}\}}{E\{f_\epsilon^2(0,\mathbf{x},\mathbf{z})|\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}\}}] : \forall \mathbf{a}(\mathbf{x},\mathbf{z}) \right).$$

We now decompose $\mathbf{S}_{\boldsymbol{\beta}}(\epsilon,\mathbf{x},\mathbf{z})$ and $\mathbf{S}_{\boldsymbol{\gamma}}(\epsilon,\mathbf{x},\mathbf{z})$ as

$$\begin{aligned}
\mathbf{S}_{\boldsymbol{\gamma}}(\epsilon,\mathbf{x},\mathbf{z}) &= -\left[\frac{f_\epsilon'(\epsilon,\mathbf{x},\mathbf{z})}{f_\epsilon(\epsilon,\mathbf{x},\mathbf{z})} - \frac{f_\epsilon(0,\mathbf{x},\mathbf{z})}{\tau(1-\tau)}\{I(\epsilon<0)-\tau\}\right]\mathbf{x} \\
&\quad - \frac{f_\epsilon(0,\mathbf{x},\mathbf{z})}{\tau(1-\tau)}\{I(\epsilon<0)-\tau\}\frac{E\{f_\epsilon^2(0,\mathbf{x},\mathbf{z})\mathbf{x}|\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}\}}{E\{f_\epsilon^2(0,\mathbf{x},\mathbf{z})|\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}\}} \\
&\quad - \frac{f_\epsilon(0,\mathbf{x},\mathbf{z})}{\tau(1-\tau)}\{I(\epsilon<0)-\tau\}\left[\mathbf{x} - \frac{E\{f_\epsilon^2(0,\mathbf{x},\mathbf{z})\mathbf{x}|\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}\}}{E\{f_\epsilon^2(0,\mathbf{x},\mathbf{z})|\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}\}}\right], \\
\mathbf{S}_{\boldsymbol{\beta}}(\epsilon,\mathbf{x},\mathbf{z}) &= -\left[\frac{f_\epsilon'(\epsilon,\mathbf{x},\mathbf{z})}{f_\epsilon(\epsilon,\mathbf{x},\mathbf{z})} - \frac{f_\epsilon(0,\mathbf{x},\mathbf{z})}{\tau(1-\tau)}\{I(\epsilon<0)-\tau\}\right]g'(\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta})\mathbf{z}_{-1} \\
&\quad - \frac{f_\epsilon(0,\mathbf{x},\mathbf{z})}{\tau(1-\tau)}\{I(\epsilon<0)-\tau\}g'(\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta})\frac{E\{f_\epsilon^2(0,\mathbf{x},\mathbf{z})\mathbf{z}_{-1}|\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}\}}{E\{f_\epsilon^2(0,\mathbf{x},\mathbf{z})|\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}\}} \\
&\quad - \frac{f_\epsilon(0,\mathbf{x},\mathbf{z})}{\tau(1-\tau)}\{I(\epsilon<0)-\tau\}g'(\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta})\left[\mathbf{z}_{-1} - \frac{E\{f_\epsilon^2(0,\mathbf{x},\mathbf{z})\mathbf{z}_{-1}|\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}\}}{E\{f_\epsilon^2(0,\mathbf{x},\mathbf{z})|\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}\}}\right].
\end{aligned}$$

It is easily verified that the first two terms of $\{\mathbf{S}_{\boldsymbol{\gamma}}^{\mathrm{T}}(\epsilon,\mathbf{x},\mathbf{z}),\mathbf{S}_{\boldsymbol{\beta}}^{\mathrm{T}}(\epsilon,\mathbf{x},\mathbf{z})\}^{\mathrm{T}}$ are elements in the nuisance tangent space Λ, while the last term belongs to the orthogonal complement $\Lambda^{\perp}$. Hence the efficient score function which is the projection of the score function is the last component.

S2 Notations

To ease the description, we first introduce several notations. Define

$$\begin{aligned}
\mathbf{x}_i^c &\equiv \mathbf{x}_i - \frac{E\{f_\epsilon^{*2}(0,\mathbf{x}_i,\mathbf{z}_i)\mathbf{x}_i|\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}\}}{E\{f_\epsilon^{*2}(0,\mathbf{x}_i,\mathbf{z}_i)|\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}\}}, \\
\mathbf{z}_{-1,i}^c &\equiv \mathbf{z}_{-1,i} - \frac{E\{f_\epsilon^{*2}(0,\mathbf{x}_i,\mathbf{z}_i)\mathbf{z}_{-1,i}|\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}\}}{E\{f_\epsilon^{*2}(0,\mathbf{x}_i,\mathbf{z}_i)|\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}\}},
\end{aligned}$$

$$P(\mathbf{x}_i,\mathbf{z}_i) \equiv \{(\mathbf{x}_i^c)^{\mathrm{T}}, g'(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})(\mathbf{z}_{-1,i}^c)^{\mathrm{T}}\}^{\mathrm{T}},$$

$$\widehat{P}(\mathbf{x}_i,\mathbf{z}_i) \equiv \{(\widehat{\mathbf{x}}_i^c)^{\mathrm{T}}, \widetilde{g}'(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}},\widetilde{\boldsymbol{\gamma}})(\widehat{\mathbf{z}}_{-1,i}^c)^{\mathrm{T}}\}^{\mathrm{T}},$$

$$\alpha \equiv \sup_{\mathbf{x},\mathbf{z}} \|P(\mathbf{x},\mathbf{z})\|,$$

$$V_0 \equiv E\left\{\frac{f_\epsilon^{*2}(0,\mathbf{x}_i,\mathbf{z}_i)}{\tau(1-\tau)}\begin{pmatrix}\mathbf{x}_i^c \\ \\ g'(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})\mathbf{z}_{-1,i}^c\end{pmatrix}\begin{pmatrix}\mathbf{x}_i^{\mathrm{T}} & g'(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})\mathbf{z}_{-1,i}^{\mathrm{T}}\end{pmatrix}\right\},$$

$$\begin{aligned}
&G_i\{\widetilde{\boldsymbol{\gamma}},\widetilde{\boldsymbol{\beta}},\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}},\widetilde{\boldsymbol{\gamma}})\}\\
\equiv\ &\frac{f_\epsilon^{*}(0,\mathbf{x}_i,\mathbf{z}_i)}{\tau(1-\tau)}[\tau - I\{Y_i - \mathbf{x}_i^{\mathrm{T}}\widetilde{\boldsymbol{\gamma}} - \widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}},\widetilde{\boldsymbol{\gamma}}) < 0\}]P(\mathbf{x}_i,\mathbf{z}_i),\\
&G_i\{\boldsymbol{\gamma},\boldsymbol{\beta},g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})\}\\
\equiv\ &\frac{f_\epsilon^{*}(0,\mathbf{x}_i,\mathbf{z}_i)}{\tau(1-\tau)}[\tau - I\{Y_i - \mathbf{x}_i^{\mathrm{T}}\boldsymbol{\gamma} - g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) < 0\}]P(\mathbf{x}_i,\mathbf{z}_i),\\
&G_{pi}\{\widetilde{\boldsymbol{\gamma}},\widetilde{\boldsymbol{\beta}},\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}},\widetilde{\boldsymbol{\gamma}})\}\\
\equiv\ &\frac{f_\epsilon^{*}(0,\mathbf{x}_i,\mathbf{z}_i)}{\tau(1-\tau)}(\tau - F_\epsilon[\{\mathbf{x}_i^{\mathrm{T}}(\widetilde{\boldsymbol{\gamma}}-\boldsymbol{\gamma}) + \widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}},\widetilde{\boldsymbol{\gamma}}) - g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})\},\mathbf{x}_i,\mathbf{z}_i])P(\mathbf{x}_i,\mathbf{z}_i),\\
&\widetilde{G}_{pi}\{\widetilde{\boldsymbol{\gamma}},\widetilde{\boldsymbol{\beta}},\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}},\widetilde{\boldsymbol{\gamma}})\}\\
\equiv\ &\frac{f_\epsilon^{*}(0,\mathbf{x}_i,\mathbf{z}_i)f_\epsilon(0,\mathbf{x}_i,\mathbf{z}_i)}{-\tau(1-\tau)}\{\mathbf{x}_i^{\mathrm{T}}(\widetilde{\boldsymbol{\gamma}}-\boldsymbol{\gamma}) + \widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}},\widetilde{\boldsymbol{\gamma}}) - g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})\}P(\mathbf{x}_i,\mathbf{z}_i).
\end{aligned}$$

Finally, let

$$\begin{aligned}
&\eta_{mi}\{\boldsymbol{\gamma},\boldsymbol{\beta},g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}),\boldsymbol{\gamma}^*,\boldsymbol{\beta}^*,\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*,\boldsymbol{\gamma}^*)\}\\
\equiv\ &G_i\{\boldsymbol{\gamma}^*,\boldsymbol{\beta}^*,\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*,\boldsymbol{\gamma}^*)\} - G_i\{\boldsymbol{\gamma},\boldsymbol{\beta},g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})\} - G_{pi}\{\boldsymbol{\gamma}^*,\boldsymbol{\beta}^*,\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*,\boldsymbol{\gamma}^*)\},
\end{aligned}$$

and $\mathscr{D} \equiv \{(\boldsymbol{\beta}^{*\mathrm{T}},\boldsymbol{\gamma}^{*\mathrm{T}})^{\mathrm{T}} : \|\boldsymbol{\beta}^* - \boldsymbol{\beta}\| \leqslant Cn^{-1/2}, \|\boldsymbol{\gamma}^* - \boldsymbol{\gamma}\| \leqslant Cn^{-1/2}, \boldsymbol{\beta}_1^* = 1\}$. Note that here, $E[G_i\{\boldsymbol{\gamma},\boldsymbol{\beta},g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})\}|\mathbf{x}_i,\mathbf{z}_i] = 0$. With the dependence of the estimates on the data ignored, $G_{pi}\{\widetilde{\boldsymbol{\gamma}},\widetilde{\boldsymbol{\beta}},\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}},\widetilde{\boldsymbol{\gamma}})\}$ can be viewed as the expectation of $G_i\{\widetilde{\boldsymbol{\gamma}},\widetilde{\boldsymbol{\beta}},\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}},\widetilde{\boldsymbol{\gamma}})\}$ given $\mathbf{x}_i$ and $\mathbf{z}_i$. Further, $\widetilde{G}_{pi}\{\widetilde{\boldsymbol{\gamma}},\widetilde{\boldsymbol{\beta}},\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}},\widetilde{\boldsymbol{\gamma}})\}$ is the first order approximation of $G_{pi}\{\widetilde{\boldsymbol{\gamma}},\widetilde{\boldsymbol{\beta}},\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}},\widetilde{\boldsymbol{\gamma}})\}$.

Define

$$
\begin{aligned}
\mathbf{x}_i^{c0} &\equiv \mathbf{x}_i - \frac{E\{f_\epsilon(0,\mathbf{x}_i,\mathbf{z}_i)\mathbf{x}_i|\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}\}}{E\{f_\epsilon(0,\mathbf{x}_i,\mathbf{z}_i)|\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}\}},\\
\mathbf{z}_{-1,i}^{c0} &\equiv \mathbf{z}_{-1,i} - \frac{E\{f_\epsilon(0,\mathbf{x}_i,\mathbf{z}_i)\mathbf{z}_{-1,i}|\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}\}}{E\{f_\epsilon(0,\mathbf{x}_i,\mathbf{z}_i)|\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}\}},\\
P^0(\mathbf{x}_i,\mathbf{z}_i) &\equiv \{(\mathbf{x}_i^{c0})^{\mathrm{T}}, g'(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})(\mathbf{z}_{-1,i}^{c0})^{\mathrm{T}}\}^{\mathrm{T}},\\
V &\equiv E\left\{\frac{f_\epsilon^*(0,\mathbf{x}_i,\mathbf{z}_i)f_\epsilon(0,\mathbf{x}_i,\mathbf{z}_i)}{\tau(1-\tau)}P(\mathbf{x}_i,\mathbf{z}_i)P^0(\mathbf{x}_i,\mathbf{z}_i)^{\mathrm{T}}\right\}.
\end{aligned}
$$

S3 Proof of Theorem 1

Define

$$
\widetilde{F}_p(\phi,\boldsymbol{\gamma},\boldsymbol{\beta}) \equiv \frac{\sum_{j\in I_1}K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}-\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})\mathrm{pr}\{Y_j-\mathbf{x}_j^{\mathrm{T}}\boldsymbol{\gamma}-\phi<0|\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}\}}{\sum_{j\in I_1}K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}-\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})}.
$$

We first introduce a new notation $\widetilde{g}_p(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})$, which is defined as the solution of ϕ that solves $\widetilde{F}_p(\phi,\boldsymbol{\gamma},\boldsymbol{\beta})=\tau$. Thus, we always have $\widetilde{F}_p\{\widetilde{g}_p(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}),\boldsymbol{\gamma},\boldsymbol{\beta}\}=\tau$. We first prove two lemmas.

<u>Lemma</u> 1. Under the Regularity Conditions C2-C5, $\widetilde{g}_p(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})-g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})=O_p\{k(n,h_1)\}$.

Proof. We aim to show that there exists a constant $C>0$ such that $\mathrm{pr}\{g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})-Ck(n,h_1)\leqslant \widetilde{g}_p(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})\leqslant g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})+Ck(n,h_1)\}\to 1$ as $n\to\infty$. It is equivalent to prove that, for arbitrary $C_1>C>0$, $\mathrm{pr}\{|\widetilde{g}_p(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})-g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})|>C_1k(n,h_1)\}\to 0$ as $n\to\infty$. We prove this by contradiction. If $|\widetilde{g}_p(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})-g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})|>C_1k(n,h_1)$, then $\widetilde{g}_p(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})>g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})+C_1k(n,h_1)$ or $\widetilde{g}_p(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})<g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})-C_1k(n,h_1)$, then by the monotonicity of the conditional distribution function, we have

$$
\mathrm{pr}\{Y_i-\mathbf{x}_i^{\mathrm{T}}\boldsymbol{\gamma}<\widetilde{g}_p(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})|\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}\}\geqslant \mathrm{pr}\{\epsilon_i<C_1k(n,h_1)|\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}\}, \tag{S3.1}
$$

or

$$\mathrm{pr}\{Y_i - \mathbf{x}_i^{\mathrm{T}}\boldsymbol{\gamma} < \widetilde{g}_p(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})|\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}\} \leqslant \mathrm{pr}\{\epsilon_i < -C_1 k(n, h_1)|\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}\}. \tag{S3.2}$$

For any C_1, under Conditions C2, C4 and C5, standard kernel regression analysis leads to

$$\begin{aligned} \widetilde{F}_p\{\widetilde{g}_p(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}), \boldsymbol{\gamma}, \boldsymbol{\beta}\} &= \frac{\sum_{j\in I_1} K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta} - \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})\mathrm{pr}\{Y_j - \mathbf{x}_j^{\mathrm{T}}\boldsymbol{\gamma} < \widetilde{g}_p(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})|\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}\}}{\sum_{j\in I_1} K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta} - \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})} \\ &= \mathrm{pr}\{Y_i - \mathbf{x}_i^{\mathrm{T}}\boldsymbol{\gamma} < \widetilde{g}_p(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})|\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}\} + O_p\{k(n, h_1)\}. \end{aligned}$$

Thus, for any $\delta > 0$, there exists a constant $C_2 > 0$, so that with probability at least $1 - \delta$,

$$\mathrm{pr}\{Y_i - \mathbf{x}_i^{\mathrm{T}}\boldsymbol{\gamma} < \widetilde{g}_p(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})|\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}\} \leqslant \widetilde{F}_p\{\widetilde{g}_p(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}), \boldsymbol{\gamma}, \boldsymbol{\beta}\} + C_2 k(n, h_1),$$

and

$$\widetilde{F}_p\{\widetilde{g}_p(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}), \boldsymbol{\gamma}, \boldsymbol{\beta}\} - C_2 k(n, h_1) \leqslant \mathrm{pr}\{Y_i - \mathbf{x}_i^{\mathrm{T}}\boldsymbol{\gamma} < \widetilde{g}_p(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})|\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}\}.$$

Note that $\widetilde{F}_p\{\widetilde{g}_p(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}), \boldsymbol{\gamma}, \boldsymbol{\beta}\} = \tau$ and $\mathrm{pr}(\epsilon_i < 0|\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) = \tau$. Then the above equations become

$$\mathrm{pr}\{Y_i - \mathbf{x}_i^{\mathrm{T}}\boldsymbol{\gamma} < \widetilde{g}_p(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})|\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}\} \leqslant \mathrm{pr}(\epsilon_i < 0|\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) + C_2 k(n, h_1), \tag{S3.3}$$

and

$$\mathrm{pr}(\epsilon_i < 0|\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - C_2 k(n, h_1) \leqslant \mathrm{pr}\{Y_i - \mathbf{x}_i^{\mathrm{T}}\boldsymbol{\gamma} < \widetilde{g}_p(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})|\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}\}. \tag{S3.4}$$

Combining (S3.1), (S3.2), (S3.3) and (S3.4), we get

$$\mathrm{pr}\{\epsilon_i < C_1 k(n,h_1)|\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}\} - \mathrm{pr}(\epsilon_i < 0|\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) \leqslant C_2 k(n,h_1),$$

or

$$-C_2 k(n,h_1) \leqslant \mathrm{pr}\{\epsilon_i < -C_1 k(n,h_1)|\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}\} - \mathrm{pr}(\epsilon_i < 0|\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}).$$

But by a first-order Taylor series expansion and Condition C2,

$$\mathrm{pr}(\epsilon_i < 0|\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) + f_{\epsilon|\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}}(0,\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})C_1 k(n,h_1)/2 \;\; \leqslant \;\; \mathrm{pr}\{\epsilon_i < C_1 k(n,h_1)|\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}\},$$

and

$$\mathrm{pr}\{\epsilon_i < -C_1 k(n,h_1)|\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}\} \;\; \leqslant \;\; \mathrm{pr}(\epsilon_i < 0|\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - f_{\epsilon|\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}}(0,\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})C_1 k(n,h_1)/2.$$

This leads to $f_{\epsilon|\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}}(0,\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})C_1 k(n,h_1)/2 \leqslant C_2 k(n,h_1)$, which is impossible to hold for a fixed C_2 and arbitrary $C_1 > C > 0$ under Condition C2. Thus,

$$\begin{aligned}\mathrm{pr}\{|\widetilde{g}_p(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})| > C_1 k(n,h_1)\} &\leqslant \mathrm{pr}\{\widetilde{g}_p(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) > g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) + C_1 k(n,h_1)\} \\ &\quad + \mathrm{pr}\{\widetilde{g}_p(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) < g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - C_1 k(n,h_1)\} \to 0.\end{aligned}$$

Lemma 1 is proved.

Lemma 2. Under the Regularity Conditions C1-C8, $\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}},\widetilde{\boldsymbol{\gamma}}) - \widetilde{g}_p(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) = O_p\{k(n,h_1)\}$.

Proof. Recall that

$$\widetilde{F}(\phi, \widetilde{\boldsymbol{\gamma}}, \widetilde{\boldsymbol{\beta}}) \equiv \frac{\sum_{j \in I_1} K_{h_1}(\mathbf{z}_j^{\mathrm{T}} \widetilde{\boldsymbol{\beta}} - \mathbf{z}_i^{\mathrm{T}} \widetilde{\boldsymbol{\beta}}) I(Y_j - \mathbf{x}_j^{\mathrm{T}} \widetilde{\boldsymbol{\gamma}} - \phi < 0)}{\sum_{j \in I_1} K_{h_1}(\mathbf{z}_j^{\mathrm{T}} \widetilde{\boldsymbol{\beta}} - \mathbf{z}_i^{\mathrm{T}} \widetilde{\boldsymbol{\beta}})}.$$

Note that $\widetilde{F}\{\widetilde{g}(\mathbf{z}_i^{\mathrm{T}} \widetilde{\boldsymbol{\beta}}, \widetilde{\boldsymbol{\gamma}}), \widetilde{\boldsymbol{\gamma}}, \widetilde{\boldsymbol{\beta}}\} = \tau$.

We aim to show that there exists a constant $C > 0$ such that $\mathrm{pr}\{\widetilde{g}_p(\mathbf{z}_i^{\mathrm{T}} \boldsymbol{\beta}) - Ck(n, h_1) \leqslant \widetilde{g}(\mathbf{z}_i^{\mathrm{T}} \widetilde{\boldsymbol{\beta}}, \widetilde{\boldsymbol{\gamma}}) \leqslant \widetilde{g}_p(\mathbf{z}_i^{\mathrm{T}} \boldsymbol{\beta}) + Ck(n, h_1)\} \to 1$ as $n \to \infty$. It is equivalent to prove that, for arbitrary $C_1 > C > 0$, $\mathrm{pr}\{|\widetilde{g}(\mathbf{z}_i^{\mathrm{T}} \widetilde{\boldsymbol{\beta}}, \widetilde{\boldsymbol{\gamma}}) - \widetilde{g}_p(\mathbf{z}_i^{\mathrm{T}} \boldsymbol{\beta})| > C_1 k(n, h_1)\} \to 0$ as $n \to \infty$. We prove this by contradiction. If $|\widetilde{g}(\mathbf{z}_i^{\mathrm{T}} \widetilde{\boldsymbol{\beta}}, \widetilde{\boldsymbol{\gamma}}) - \widetilde{g}_p(\mathbf{z}_i^{\mathrm{T}} \boldsymbol{\beta})| > C_1 k(n, h_1)$, then either $\widetilde{g}(\mathbf{z}_i^{\mathrm{T}} \widetilde{\boldsymbol{\beta}}, \widetilde{\boldsymbol{\gamma}}) > \widetilde{g}_p(\mathbf{z}_i^{\mathrm{T}} \boldsymbol{\beta}) + C_1 k(n, h_1)$ or $\widetilde{g}(\mathbf{z}_i^{\mathrm{T}} \widetilde{\boldsymbol{\beta}}, \widetilde{\boldsymbol{\gamma}}) < \widetilde{g}_p(\mathbf{z}_i^{\mathrm{T}} \boldsymbol{\beta}) - C_1 k(n, h_1)$.

Consider first the case when $\widetilde{g}(\mathbf{z}_i^{\mathrm{T}} \widetilde{\boldsymbol{\beta}}, \widetilde{\boldsymbol{\gamma}}) < \widetilde{g}_p(\mathbf{z}_i^{\mathrm{T}} \boldsymbol{\beta}) - C_1 k(n, h_1)$. By the monotonicity of the conditional distribution function, we have

$$\begin{aligned} & \mathrm{pr}\{Y_i - \mathbf{x}_i^{\mathrm{T}} \boldsymbol{\gamma} < \widetilde{g}(\mathbf{z}_i^{\mathrm{T}} \widetilde{\boldsymbol{\beta}}, \widetilde{\boldsymbol{\gamma}}) | \mathbf{z}_i^{\mathrm{T}} \boldsymbol{\beta}\} \\ \leqslant \; & \mathrm{pr}\{Y_i - \mathbf{x}_i^{\mathrm{T}} \boldsymbol{\gamma} < \widetilde{g}_p(\mathbf{z}_i^{\mathrm{T}} \boldsymbol{\beta}) - C_1 k(n, h_1) | \mathbf{z}_i^{\mathrm{T}} \boldsymbol{\beta}\}. \end{aligned} \tag{S3.5}$$

Under Conditions C2, C4 and C5, standard kernel regression analysis leads to

$$\begin{aligned} \widetilde{F}_p\{\widetilde{g}(\mathbf{z}_i^{\mathrm{T}} \widetilde{\boldsymbol{\beta}}, \widetilde{\boldsymbol{\gamma}}), \boldsymbol{\gamma}, \boldsymbol{\beta}\} &= \frac{\sum_{j \in I_1} K_{h_1}(\mathbf{z}_j^{\mathrm{T}} \boldsymbol{\beta} - \mathbf{z}_i^{\mathrm{T}} \boldsymbol{\beta}) \mathrm{pr}\{Y_j - \mathbf{x}_j^{\mathrm{T}} \boldsymbol{\gamma} < \widetilde{g}(\mathbf{z}_i^{\mathrm{T}} \widetilde{\boldsymbol{\beta}}, \widetilde{\boldsymbol{\gamma}}) | \mathbf{z}_j^{\mathrm{T}} \boldsymbol{\beta}\}}{\sum_{j \in I_1} K_{h_1}(\mathbf{z}_j^{\mathrm{T}} \boldsymbol{\beta} - \mathbf{z}_i^{\mathrm{T}} \boldsymbol{\beta})} \\ &= \mathrm{pr}\{Y_i - \mathbf{x}_i^{\mathrm{T}} \boldsymbol{\gamma} < \widetilde{g}(\mathbf{z}_i^{\mathrm{T}} \widetilde{\boldsymbol{\beta}}, \widetilde{\boldsymbol{\gamma}}) | \mathbf{z}_i^{\mathrm{T}} \boldsymbol{\beta}\} + O_p\{k(n, h_1)\}. \end{aligned}$$

$$\begin{aligned} & \widetilde{F}_p\{\widetilde{g}_p(\mathbf{z}_i^{\mathrm{T}} \boldsymbol{\beta}) - C_1 k(n, h_1), \boldsymbol{\gamma}, \boldsymbol{\beta}\} \\ = \; & \frac{\sum_{j \in I_1} K_{h_1}(\mathbf{z}_j^{\mathrm{T}} \boldsymbol{\beta} - \mathbf{z}_i^{\mathrm{T}} \boldsymbol{\beta}) \mathrm{pr}\{Y_j - \mathbf{x}_j^{\mathrm{T}} \boldsymbol{\gamma} < \widetilde{g}_p(\mathbf{z}_i^{\mathrm{T}} \boldsymbol{\beta}) - C_1 k(n, h_1) | \mathbf{z}_j^{\mathrm{T}} \boldsymbol{\beta}\}}{\sum_{j \in I_1} K_{h_1}(\mathbf{z}_j^{\mathrm{T}} \boldsymbol{\beta} - \mathbf{z}_i^{\mathrm{T}} \boldsymbol{\beta})} \\ = \; & \mathrm{pr}\{Y_i - \mathbf{x}_i^{\mathrm{T}} \boldsymbol{\gamma} < \widetilde{g}_p(\mathbf{z}_i^{\mathrm{T}} \boldsymbol{\beta}) - C_1 k(n, h_1) | \mathbf{z}_i^{\mathrm{T}} \boldsymbol{\beta}\} + O_p\{k(n, h_1)\}. \end{aligned}$$

Thus, for any $\delta > 0$, there exists a constant $C_2 > 0$, so that with probability at least $1 - \delta$,

$$\widetilde{F}_p\{\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}, \widetilde{\gamma}), \boldsymbol{\gamma}, \boldsymbol{\beta}\} - C_2 k(n, h_1) \leqslant \mathrm{pr}\{Y_i - \mathbf{x}_i^{\mathrm{T}}\boldsymbol{\gamma} < \widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}, \widetilde{\boldsymbol{\gamma}}) | \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}\}, \tag{S3.6}$$

$$\begin{aligned} & \mathrm{pr}\{Y_i - \mathbf{x}_i^{\mathrm{T}}\boldsymbol{\gamma} < \widetilde{g}_p(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - C_1 k(n, h_1) | \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}\} \\ \leqslant \quad & \widetilde{F}_p\{\widetilde{g}_p(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - C_1 k(n, h_1), \boldsymbol{\gamma}, \boldsymbol{\beta}\} + C_2 k(n, h_1). \end{aligned} \tag{S3.7}$$

Combining (S3.5), (S3.6) and (S3.7), we get

$$\widetilde{F}_p\{\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}, \widetilde{\gamma}), \boldsymbol{\gamma}, \boldsymbol{\beta}\} \leqslant \widetilde{F}_p\{\widetilde{g}_p(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - C_1 k(n, h_1), \boldsymbol{\gamma}, \boldsymbol{\beta}\} + 2C_2 k(n, h_1). \tag{S3.8}$$

By direct calculation, we get

$$\begin{aligned} & \widetilde{F}\{\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}, \widetilde{\gamma}), \widetilde{\gamma}, \widetilde{\boldsymbol{\beta}}\} - \widetilde{F}\{\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}, \widetilde{\gamma}), \widetilde{\gamma}, \boldsymbol{\beta}\} \\ = \quad & \frac{\sum_{j \in I_1}\{K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\widetilde{\boldsymbol{\beta}} - \mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}) - K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta} - \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})\} I\{Y_j - \mathbf{x}_j^{\mathrm{T}}\widetilde{\boldsymbol{\gamma}} < \widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}, \widetilde{\gamma})\}}{\sum_{j \in I_1} K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta} - \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})} \\ & + \frac{\sum_{j \in I_1} K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\widetilde{\boldsymbol{\beta}} - \mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}) I\{Y_j - \mathbf{x}_j^{\mathrm{T}}\widetilde{\boldsymbol{\gamma}} < \widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}, \widetilde{\gamma})\}}{\sum_{j \in I_1} K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\widetilde{\boldsymbol{\beta}} - \mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}})} \\ & \times \frac{\sum_{j \in I_1}\{K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta} - \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\widetilde{\boldsymbol{\beta}} - \mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}})\}}{\sum_{j \in I_1} K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta} - \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})}. \end{aligned} \tag{S3.9}$$

For notational brevity, let

$$\xi_j(\mathbf{z}_i, \boldsymbol{\beta}) \equiv K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta} - \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) I\{Y_j - \mathbf{x}_j^{\mathrm{T}}\widetilde{\boldsymbol{\gamma}} < \widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}, \widetilde{\gamma})\}, j \in I_1.$$

Under Conditions C1,C5, and C6, for all $\mathbf{z}_i$ and $||\boldsymbol{\beta}^* - \boldsymbol{\beta}|| = O_p(n^{-1/2})$,

$$E\{\xi_j(\mathbf{z}_i, \boldsymbol{\beta}^*) - \xi_j(\mathbf{z}_i, \boldsymbol{\beta})\}^2 = O_p(n^{-1}h_1^{-4}). \tag{S3.10}$$

By Chebyshev's inequality, we can get

$$\mathrm{pr}(|\xi_j(\mathbf{z}_i, \boldsymbol{\beta}^*) - \xi_j(\mathbf{z}_i, \boldsymbol{\beta})| \geqslant 1) \leqslant E\{\xi_j(\mathbf{z}_i, \boldsymbol{\beta}^*) - \xi_j(\mathbf{z}_i, \boldsymbol{\beta})\}^2. \tag{S3.11}$$

Let $\zeta(\mathbf{z}_i, \boldsymbol{\beta}^*, \boldsymbol{\beta}) = 1/n\sum_{j\in I_1}\{\xi_j(\mathbf{z}_i, \boldsymbol{\beta}^*) - \xi_j(\mathbf{z}_i, \boldsymbol{\beta})\}$. Under Condition C4, (S3.10) and (S3.11), together with Lemma 18 (page 20), Lemma 25 (page 27), Lemma 28 (page 30), the proof of Lemma 33 (page 31-33) and Theorem 37 (page 34) in Pollard (1984), yield

$$\sup_{\|\boldsymbol{\beta}^*-\boldsymbol{\beta}\|=O_p(n^{-1/2})} \sup_{\mathbf{z}_i} |\zeta(\mathbf{z}_i, \boldsymbol{\beta}^*, \boldsymbol{\beta}) - E\zeta(\mathbf{z}_i, \boldsymbol{\beta}^*, \boldsymbol{\beta})| = O_p(n^{-1}h_1^{-2}\mathrm{log}n) = o_p(n^{-1/2}) \tag{S3.12}$$

under Condition C4. By direct derivation, at fixed $\mathbf{z}_i$, for any $\boldsymbol{\beta}^*$ satisfying $\|\boldsymbol{\beta}^* - \boldsymbol{\beta}\| = O_p(n^{-1/2})$, we have

$$E\{\xi_j(\mathbf{z}_i, \boldsymbol{\beta}^*)\} = \int K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}^* - \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*)w(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}^*)\mathrm{d}\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}^* = w(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*) + O_p(h_1^4). \tag{S3.13}$$

where $w(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}^*) \equiv \mathrm{pr}\{Y_j - \mathbf{x}_j^{\mathrm{T}}\widetilde{\boldsymbol{\gamma}} < \widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}, \widetilde{\boldsymbol{\gamma}})|\mathbf{z}_i, \mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}^*\}f_{\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}^*}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}^*)$. Similarly, it can be shown that $E\xi_j(\mathbf{z}_i, \boldsymbol{\beta}) = w(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) + O_p(h_1^4)$. Then under Conditions C7 and C8, we have $|E\zeta(\mathbf{z}_i, \boldsymbol{\beta}^*, \boldsymbol{\beta})\}| = |E\{\xi_j(\mathbf{z}_i, \boldsymbol{\beta}^*) - \xi_j(\mathbf{z}_i, \boldsymbol{\beta})\}| = |w(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*) - w(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})| + O_p(h_1^4)$. This means

$$\sup_{\|\boldsymbol{\beta}^*-\boldsymbol{\beta}\|=O_p(n^{-1/2})} \sup_{\mathbf{z}_i} |E\zeta(\mathbf{z}_i, \boldsymbol{\beta}^*, \boldsymbol{\beta})| = O_p(n^{-1/2} + h_1^4). \tag{S3.14}$$

Under Condition C4, combining (S3.12) and (S3.14), it is obtained that the first term in

(S3.9) is $O_p(n^{-1/2})$. Similarly, we can prove the second term in (S3.9) is $O_p(n^{-1/2})$. Then we have

$$\widetilde{F}\{\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}},\widetilde{\boldsymbol{\gamma}}),\widetilde{\boldsymbol{\gamma}},\widetilde{\boldsymbol{\beta}}\}-\widetilde{F}\{\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}},\widetilde{\boldsymbol{\gamma}}),\widetilde{\boldsymbol{\gamma}},\boldsymbol{\beta}\}=O_p(n^{-1/2}). \tag{S3.15}$$

Under Conditions C2, C4 and C5, standard kernel regression analysis leads to

$$\widetilde{F}\{\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}},\widetilde{\boldsymbol{\gamma}}),\widetilde{\boldsymbol{\gamma}},\boldsymbol{\beta}\}-\widetilde{F}_p\{\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}},\widetilde{\boldsymbol{\gamma}}),\widetilde{\boldsymbol{\gamma}},\boldsymbol{\beta}\}=O_p\{k(n,h_1)\}. \tag{S3.16}$$

Under Conditions C1 and C3, we also have

$$\widetilde{F}_p\{\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}},\widetilde{\boldsymbol{\gamma}}),\widetilde{\boldsymbol{\gamma}},\boldsymbol{\beta}\}-\widetilde{F}_p\{\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}},\widetilde{\boldsymbol{\gamma}}),\boldsymbol{\gamma},\boldsymbol{\beta}\}=O_p(n^{-1/2}). \tag{S3.17}$$

Combining (S3.15), (S3.16) and (S3.17), we get

$$\widetilde{F}\{\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}},\widetilde{\boldsymbol{\gamma}}),\widetilde{\boldsymbol{\gamma}},\widetilde{\boldsymbol{\beta}}\}-\widetilde{F}_p\{\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}},\widetilde{\boldsymbol{\gamma}}),\boldsymbol{\gamma},\boldsymbol{\beta}\}=O_p\{k(n,h_1)\}.$$

Thus, for any $\delta>0$, there exists a constant $C_3>0$, so that with probability at least $1-\delta$,

$$\widetilde{F}\{\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}},\widetilde{\boldsymbol{\gamma}}),\widetilde{\boldsymbol{\gamma}},\widetilde{\boldsymbol{\beta}}\}-C_3k(n,h_1)\leqslant\widetilde{F}_p\{\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}},\widetilde{\boldsymbol{\gamma}}),\boldsymbol{\gamma},\boldsymbol{\beta}\}. \tag{S3.18}$$

Combining (S3.8) and (S3.18), we get

$$\widetilde{F}\{\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}},\widetilde{\boldsymbol{\gamma}}),\widetilde{\boldsymbol{\gamma}},\widetilde{\boldsymbol{\beta}}\}-C_3k(n,h_1)\leqslant\widetilde{F}_p\{\widetilde{g}_p(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})-C_1k(n,h_1),\boldsymbol{\gamma},\boldsymbol{\beta}\}+2C_2k(n,h_1).$$

That is

$$\widetilde{F}\{\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}},\widetilde{\boldsymbol{\gamma}}),\widetilde{\boldsymbol{\gamma}},\widetilde{\boldsymbol{\beta}}\} \leqslant \widetilde{F}_p\{\widetilde{g}_p(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - C_1 k(n,h_1),\boldsymbol{\gamma},\boldsymbol{\beta}\} + C_4 k(n,h_1),$$

where $C_4 = 2C_2 + C_3$. Note that $\widetilde{F}\{\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}},\widetilde{\boldsymbol{\gamma}}),\widetilde{\boldsymbol{\gamma}},\widetilde{\boldsymbol{\beta}}\} = \tau$ and $\widetilde{F}_p\{\widetilde{g}_p(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}),\boldsymbol{\gamma},\boldsymbol{\beta}\} = \tau$. Combining this and the above equation, we get

$$\widetilde{F}_p\{\widetilde{g}_p(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}),\boldsymbol{\gamma},\boldsymbol{\beta}\} \leqslant \widetilde{F}_p\{\widetilde{g}_p(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - C_1 k(n,h_1),\boldsymbol{\gamma},\boldsymbol{\beta}\} + C_4 k(n,h_1). \tag{S3.19}$$

By a first-order Taylor series expansion, we get

$$\begin{aligned} & \widetilde{F}_p\{\widetilde{g}_p(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - C_1 k(n,h_1),\boldsymbol{\gamma},\boldsymbol{\beta}\} \\ = \; & \widetilde{F}_p\{\widetilde{g}_p(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}),\boldsymbol{\gamma},\boldsymbol{\beta}\} - \widetilde{F}_p'\{\widetilde{g}^*(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}),\boldsymbol{\gamma},\boldsymbol{\beta}\} C_1 k(n,h_1), \end{aligned} \tag{S3.20}$$

where

$$\widetilde{F}_p'\{\widetilde{g}^*(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}),\boldsymbol{\gamma},\boldsymbol{\beta}\} = \frac{\sum_{j\in I_1} K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta} - \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) f_{\epsilon|\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}}\{\widetilde{g}^*(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})|\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}\}}{\sum_{j\in I_1} K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta} - \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})},$$

and $\widetilde{g}^*(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})$ is between $\widetilde{g}_p(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - C_1 k(n,h_1)$ and $\widetilde{g}_p(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})$. Combining (S3.20), using Lemma 1 and standard kernel regression analysis, we get

$$\begin{aligned} & \widetilde{F}_p\{\widetilde{g}_p(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - C_1 k(n,h_1),\boldsymbol{\gamma},\boldsymbol{\beta}\} \\ = \; & \widetilde{F}_p\{\widetilde{g}_p(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}),\boldsymbol{\gamma},\boldsymbol{\beta}\} - f_{\epsilon|\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}}(0|\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) C_1 k(n,h_1) + O_p\{k(n,h_1)\}. \end{aligned} \tag{S3.21}$$

So for any $\delta > 0$, there exists a constant $C_5 > 0$, so that with probability at least $1-\delta$, under Condition C2,

$$\begin{aligned}
& \widetilde{F}_p\{\widetilde{g}_p(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - C_1 k(n,h_1), \boldsymbol{\gamma}, \boldsymbol{\beta}\} \\
\leqslant \; & \widetilde{F}_p\{\widetilde{g}_p(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}), \boldsymbol{\gamma}, \boldsymbol{\beta}\} - 1/2 f_{\epsilon|\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}}(0|\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) C_1 k(n,h_1) + C_5 k(n,h_1).
\end{aligned} \tag{S3.22}$$

(S3.19) and (S3.22) lead to $1/2 f_{\epsilon|\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}}(0|\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) C_1 k(n,h_1) \leqslant (C_4 + C_5) k(n,h_1)$, which is impossible to hold for fixed constants C_4 and C_5 and arbitrary $C_1 > C > 0$ under Condition C2. This means

$$\mathrm{pr}\{\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}, \widetilde{\boldsymbol{\gamma}}) < \widetilde{g}_p(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - C_1 k(n,h_1)\} \to 0.$$

Similarly we can consider the case $\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}, \widetilde{\boldsymbol{\gamma}}) > \widetilde{g}_p(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) + C_1 k(n,h_1)$ and prove $\mathrm{pr}\{\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}, \widetilde{\boldsymbol{\gamma}}) > \widetilde{g}_p(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) + C_1 k(n,h_1)\} \to 0$. Thus, for arbitrary $C_1 > C > 0$, $\mathrm{pr}\{|\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}, \widetilde{\boldsymbol{\gamma}}) - \widetilde{g}_p(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})| > C_1 k(n,h_1)\} \to 0$ as $n \to \infty$. Lemma 2 is proved.

Proof of Theorem 1. Combining the results in Lemmas 1 and 2, Theorem 1 is shown.

S4 Proof of Theorem 2

We first define some notations used in the proof of Theorem 2. Let $a_n \equiv k(n,h_1)\log n$. For any possibly random $\boldsymbol{\beta}^*$ and $\boldsymbol{\gamma}^*$ satisfying $\|\boldsymbol{\beta}^* - \boldsymbol{\beta}\| \leqslant Cn^{-1/2}$, $\|\boldsymbol{\gamma}^* - \boldsymbol{\gamma}\| \leqslant Cn^{-1/2}$, let

$$\begin{aligned}
T_n(\eta, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*) \;\; &\equiv \;\; \widetilde{U}(\eta, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*) - \widetilde{U}\{g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*), \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*\} \\
& \quad - [\widetilde{U}_p(\eta, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*) - \widetilde{U}_p\{g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*), \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*\}], \\
T_n^-(\eta, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*) \;\; &\equiv \;\; \widetilde{U}^-(\eta, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*) - \widetilde{U}^-\{g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*), \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*\} \\
& \quad - [\widetilde{U}_p^-(\eta, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*) - \widetilde{U}_p^-\{g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*), \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*\}],
\end{aligned}$$

$$
\begin{aligned}
\widetilde{U}(\eta, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*) &\equiv \frac{1}{n}\sum_{j\in I_1} K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}^* - \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*) I(Y_j - \mathbf{x}_j^{\mathrm{T}}\boldsymbol{\gamma}^* - \eta < 0),\\
\widetilde{U}_p(\eta, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*) &\equiv \frac{1}{n}\sum_{j\in I_1} K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}^* - \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*) \mathrm{pr}(Y_j - \mathbf{x}_j^{\mathrm{T}}\boldsymbol{\gamma}^* - \eta < 0 | \mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}^*),\\
\widetilde{U}^+(\eta, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*) &\equiv \frac{1}{n}\sum_{j\in I_1} K_{h_1}^+(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}^* - \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*) I(Y_j - \mathbf{x}_j^{\mathrm{T}}\boldsymbol{\gamma}^* - \eta < 0),\\
\widetilde{U}^-(\eta, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*) &\equiv \frac{1}{n}\sum_{j\in I_1} K_{h_1}^-(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}^* - \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*) I(Y_j - \mathbf{x}_j^{\mathrm{T}}\boldsymbol{\gamma}^* - \eta < 0),\\
\widetilde{U}_p^+(\eta, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*) &\equiv \frac{1}{n}\sum_{j\in I_1} K_{h_1}^+(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}^* - \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*) \mathrm{pr}(Y_j - \mathbf{x}_j^{\mathrm{T}}\boldsymbol{\gamma}^* - \eta < 0 | \mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}^*),\\
\widetilde{U}_p^-(\eta, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*) &\equiv \frac{1}{n}\sum_{j\in I_1} K_{h_1}^-(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}^* - \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*) \mathrm{pr}(Y_j - \mathbf{x}_j^{\mathrm{T}}\boldsymbol{\gamma}^* - \eta < 0 | \mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}^*),
\end{aligned}
$$

where, for $j \in I_1$,

$$
K_{h_1}^+(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}^* - \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*) \equiv \begin{cases} K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}^* - \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*), K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}^* - \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*) \geqslant 0, \\ 0, K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}^* - \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*) < 0. \end{cases}
$$

$$
K_{h_1}^-(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}^* - \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*) \equiv \begin{cases} -K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}^* - \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*), K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}^* - \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*) \leqslant 0, \\ 0, K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}^* - \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*) > 0. \end{cases}
$$

It is obvious that

$$
\begin{aligned}
&\widetilde{U}(\cdot, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*) = \widetilde{U}^+(\cdot, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*) - \widetilde{U}^-(\cdot, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*),\\
&\widetilde{U}_p(\cdot, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*) = \widetilde{U}_p^+(\cdot, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*) - \widetilde{U}_p^-(\cdot, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*),\\
&T_n(\eta, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*) = \widetilde{U}^+(\eta, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*) - \widetilde{U}^-(\eta, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*) - \widetilde{U}\{g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*), \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*\}\\
&-[\widetilde{U}_p^+(\eta, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*) - \widetilde{U}_p^-(\eta, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*) - \widetilde{U}_p\{g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*), \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*\}]. \qquad \text{(S4.1)}
\end{aligned}
$$

Define $b_n \equiv \lceil 1/h_1 \rceil$. The interval $[g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*) - a_n, g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*) + a_n]$ is devided into $2b_n$ eqaual intervals:

$$J_{n,r} = [g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*) + ra_n/b_n, g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*) + (r+1)a_n/b_n] = [\eta_{n,r}, \eta_{n,r+1}],$$

$r = -b_n, \cdots, -1, 0, 1, , \cdots, b_n - 1$, each length a_n/b_n satisfies $k(n, h_1)h_1 \log n/ (1 + h_1) \leqslant a_n/b_n \leqslant k(n, h_1)h_1 \log n$. Finally, define

$$\begin{aligned}
& \widetilde{F}^*\{g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*), \boldsymbol{\gamma}^*, \boldsymbol{\beta}\} \\
\equiv\ & \frac{\sum_{j \in I_1} K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta} - \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) \mathrm{pr}\{Y_j - \mathbf{x}_j^{\mathrm{T}}\boldsymbol{\gamma}^* - g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*) < 0 | \mathbf{x}_j, \mathbf{z}_j\}}{\sum_{j \in I_1} K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta} - \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})}, \\
\Delta(\boldsymbol{\gamma}^*, \boldsymbol{\beta}^*) =\ & \widetilde{F}\{g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*), \boldsymbol{\gamma}^*, \boldsymbol{\beta}\} - \widetilde{F}\{g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}), \boldsymbol{\gamma}, \boldsymbol{\beta}\} \\
& - [\widetilde{F}^*\{g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*), \boldsymbol{\gamma}^*, \boldsymbol{\beta}\} - \widetilde{F}^*\{g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}), \boldsymbol{\gamma}, \boldsymbol{\beta}\}].
\end{aligned}$$

We now prove three lemmas which will be used to prove Theorem 2.

Lemma 3. Under the Regularity Conditions C3, C4 and C5, we have

$$T_n^{*-}(\boldsymbol{\gamma}^*, \boldsymbol{\beta}^*) \equiv \sup_{|\eta - g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*)| \leqslant a_n} |T_n^-(\eta, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*)| = O_p\{k(n, h_1)h_1 \log n\},$$

for any $\boldsymbol{\beta}^*$ and $\boldsymbol{\gamma}^*$ satisfying $\|\boldsymbol{\beta}^* - \boldsymbol{\beta}\| \leqslant Cn^{-1/2}$, $\|\boldsymbol{\gamma}^* - \boldsymbol{\gamma}\| \leqslant Cn^{-1/2}$.

Proof. By the monotonicity of $\widetilde{U}^-(\cdot, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*)$ and $\widetilde{U}_p^-(\cdot, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*)$, for $\eta \in J_{n,r} = [\eta_{n,r}, \eta_{n,r+1}]$, we get

$$\begin{aligned}
& \widetilde{U}^-\{\eta_{n,r}, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*\} - \widetilde{U}^-\{g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*), \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*\} - [\widetilde{U}_p^-\{\eta_{n,r+1}, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*\} \\
& - \widetilde{U}_p^-\{g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*), \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*\}] \leqslant T_n^-(\eta, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*) \leqslant \widetilde{U}^-\{\eta_{n,r+1}, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*\} \\
& - \widetilde{U}^-\{g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*), \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*\} - [\widetilde{U}_p^-\{\eta_{n,r}, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*\} - \widetilde{U}_p^-\{g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*), \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*\}].
\end{aligned}$$

This can be written as

$$T_n^-(\eta_{n,r},\boldsymbol{\gamma}^*,\boldsymbol{\beta}^*)-\alpha_{nr}^-\leqslant T_n^-\{\eta,\boldsymbol{\gamma}^*,\boldsymbol{\beta}^*\}\leqslant T_n^-(\eta_{n,r+1},\boldsymbol{\gamma}^*,\boldsymbol{\beta}^*)+\alpha_{nr}^-,$$

where $\alpha_{nr}^-=\widetilde{U}_p^-(\eta_{n,r+1},\boldsymbol{\gamma}^*,\boldsymbol{\beta}^*)-\widetilde{U}_p^-(\eta_{n,r},\boldsymbol{\gamma}^*,\boldsymbol{\beta}^*)$. Hence

$$T_n^{*-}(\boldsymbol{\gamma}^*,\boldsymbol{\beta}^*)\leqslant\max_{-b_n\leqslant r\leqslant b_n}|T_n^-(\eta_{n,r},\boldsymbol{\gamma}^*,\boldsymbol{\beta}^*)|+\max_{-b_n\leqslant r\leqslant b_n-1}\alpha_{nr}^-. \tag{S4.2}$$

It is easy to show that

$$\max_{-b_n\leqslant r\leqslant b_n-1}\alpha_{nr}^-=O_p\{k(n,h_1)h_1\log n\}. \tag{S4.3}$$

Let $\mathscr{A}=\sigma(\mathbf{z}_1^{\mathrm{T}}\boldsymbol{\beta}^*,\mathbf{z}_2^{\mathrm{T}}\boldsymbol{\beta}^*,\cdots,\mathbf{z}_n^{\mathrm{T}}\boldsymbol{\beta}^*)$. For notational brevity, let

$$\begin{aligned}L_{nj}^-\;\equiv\;&K_{h_1}^-(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}^*-\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*)([I(Y_j-\mathbf{x}_j^{\mathrm{T}}\boldsymbol{\gamma}^*-\eta_{n,r}<0)\\&-I\{Y_j-\mathbf{x}_j^{\mathrm{T}}\boldsymbol{\gamma}^*-g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*)<0\}]-[\mathrm{pr}(Y_j-\mathbf{x}_j^{\mathrm{T}}\boldsymbol{\gamma}^*-\eta_{n,r}<0|\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}^*)\\&-\mathrm{pr}\{Y_j-\mathbf{x}_j^{\mathrm{T}}\boldsymbol{\gamma}^*-g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*)<0|\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}^*\}]).\end{aligned}$$

Then,

$$T_n^-(\eta_{n,r},\boldsymbol{\gamma}^*,\boldsymbol{\beta}^*)=\frac{1}{n}\sum_{j\in I_1}L_{nj}^-.$$

By Bernstein inequality on page 95 of Serfling (1980), we have

$$\begin{aligned}&\mathrm{pr}\{|T_n^-(\eta_{n,r},\boldsymbol{\gamma}^*,\boldsymbol{\beta}^*)|\geqslant k(n,h_1)h_1\log n|\mathscr{A}\}\\\leqslant\;&2\exp[-\frac{\{nk(n,h_1)h_1^2\log n\}^2}{2\sum_{j\in I_1}var(h_1L_{nj}^-|\mathscr{A})+4C_1nk(n,h_1)h_1^2\log n/3}].\end{aligned} \tag{S4.4}$$

By simple derivation, we get

$$var(h_1 L_{nj}^{-}|\mathscr{A}) = E\{h_1^2 (L_{nj}^{-})^2|\mathscr{A}\} \leqslant C_2 h_1^2 \{K_{h_1}^{-}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}^* - \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*)\}^2 |\eta_{n,r} - g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*)|.$$

The above inequality holds under Condition C3. Then, under Conditions C4 and C5, it can be shown that

$$\sum_{j\in I_1} var(h_1 L_{nj}^{-}|\mathscr{A}) \leqslant C_3 n h_1 |\eta_{n,r} - g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*)| \leqslant C_3 n h_1 k(n, h_1)\mathrm{log}n.$$

Then, under Condition C4, we have

$$\begin{aligned} & \frac{\{nk(n,h_1)h_1^2\mathrm{log}n\}^2}{2\sum_{j\in I_1} var(h_1 L_{nj}^{-}|\mathscr{A}) + 4C_1 nk(n,h_1)h_1^2\mathrm{log}n/3} \\ \geqslant\ & C_4\{nk(n,h_1)h_1^2\mathrm{log}n\}^2/\{nk(n,h_1)h_1\mathrm{log}n\} = C_4 n h_1^3 k(n,h_1)\mathrm{log}n. \qquad \text{(S4.5)} \end{aligned}$$

Also we have

$$\begin{aligned} & \mathrm{pr}\{\max_{-b_n\leqslant r\leqslant b_n} |T_n^{-}(\eta_{n,r}, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*)| \geqslant k(n,h_1)h_1\mathrm{log}n\} \\ \leqslant\ & 2b_n \max_{-b_n\leqslant r\leqslant b_n} \mathrm{pr}\{|T_n^{-}(\eta_{n,r}, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*)| \geqslant k(n,h_1)h_1\mathrm{log}n\} \\ =\ & 2b_n \max_{-b_n\leqslant r\leqslant b_n} E[\mathrm{pr}\{|T_n^{-}(\eta_{n,r}, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*)| \geqslant k(n,h_1)h_1\mathrm{log}n|\mathscr{A}\}]. \qquad \text{(S4.6)} \end{aligned}$$

Then combining (S4.4), (S4.5), and (S4.6), recalling that $b_n = \lceil h_1^{-1} \rceil$, we get

$$\begin{aligned} & \mathrm{pr}\{\max_{-b_n\leqslant r\leqslant b_n} |T_n^{-}(\eta_{n,r}, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*)| \geqslant k(n,h_1)h_1\mathrm{log}n\} \\ \leqslant\ & 2(h_1^{-1}+1)\exp\{-C_4 n h_1^3 k(n,h_1)\mathrm{log}n\} \leqslant 3\exp\{-\mathrm{log}h_1 - C_4 n h_1^3 k(n,h_1)\mathrm{log}n\}, \end{aligned}$$

which goes to zero under C4. Combining this result, (S4.2) and (S4.3), Lemma 3 is proved.

Lemma 4. Under the Regularity Conditions C3, C4 and C5, we have

$$T_n^*(\boldsymbol{\gamma}^*, \boldsymbol{\beta}^*) \equiv \sup_{|\eta - g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*)| \leqslant a_n} |T_n(\eta, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*)| = O_p\{k(n, h_1)h_1 \mathrm{log} n\},$$

for any $\boldsymbol{\beta}^*$ and $\boldsymbol{\gamma}^*$ satisfying $\|\boldsymbol{\beta}^* - \boldsymbol{\beta}\| \leqslant Cn^{-1/2}$, $\|\boldsymbol{\gamma}^* - \boldsymbol{\gamma}\| \leqslant Cn^{-1/2}$.

Proof. By the monotonicity of $\widetilde{U}^+(\cdot, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*)$, $\widetilde{U}^-(\cdot, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*)$, $\widetilde{U}_p^+(\cdot, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*)$ and $\widetilde{U}_p^-(\cdot, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*)$, for $\eta \in J_{n,r} = [\eta_{n,r}, \eta_{n,r+1}]$, we have

$$\begin{aligned}
\widetilde{U}^+(\eta_{n,r}, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*) \leqslant \widetilde{U}^+(\eta, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*) \leqslant \widetilde{U}^+(\eta_{n,r+1}, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*), \\
\widetilde{U}^-(\eta_{n,r}, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*) \leqslant \widetilde{U}^-(\eta, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*) \leqslant \widetilde{U}^-(\eta_{n,r+1}, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*), \\
\widetilde{U}_p^+(\eta_{n,r}, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*) \leqslant \widetilde{U}_p^+(\eta, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*) \leqslant \widetilde{U}_p^+(\eta_{n,r+1}, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*), \\
\widetilde{U}_p^-(\eta_{n,r}, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*) \leqslant \widetilde{U}_p^-(\eta, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*) \leqslant \widetilde{U}_p^-(\eta_{n,r+1}, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*).
\end{aligned}$$

Then we have

$$\begin{aligned}
&\widetilde{U}^+(\eta_{n,r}, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*) - \widetilde{U}^-(\eta_{n,r+1}, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*) - \{\widetilde{U}_p^+(\eta_{n,r+1}, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*) \\
&-\widetilde{U}_p^-(\eta_{n,r}, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*)\} \leqslant \widetilde{U}^+(\eta_{n,r+1}, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*) - \widetilde{U}^-(\eta_{n,r}, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*) \\
&-\{\widetilde{U}_p^+(\eta_{n,r}, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*) - \widetilde{U}_p^-(\eta_{n,r+1}, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*)\}. \qquad \text{(S4.7)}
\end{aligned}$$

Note that

$$\begin{aligned}
& \widetilde{U}^+(\eta_{n,r}, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*) - \widetilde{U}^-(\eta_{n,r+1}, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*) - \{\widetilde{U}_p^+(\eta_{n,r+1}, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*) - \widetilde{U}_p^-(\eta_{n,r}, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*)\} \\
= \;& \widetilde{U}(\eta_{n,r}, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*) - \widetilde{U}_p(\eta_{n,r}, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*) - \alpha_{nr},
\end{aligned}$$

and

$$\begin{aligned}&\widetilde{U}^{+}(\eta_{n,r+1},\boldsymbol{\gamma}^*,\boldsymbol{\beta}^*)-\widetilde{U}^{-}(\eta_{n,r},\boldsymbol{\gamma}^*,\boldsymbol{\beta}^*)-\{\widetilde{U}_p^{+}(\eta_{n,r},\boldsymbol{\gamma}^*,\boldsymbol{\beta}^*)\\&-\widetilde{U}_p^{-}(\eta_{n,r+1},\boldsymbol{\gamma}^*,\boldsymbol{\beta}^*)\}=\widetilde{U}(\eta_{n,r+1},\boldsymbol{\gamma}^*,\boldsymbol{\beta}^*)-\widetilde{U}_p(\eta_{n,r+1},\boldsymbol{\gamma}^*,\boldsymbol{\beta}^*)+\alpha_{nr},\end{aligned}$$

where $\alpha_{nr}=\widetilde{U}^{-}(\eta_{n,r+1},\boldsymbol{\gamma}^*,\boldsymbol{\beta}^*)-\widetilde{U}^{-}(\eta_{n,r},\boldsymbol{\gamma}^*,\boldsymbol{\beta}^*)+\widetilde{U}_p^{+}(\eta_{n,r+1},\boldsymbol{\gamma}^*,\boldsymbol{\beta}^*)-\widetilde{U}_p^{+}(\eta_{n,r},\boldsymbol{\gamma}^*,\boldsymbol{\beta}^*)$ and $\alpha_{nr}\geqslant 0$. Combining the above results, (S4.1), (S4.7) and the definition of $T_n(\cdot,\boldsymbol{\gamma}^*,\boldsymbol{\beta}^*)$, we get

$$T_n(\eta_{n,r},\boldsymbol{\gamma}^*,\boldsymbol{\beta}^*)-\alpha_{nr}\leqslant T_n\{\eta,\boldsymbol{\gamma}^*,\boldsymbol{\beta}^*\}\leqslant T_n(\eta_{n,r+1},\boldsymbol{\gamma}^*,\boldsymbol{\beta}^*)+\alpha_{nr}.$$

Hence

$$T_n^*(\boldsymbol{\gamma}^*,\boldsymbol{\beta}^*)\leqslant\max_{-b_n\leqslant r\leqslant b_n}|T_n(\eta_{n,r},\boldsymbol{\gamma}^*,\boldsymbol{\beta}^*)|+\max_{-b_n\leqslant r\leqslant b_n-1}\alpha_{nr}. \tag{S4.8}$$

Note that α_{nr} can be written as

$$\begin{aligned}\alpha_{nr}\ &=\ T_n^{-}(\eta_{n,r+1},\boldsymbol{\gamma}^*,\boldsymbol{\beta}^*)-T_n^{-}(\eta_{n,r},\boldsymbol{\gamma}^*,\boldsymbol{\beta}^*)+\{\widetilde{U}_p^{+}(\eta_{n,r+1},\boldsymbol{\gamma}^*,\boldsymbol{\beta}^*)\\&\quad-\widetilde{U}_p^{+}(\eta_{n,r},\boldsymbol{\gamma}^*,\boldsymbol{\beta}^*)\}+\{\widetilde{U}_p^{-}(\eta_{n,r+1},\boldsymbol{\gamma}^*,\boldsymbol{\beta}^*)-\widetilde{U}_p^{-}(\eta_{n,r},\boldsymbol{\gamma}^*,\boldsymbol{\beta}^*)\}.\end{aligned}$$

Combining the above relation, Lemma 3, and Conditions C3, C4 and C5, we get

$$\max_{-b_n\leqslant r\leqslant b_n-1}\alpha_{nr}=O_p\{k(n,h_1)h_1\mathrm{log}n\}. \tag{S4.9}$$

For notational brevity, let

$$\begin{aligned} L_{nj} &\equiv K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}^* - \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*)([I(Y_j - \mathbf{x}_j^{\mathrm{T}}\boldsymbol{\gamma}^* - \eta_{n,r} < 0) \\ &\quad -I\{Y_j - \mathbf{x}_j^{\mathrm{T}}\boldsymbol{\gamma}^* - g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*) < 0\}] - [\mathrm{pr}(Y_j - \mathbf{x}_j^{\mathrm{T}}\boldsymbol{\gamma}^* - \eta_{n,r} < 0|\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}^*) \\ &\quad -\mathrm{pr}\{Y_j - \mathbf{x}_j^{\mathrm{T}}\boldsymbol{\gamma}^* - g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*) < 0|\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}^*\}]). \end{aligned}$$

Then,

$$T_n(\eta_{n,r}, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*) = \frac{1}{n}\sum_{j\in I_1} L_{nj}.$$

Similar to the analysis of $T_n^-(\eta_{n,r}, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*)$ in Lemma 3, we can get

$$\mathrm{pr}\{\max_{-b_n\leqslant r\leqslant b_n} |T_n(\eta_{n,r}, \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*)| \geqslant k(n,h_1)h_1\mathrm{log}n\} \leqslant 3\exp\{-\mathrm{log}h_1 - C_4 n h_1^3 k(n,h_1)\mathrm{log}n\},$$

which goes to zero under C4. Combining this result, (S4.8) and (S4.9), Lemma 4 is proved.

<u>**Lemma**</u> **5.** Under the Regularity Conditions C1, C3, C4 and C5,

$$\sup_{\substack{\|\boldsymbol{\beta}^*-\boldsymbol{\beta}\|\leqslant Cn^{-1/2} \\ \|\boldsymbol{\gamma}^*-\boldsymbol{\gamma}\|\leqslant Cn^{-1/2}}} |\Delta(\boldsymbol{\gamma}^*, \boldsymbol{\beta}^*)| = o_p(n^{-1/2}).$$

Proof. For notational brevity, let

$$\begin{aligned} & s_j(\boldsymbol{\gamma}^*, \boldsymbol{\beta}^*, \boldsymbol{\gamma}, \boldsymbol{\beta}, \mathbf{z}_i) \\ = \; & K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta} - \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})([I\{Y_j - \mathbf{x}_j^{\mathrm{T}}\boldsymbol{\gamma}^* - g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*) < 0\} \\ & -I\{Y_j - \mathbf{x}_j^{\mathrm{T}}\boldsymbol{\gamma} - g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) < 0\}] - [\mathrm{pr}\{Y_j - \mathbf{x}_j^{\mathrm{T}}\boldsymbol{\gamma}^* - g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*) < 0|\mathbf{x}_j, \mathbf{z}_j\} \\ & -\mathrm{pr}\{Y_j - \mathbf{x}_j^{\mathrm{T}}\boldsymbol{\gamma} - g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) < 0|\mathbf{x}_j, \mathbf{z}_j\}]) \end{aligned}$$

and $\delta_j' = \mathbf{x}_j^{\mathrm{T}}(\boldsymbol{\gamma}^* - \boldsymbol{\gamma}) + \{g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*) - g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})\}$.

$$
\begin{aligned}
& E[\sup_{\mathscr{D}} |s_j(\boldsymbol{\gamma}^*, \boldsymbol{\beta}^*, \boldsymbol{\gamma}, \boldsymbol{\beta}, \mathbf{z}_i)|^2] \\
= \;& E[K_{h_1}^2(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta} - \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) \sup_{\mathscr{D}} |I\{\epsilon_j < \delta_j' + g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})\} \\
& -I\{\epsilon_j < g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})\} - F_\epsilon\{\delta_j' + g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}), \mathbf{x}_j, \mathbf{z}_j\} \\
& +F_\epsilon\{g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}), \mathbf{x}_j, \mathbf{z}_j\}|^2] \\
\leqslant \;& Ch_1^{-2} E[\sup_{\mathscr{D}} |I\{\epsilon_j < |\delta_j'| + g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})\} \\
& -I\{\epsilon_j < -|\delta_j'| + g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})\} - F_\epsilon\{-|\delta_j'| + g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}), \mathbf{x}_j, \mathbf{z}_j\} \\
& +F_\epsilon\{|\delta_j'| + g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}), \mathbf{x}_j, \mathbf{z}_j\}|^2] \\
\leqslant \;& 2Ch_1^{-2} E\{\sup_{\mathscr{D}}([I\{\epsilon_j < |\delta_j'| + g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})\} \\
& -I\{\epsilon_j < -|\delta_j'| + g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})\}]^2 + [F_\epsilon\{|\delta_j'| + g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}), \mathbf{x}_j, \mathbf{z}_j\} \\
& -F_\epsilon\{-|\delta_j'| + g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}), \mathbf{x}_j, \mathbf{z}_j\}]^2)\} \\
= \;& 2Ch_1^{-2} E\{\sup_{\mathscr{D}}([I\{\epsilon_j < |\delta_j'| + g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})\} \\
& -I\{\epsilon_j < -|\delta_j'| + g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})\}] + [F_\epsilon\{|\delta_j'| + g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}), \mathbf{x}_j, \mathbf{z}_j\} \\
& -F_\epsilon\{-|\delta_j'| + g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}), \mathbf{x}_j, \mathbf{z}_j\}]^2)\} \\
= \;& 2Ch_1^{-2} E\{\sup_{\mathscr{D}}(F_\epsilon\{|\delta_j'| + g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}), \mathbf{x}_j, \mathbf{z}_j\} \\
& -F_\epsilon\{-|\delta_j'| + g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}), \mathbf{x}_j, \mathbf{z}_j\} + [F_\epsilon\{|\delta_j'| + g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}), \mathbf{x}_j, \mathbf{z}_j\} \\
& -F_\epsilon\{-|\delta_j'| + g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}), \mathbf{x}_j, \mathbf{z}_j\}]^2)\} \\
\leqslant \;& 6Ch_1^{-2} E\Bigg(\sup_{\mathscr{D}} \Bigg[F_\epsilon\{|\delta_j'| + g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}), \mathbf{x}_j, \mathbf{z}_j\} \\
& -F_\epsilon\{-|\delta_j'| + g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}), \mathbf{x}_j, \mathbf{z}_j\}\Bigg]\Bigg) \\
\leqslant \;& C' h_1^{-2} E(\sup_{\mathscr{D}} |\delta_j'|)
\end{aligned}
$$

$$\leqslant \quad C' h_1^{-2} E[\sup_{\mathscr{D}} \|\mathbf{x}_j\| \|\boldsymbol{\gamma}^* - \boldsymbol{\gamma}\| + |g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*) - g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})|] = O_p\{(nh_1^4)^{-1/2}\}. \tag{S4.10}$$

In the above derivation, the first inequality uses the monotonicity of indicator function and $F_\epsilon(\cdot, \mathbf{x}_j, \mathbf{z}_j)$, as well as Condition C5. The fourth inequality uses Condition C3. The last equality uses Conditions C1 and C9. Note that $E\{\sup_{\mathscr{D}} |s_j(\boldsymbol{\gamma}^*, \boldsymbol{\beta}^*, \boldsymbol{\gamma}, \boldsymbol{\beta}, \mathbf{z}_i)|^2\}$ actually does not depend on j, and, under Condition C4, we can get

$$\max_{j \in I_1} E\{\sup_{\mathscr{D}} |s_j(\boldsymbol{\gamma}^*, \boldsymbol{\beta}^*, \boldsymbol{\gamma}, \boldsymbol{\beta}, \mathbf{z}_i)|^2\} = E\{\sup_{\mathscr{D}} |s_j(\boldsymbol{\gamma}^*, \boldsymbol{\beta}^*, \boldsymbol{\gamma}, \boldsymbol{\beta}, \mathbf{z}_i)|^2\} \leqslant C, \tag{S4.11}$$

$$\begin{aligned} & \sum_{j \in I_1} E\{\sup_{\mathscr{D}} |s_j(\boldsymbol{\gamma}^*, \boldsymbol{\beta}^*, \boldsymbol{\gamma}, \boldsymbol{\beta}, \mathbf{z}_i)|^2\} \\ = \quad & nE\{\sup_{\mathscr{D}} |s_j(\boldsymbol{\gamma}^*, \boldsymbol{\beta}^*, \boldsymbol{\gamma}, \boldsymbol{\beta}, \mathbf{z}_i)|^2\} = O_p\{n(nh_1^4)^{-1/2}\}, \end{aligned} \tag{S4.12}$$

where C is a constant. By Markov's inequality and (S4.10), we get

$$\sup_{\mathscr{D}} |s_j(\boldsymbol{\gamma}^*, \boldsymbol{\beta}^*, \boldsymbol{\gamma}, \boldsymbol{\beta}, \mathbf{z}_i)|^2 = O_p\{(nh_1^4)^{-1/2}\}.$$

Then we get

$$\sum_{j \in I_1} \sup_{\mathscr{D}} |s_j(\boldsymbol{\gamma}^*, \boldsymbol{\beta}^*, \boldsymbol{\gamma}, \boldsymbol{\beta}, \mathbf{z}_i)|^2 = n \sup_{\mathscr{D}} |s_j(\boldsymbol{\gamma}^*, \boldsymbol{\beta}^*, \boldsymbol{\gamma}, \boldsymbol{\beta}, \mathbf{z}_i)|^2 = O_p\{n(nh_1^4)^{-1/2}\}. \tag{S4.13}$$

Based on (S4.11), (S4.12) , (S4.13) and Lemma 3.3 of He and Shao (2000), we get

$$\sup_{\mathscr{D}} |\sum_{j \in I_1} s_j(\boldsymbol{\gamma}^*, \boldsymbol{\beta}^*, \boldsymbol{\gamma}, \boldsymbol{\beta}, \mathbf{z}_i)| = O_p\{n^{1/2}(nh_1^4)^{-1/4}(\mathrm{log} n)^{1/2}\}.$$

By Condition C4, Lemma 5 is shown.

<u>Proof of Theorem 2</u>. By Theorem 1, Lemma 4 and the convergence rate of $\widetilde{\boldsymbol{\beta}}$ and $\widetilde{\boldsymbol{\gamma}}$, we have

$$T_n\{\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol\beta},\widetilde{\boldsymbol\gamma}),\widetilde{\boldsymbol\gamma},\widetilde{\boldsymbol\beta}\} = O_p\{k(n,h_1)h_1\mathrm{log}n\}.$$

So under Conditions C2, C4, C5 and the consistency of $\widetilde{\boldsymbol\beta}$, we have

$$\begin{aligned}
& \widetilde{F}\{\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol\beta},\widetilde{\boldsymbol\gamma}),\widetilde{\boldsymbol\gamma},\widetilde{\boldsymbol\beta}\} - \widetilde{F}\{g(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol\beta}),\widetilde{\boldsymbol\gamma},\widetilde{\boldsymbol\beta}\} - [\widetilde{F}_p\{\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol\beta},\widetilde{\boldsymbol\gamma}),\widetilde{\boldsymbol\gamma},\widetilde{\boldsymbol\beta}\} - \widetilde{F}_p\{g(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol\beta}),\widetilde{\boldsymbol\gamma},\widetilde{\boldsymbol\beta}\}] \\
= \;& T_n\{\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol\beta},\widetilde{\boldsymbol\gamma}),\widetilde{\boldsymbol\gamma},\widetilde{\boldsymbol\beta}\}/\{\sum_{j\in I_1} K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\widetilde{\boldsymbol\beta} - \mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol\beta})/n\} = O_p\{k(n,h_1)h_1\mathrm{log}n\},
\end{aligned}$$

where

$$\begin{aligned}
& \widetilde{F}\{\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol\beta},\widetilde{\boldsymbol\gamma}),\widetilde{\boldsymbol\gamma},\widetilde{\boldsymbol\beta}\} \\
= \;& \frac{\sum_{j\in I_1} K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\widetilde{\boldsymbol\beta} - \mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol\beta}) I\{Y_j - \mathbf{x}_j^{\mathrm{T}}\widetilde{\boldsymbol\gamma} - \widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol\beta},\widetilde{\boldsymbol\gamma}) < 0\}}{\sum_{j\in I_1} K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\widetilde{\boldsymbol\beta} - \mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol\beta})} = \tau, \\
& \widetilde{F}\{g(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol\beta}),\widetilde{\boldsymbol\gamma},\widetilde{\boldsymbol\beta}\} = \frac{\sum_{j\in I_1} K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\widetilde{\boldsymbol\beta} - \mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol\beta}) I\{Y_j - \mathbf{x}_j^{\mathrm{T}}\widetilde{\boldsymbol\gamma} - g(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol\beta}) < 0\}}{\sum_{j\in I_1} K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\widetilde{\boldsymbol\beta} - \mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol\beta})}, \\
& \widetilde{F}_p\{\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol\beta},\widetilde{\boldsymbol\gamma}),\widetilde{\boldsymbol\gamma},\widetilde{\boldsymbol\beta}\} \\
= \;& \frac{\sum_{j\in I_1} K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\widetilde{\boldsymbol\beta} - \mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol\beta}) \mathrm{pr}\{Y_j - \mathbf{x}_j^{\mathrm{T}}\widetilde{\boldsymbol\gamma} - \widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol\beta},\widetilde{\boldsymbol\gamma}) < 0 | \mathbf{z}_j^{\mathrm{T}}\widetilde{\boldsymbol\beta}\}}{\sum_{j\in I_1} K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\widetilde{\boldsymbol\beta} - \mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol\beta})}, \\
& \widetilde{F}_p\{g(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol\beta}),\widetilde{\boldsymbol\gamma},\widetilde{\boldsymbol\beta}\} \\
= \;& \frac{\sum_{j\in I_1} K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\widetilde{\boldsymbol\beta} - \mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol\beta}) \mathrm{pr}\{Y_j - \mathbf{x}_j^{\mathrm{T}}\widetilde{\boldsymbol\gamma} - g(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol\beta}) < 0 | \mathbf{z}_j^{\mathrm{T}}\widetilde{\boldsymbol\beta}\}}{\sum_{j\in I_1} K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\widetilde{\boldsymbol\beta} - \mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol\beta})}.
\end{aligned}$$

Because

$$\begin{aligned}
& \widetilde{F}_p\{\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol\beta},\widetilde{\boldsymbol\gamma}),\widetilde{\boldsymbol\gamma},\widetilde{\boldsymbol\beta}\} - \widetilde{F}_p\{g(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol\beta}),\widetilde{\boldsymbol\gamma},\widetilde{\boldsymbol\beta}\} \\
= \;& f_{\epsilon|\mathbf{z}^{\mathrm{T}}\boldsymbol\beta}(0,\mathbf{z}_i^{\mathrm{T}}\boldsymbol\beta)\{\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol\beta},\widetilde{\boldsymbol\gamma}) - g(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol\beta})\} + O_p\{k(n,h_1)^2\},
\end{aligned}$$

hence

$$\begin{aligned}
& \widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}, \widetilde{\boldsymbol{\gamma}}) - g(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}) \\
= \ & \{f_{\epsilon|\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}}(0, \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})\}^{-1}[\widetilde{F}_p\{\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}, \widetilde{\boldsymbol{\gamma}}), \widetilde{\boldsymbol{\gamma}}, \widetilde{\boldsymbol{\beta}}\} - \widetilde{F}_p\{g(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}), \widetilde{\boldsymbol{\gamma}}, \widetilde{\boldsymbol{\beta}}\}] + O_p\{k(n, h_1)^2\} \\
= \ & \{f_{\epsilon|\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}}(0, \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})\}^{-1}[\widetilde{F}\{\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}, \widetilde{\boldsymbol{\gamma}}), \widetilde{\boldsymbol{\gamma}}, \widetilde{\boldsymbol{\beta}}\} - \widetilde{F}\{g(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}), \widetilde{\boldsymbol{\gamma}}, \widetilde{\boldsymbol{\beta}}\} \\
& -T_n\{\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}, \widetilde{\boldsymbol{\gamma}}), \widetilde{\boldsymbol{\gamma}}, \widetilde{\boldsymbol{\beta}}\}/\{\sum_{j\in I_1} K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\widetilde{\boldsymbol{\beta}} - \mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}})/n\}] + O_p\{k(n, h_1)^2\} \\
= \ & \{f_{\epsilon|\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}}(0, \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})\}^{-1}[\tau - \widetilde{F}\{g(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}), \widetilde{\boldsymbol{\gamma}}, \widetilde{\boldsymbol{\beta}}\}] + O_p\{k(n, h_1)h_1 \mathrm{log} n\}. \qquad \text{(S4.14)}
\end{aligned}$$

The last equality holds under Condition C4.

By direct calculation, we have

$$\begin{aligned}
& \widetilde{F}\{g(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}), \widetilde{\boldsymbol{\gamma}}, \widetilde{\boldsymbol{\beta}}\} \\
= \ & \frac{\sum_{j\in I_1} K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\widetilde{\boldsymbol{\beta}} - \mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}) I\{Y_j - \mathbf{x}_j^{\mathrm{T}}\widetilde{\boldsymbol{\gamma}} - g(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}) < 0\}}{\sum_{j\in I_1} K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta} - \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})} \\
& + \frac{\sum_{j\in I_1} K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\widetilde{\boldsymbol{\beta}} - \mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}) I\{Y_j - \mathbf{x}_j^{\mathrm{T}}\widetilde{\boldsymbol{\gamma}} - g(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}) < 0\}}{\sum_{j\in I_1} K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta} - \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})} \\
& \times \frac{\sum_{j\in I_1}\{K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta} - \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\widetilde{\boldsymbol{\beta}} - \mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}})\}}{\sum_{j\in I_1} K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\widetilde{\boldsymbol{\beta}} - \mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}})}.
\end{aligned}$$

Thus,

$$\begin{aligned}
& \widetilde{F}\{g(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}), \widetilde{\boldsymbol{\gamma}}, \widetilde{\boldsymbol{\beta}}\} - \widetilde{F}\{g(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}), \widetilde{\boldsymbol{\gamma}}, \boldsymbol{\beta}\} \\
= \ & \frac{\sum_{j\in I_1}\{K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\widetilde{\boldsymbol{\beta}} - \mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}) - K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta} - \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})\} I\{Y_j - \mathbf{x}_j^{\mathrm{T}}\widetilde{\boldsymbol{\gamma}} - g(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}) < 0\}}{\sum_{j\in I_1} K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta} - \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})} \\
& + \frac{\sum_{j\in I_1} K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\widetilde{\boldsymbol{\beta}} - \mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}) I\{Y_j - \mathbf{x}_j^{\mathrm{T}}\widetilde{\boldsymbol{\gamma}} - g(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}) < 0\}}{\sum_{j\in I_1} K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\widetilde{\boldsymbol{\beta}} - \mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}})} \\
& \times \frac{\sum_{j\in I_1}\{K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta} - \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\widetilde{\boldsymbol{\beta}} - \mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}})\}}{\sum_{j\in I_1} K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta} - \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})}
\end{aligned}$$

$$= \quad A_1 + A_2. \tag{S4.15}$$

Similar to the derivation of (S3.12) and (S3.13) from (S3.9), we get

$$\begin{aligned}
A_1 &= \frac{\mathrm{pr}\{Y_i - \mathbf{x}_i^{\mathrm{T}}\widetilde{\boldsymbol{\gamma}} - g(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}) < 0|\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}\}f(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}) - \mathrm{pr}\{Y_i - \mathbf{x}_i^{\mathrm{T}}\widetilde{\boldsymbol{\gamma}} - g(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}) < 0|\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}\}f(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})}{f(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})} \\
&\quad + o_p(n^{-1/2}) + O_p(h_1^4), \\
A_2 &= \frac{\mathrm{pr}\{Y_i - \mathbf{x}_i^{\mathrm{T}}\widetilde{\boldsymbol{\gamma}} - g(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}) < 0|\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}\}\{f(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - f(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}})\}}{f(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})} + o_p(n^{-1/2}) + O_p(h_1^4) \\
&= \frac{\mathrm{pr}\{Y_i - \mathbf{x}_i^{\mathrm{T}}\widetilde{\boldsymbol{\gamma}} - g(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}) < 0|\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}\}\{f(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - f(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}})\}}{f(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})} + o_p(n^{-1/2}) + O_p(h_1^4).
\end{aligned}$$

Then, under Condition C4, we have

$$\begin{aligned}
& A_1 + A_2 \\
&= \mathrm{pr}\{Y_i - \mathbf{x}_i^{\mathrm{T}}\widetilde{\boldsymbol{\gamma}} - g(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}) < 0|\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}\} - \mathrm{pr}\{Y_i - \mathbf{x}_i^{\mathrm{T}}\widetilde{\boldsymbol{\gamma}} - g(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}) < 0|\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}\} + o_p(n^{-1/2}).
\end{aligned} \tag{S4.16}$$

Note that for any $i \in I_2$,

$$\begin{aligned}
& \mathrm{pr}\{Y_i - \mathbf{x}_i^{\mathrm{T}}\widetilde{\boldsymbol{\gamma}} - g(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}) < 0|\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}\} \\
&= \tau + E\{f_\epsilon(0, \mathbf{x}_i, \mathbf{z}_i)\mathbf{x}_i^{\mathrm{T}}|\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}\}(\widetilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma}) + g'(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})E(\mathbf{z}_{-1,i}^{\mathrm{T}}|\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})(\widetilde{\boldsymbol{\beta}}_{-1} - \boldsymbol{\beta}_{-1}) + o_p(n^{-1/2}),
\end{aligned}$$

and

$$\begin{aligned}
& \mathrm{pr}\{Y_i - \mathbf{x}_i^{\mathrm{T}}\widetilde{\boldsymbol{\gamma}} - g(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}) < 0|\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}\} \\
&= \tau + E\{f_\epsilon(0, \mathbf{x}_i, \mathbf{z}_i)\mathbf{x}_i^{\mathrm{T}}|\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}\}(\widetilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma}) + g'(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})E(\mathbf{z}_{-1,i}^{\mathrm{T}}|\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})(\widetilde{\boldsymbol{\beta}}_{-1} - \boldsymbol{\beta}_{-1}) + o_p(n^{-1/2}).
\end{aligned}$$

Thus

$$\mathrm{pr}\{Y_i - \mathbf{x}_i^{\mathrm{T}}\widetilde{\boldsymbol{\gamma}} - g(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}) < 0|\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}\} - \mathrm{pr}\{Y_i - \mathbf{x}_i^{\mathrm{T}}\widetilde{\boldsymbol{\gamma}} - g(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}) < 0|\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}\} = o_p(n^{-1/2}). \quad \text{(S4.17)}$$

Combining (S4.15), (S4.16) and (S4.17), we get

$$\widetilde{F}\{g(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}), \widetilde{\boldsymbol{\gamma}}, \widetilde{\boldsymbol{\beta}}\} - \widetilde{F}\{g(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}), \widetilde{\boldsymbol{\gamma}}, \boldsymbol{\beta}\} = o_p(n^{-1/2}). \quad \text{(S4.18)}$$

By (S4.14) and (S4.18), we get

$$\begin{aligned} \widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}, \widetilde{\boldsymbol{\gamma}}) - g(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}) &= \{f_{\epsilon|\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}}(0, \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})\}^{-1}[\tau - \widetilde{F}\{g(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}), \widetilde{\boldsymbol{\gamma}}, \boldsymbol{\beta}\}] \\ &\quad + o_p(n^{-1/2}) + O_p\{k(n, h_1)h_1 \log n\}, \quad \text{(S4.19)} \end{aligned}$$

Combining $\Delta(\widetilde{\boldsymbol{\gamma}}, \widetilde{\boldsymbol{\beta}}) = o_p(n^{-1/2})$ from Lemma 5 and (S4.19), we obtain

$$\begin{aligned} &\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}, \widetilde{\boldsymbol{\gamma}}) - g(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}) \\ = \ &\{f_{\epsilon|\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}}(0, \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})\}^{-1}[\tau - \widetilde{F}\{g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}), \boldsymbol{\gamma}, \boldsymbol{\beta}\}] \\ &-\{f_{\epsilon|\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}}(0, \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})\}^{-1}[\widetilde{F}^*\{g(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}), \widetilde{\boldsymbol{\gamma}}, \boldsymbol{\beta}\} - \widetilde{F}^*\{g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}), \boldsymbol{\gamma}, \boldsymbol{\beta}\}] + o_p(n^{-1/2}). \quad \text{(S4.20)} \end{aligned}$$

Also we can show

$$\begin{aligned} &\widetilde{F}^*\{g(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}), \widetilde{\boldsymbol{\gamma}}, \boldsymbol{\beta}\} \\ = \ &\sum_{j \in I_1} K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta} - \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) \mathrm{pr}[\epsilon_j < \mathbf{x}_j^{\mathrm{T}}(\widetilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma}) + \{g(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}) - g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})\} \\ &+\{g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})\}|\mathbf{x}_j, \mathbf{z}_j] / \sum_{j \in I_1} K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta} - \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}), \end{aligned}$$

and

$$\begin{aligned}\widetilde{F}^*\{g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}),\boldsymbol{\gamma},\boldsymbol{\beta}\} &= \sum_{j\in I_1} K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}-\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})\mathrm{pr}\{\epsilon_j < g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})-g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})|\mathbf{x}_j,\mathbf{z}_j\}] \\ &\quad /\sum_{j\in I_1} K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}-\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}).\end{aligned}$$

Then combining the above results, under Conditions C1 and C2, we have

$$\begin{aligned}&\widetilde{F}^*\{g(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}),\widetilde{\boldsymbol{\gamma}},\boldsymbol{\beta}\}-\widetilde{F}^*\{g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}),\boldsymbol{\gamma},\boldsymbol{\beta}\} \\ =\ & E\{f_\epsilon(0,\mathbf{x}_i,\mathbf{z}_i)\mathbf{x}_i^{\mathrm{T}}|\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}\}(\widetilde{\boldsymbol{\gamma}}-\boldsymbol{\gamma})+g'(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})E\{f_\epsilon(0,\mathbf{x}_i,\mathbf{z}_i)\mathbf{z}_{-1,i}^{\mathrm{T}}|\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}\} \\ & \times(\widetilde{\boldsymbol{\beta}}_{-1}-\boldsymbol{\beta}_{-1})+o_p(n^{-1/2}).\end{aligned} \tag{S4.21}$$

Combining (S4.20) and (S4.21), we get

$$\begin{aligned}&\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}},\widetilde{\boldsymbol{\gamma}})-g(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}) \\ =\ & \{f_{\epsilon|\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}}(0,\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})\}^{-1}[\tau-\widetilde{F}\{g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}),\boldsymbol{\gamma},\boldsymbol{\beta}\}] \\ & -\{f_{\epsilon|\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}}(0,\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})\}^{-1}E\{f_\epsilon(0,\mathbf{x}_i,\mathbf{z}_i)\mathbf{x}_i^{\mathrm{T}}|\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}\}(\widetilde{\boldsymbol{\gamma}}-\boldsymbol{\gamma}) \\ & -\{f_{\epsilon|\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}}(0,\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})\}^{-1}g'(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})E\{f_\epsilon(0,\mathbf{x}_i,\mathbf{z}_i)\mathbf{z}_{-1,i}^{\mathrm{T}}|\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}\}(\widetilde{\boldsymbol{\beta}}_{-1}-\boldsymbol{\beta}_{-1})+o_p(n^{-1/2}).\end{aligned}$$

Theorem 2 is proved.

S5 Partial Results for the Proof of Theorem 3

We first state a result regarding $g'(\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta})$.

Proposition 2. Under the Regularity Conditions C1, C3, C4, C5-C8, we have

$$\widetilde{g}'(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}, \widetilde{\boldsymbol{\gamma}}) - g'(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) = o_p(1),$$

for all $\mathbf{z}_i \in I_2$.

The proof of Proposition 2 is similar to that of Theorem 2, so we omit it here.

The following five lemmas are used to prove Lemma 11.

Lemma 6. Under the Regularity Conditions C1, C3 and C4,

$$\sup_{\mathscr{D}} \|\frac{1}{m}\sum_{i\in I_2} \eta_{mi}\{\boldsymbol{\gamma}, \boldsymbol{\beta}, g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}), \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*, \widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*, \boldsymbol{\gamma}^*)\}\| = o_p(m^{-1/2}).$$

Proof. For convenience, denote $\delta_i \equiv \mathbf{x}_i^{\mathrm{T}}(\boldsymbol{\gamma}^* - \boldsymbol{\gamma}) + \{\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*, \boldsymbol{\gamma}^*) - g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})\}$. First, by the definition of $\eta_{mi}(\cdot)$ and $G_i(\cdot)$,

$$\begin{aligned}
& E[\sup_{\mathscr{D}} \|\eta_{mi}\{\boldsymbol{\gamma}, \boldsymbol{\beta}, g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}), \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*, \widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*, \gamma^*)\}\|^2] \\
= \; & E[P(\mathbf{x}_i, \mathbf{z}_i)^{\mathrm{T}} P(\mathbf{x}_i, \mathbf{z}_i) \frac{f_\epsilon^{*2}(0, \mathbf{x}_i, \mathbf{z}_i)}{\tau^2(1-\tau)^2} \sup_{\mathscr{D}} |I\{Y_i - \mathbf{x}_i^{\mathrm{T}}\boldsymbol{\gamma}^* - \widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*, \boldsymbol{\gamma}^*) < 0\} \\
& - I\{Y_i - \mathbf{x}_i^{\mathrm{T}}\boldsymbol{\gamma} - g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) < 0\} \\
& - F_\epsilon([\mathbf{x}_i^{\mathrm{T}}(\boldsymbol{\gamma}^* - \boldsymbol{\gamma}) + \{\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*, \boldsymbol{\gamma}^*) - g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})\}], \mathbf{x}_i, \mathbf{z}_i) + \tau|^2] \\
= \; & E[P(\mathbf{x}_i, \mathbf{z}_i)^{\mathrm{T}} P(\mathbf{x}_i, \mathbf{z}_i) \frac{f_\epsilon^{*2}(0, \mathbf{x}_i, \mathbf{z}_i)}{\tau^2(1-\tau)^2} \\
& \times \sup_{\mathscr{D}} |I\{\epsilon_i + \mathbf{x}_i^{\mathrm{T}}\boldsymbol{\gamma} + g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - \mathbf{x}_i^{\mathrm{T}}\boldsymbol{\gamma}^* - \widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*, \boldsymbol{\gamma}^*) < 0\} - I(\epsilon_i < 0) \\
& - F_\epsilon([\mathbf{x}_i^{\mathrm{T}}(\boldsymbol{\gamma}^* - \boldsymbol{\gamma}) + \{\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*, \boldsymbol{\gamma}^*) - g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})\}], \mathbf{x}_i, \mathbf{z}_i) + F_\epsilon(0, \mathbf{x}_i, \mathbf{z}_i)|^2] \\
= \; & E[P(\mathbf{x}_i, \mathbf{z}_i)^{\mathrm{T}} P(\mathbf{x}_i, \mathbf{z}_i) \frac{f_\epsilon^{*2}(0, \mathbf{x}_i, \mathbf{z}_i)}{\tau^2(1-\tau)^2} \sup_{\mathscr{D}} |I(\epsilon_i < \delta_i) - I(\epsilon_i < 0) \\
& - F_\epsilon(\delta_i, \mathbf{x}_i, \mathbf{z}_i) + F_\epsilon(0, \mathbf{x}_i, \mathbf{z}_i)|^2]
\end{aligned}$$

$$
\begin{aligned}
&\leqslant CE\{\sup_{\mathscr{D}}|I(\epsilon_i<|\delta_i|)-I(\epsilon_i<-|\delta_i|)-F_\epsilon(-|\delta_i|,\mathbf{x}_i,\mathbf{z}_i)+F_\epsilon(|\delta_i|,\mathbf{x}_i,\mathbf{z}_i)|^2\}\\
&\leqslant 2CE\left(\sup_{\mathscr{D}}\left[\{I(\epsilon_i<|\delta_i|)-I(\epsilon_i<-|\delta_i|)\}^2+\{F_\epsilon(|\delta_i|,\mathbf{x}_i,\mathbf{z}_i)-F_\epsilon(-|\delta_i|,\mathbf{x}_i,\mathbf{z}_i)\}^2\right]\right)\\
&= 2CE\left(\sup_{\mathscr{D}}\left[I(\epsilon_i<|\delta_i|)-I(\epsilon_i<-|\delta_i|)+\{F_\epsilon(|\delta_i|,\mathbf{x}_i,\mathbf{z}_i)-F_\epsilon(-|\delta_i|,\mathbf{x}_i,\mathbf{z}_i)\}^2\right]\right)\\
&= 2CE\left(\sup_{\mathscr{D}}\left[F_\epsilon(|\delta_i|,\mathbf{x}_i,\mathbf{z}_i)-F_\epsilon(-|\delta_i|,\mathbf{x}_i,\mathbf{z}_i)+\{F_\epsilon(|\delta_i|,\mathbf{x}_i,\mathbf{z}_i)-F_\epsilon(-|\delta_i|,\mathbf{x}_i,\mathbf{z}_i)\}^2\right]\right)\\
&\leqslant 6CE\left[\sup_{\mathscr{D}}\{F_\epsilon(|\delta_i|,\mathbf{x}_i,\mathbf{z}_i)-F_\epsilon(-|\delta_i|,\mathbf{x}_i,\mathbf{z}_i)\}\right]\\
&\leqslant C'E(\sup_{\mathscr{D}}|\delta_i|)\\
&\leqslant C'E[\sup_{\mathscr{D}}\|\mathbf{x}_i\|\|\boldsymbol{\gamma}^*-\boldsymbol{\gamma}\|+|\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*,\boldsymbol{\gamma}^*)-g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})|]\\
&= O_p\{(nh_1)^{-1/2}+h_1^4\}. \qquad \text{(S5.22)}
\end{aligned}
$$

In the above derivation, the first inequality uses the monotonicity of indicator function and $F_\epsilon(\cdot,\mathbf{x}_i,\mathbf{z}_i)$, as well as Condition C1 and the boundedness of the proposed $f_\epsilon^*(0,\mathbf{x}_i,\mathbf{z}_i)$. The fourth inequality uses Condition C3. The last equality uses Conditions C1, C4 and Theorem 1. Note that $E[\sup_{\mathscr{D}}\|\eta_{mi}\{\boldsymbol{\gamma},\boldsymbol{\beta},g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}),\boldsymbol{\gamma}^*,\boldsymbol{\beta}^*,\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*,\gamma^*)\}\|^2]$ actually does not depend on i, and, under Condition C4, we can get

$$
\begin{aligned}
&\max_{i\in I_2}E[\sup_{\mathscr{D}}\|\eta_{mi}\{\boldsymbol{\gamma},\boldsymbol{\beta},g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}),\boldsymbol{\gamma}^*,\boldsymbol{\beta}^*,\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*,\gamma^*)\}\|^2]\\
=\ &E[\sup_{\mathscr{D}}\|\eta_{mi}\{\boldsymbol{\gamma},\boldsymbol{\beta},g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}),\boldsymbol{\gamma}^*,\boldsymbol{\beta}^*,\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*,\gamma^*)\}\|^2]\leqslant C, \qquad \text{(S5.23)}\\
&\sum_{i\in I_2}E[\sup_{\mathscr{D}}\|\eta_{mi}\{\boldsymbol{\gamma},\boldsymbol{\beta},g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}),\boldsymbol{\gamma}^*,\boldsymbol{\beta}^*,\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*,\boldsymbol{\gamma}^*)\}\|^2]\\
=\ &mE[\sup_{\mathscr{D}}\|\eta_{mi}\{\boldsymbol{\gamma},\boldsymbol{\beta},g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}),\boldsymbol{\gamma}^*,\boldsymbol{\beta}^*,\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*,\boldsymbol{\gamma}^*)\}\|^2]\\
=\ &O_p[m\{h_1^4+(nh_1)^{-1/2}\}], \qquad \text{(S5.24)}
\end{aligned}
$$

where C is a constant. By Markov's inequality and (S5.22), we get

$$\sup_{\mathscr{D}} \|\eta_{ni}\{\boldsymbol{\gamma}, \boldsymbol{\beta}, g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}), \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*, \widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*, \boldsymbol{\gamma}^*)\}\|^2 = O_p\{h_1^4 + (nh_1)^{-1/2}\}.$$

Then we have

$$\begin{aligned}
& \sum_{i\in I_2} \sup_{\mathscr{D}} \|\eta_{ni}\{\boldsymbol{\gamma}, \boldsymbol{\beta}, g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}), \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*, \widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*, \boldsymbol{\gamma}^*)\}\|^2 \\
= \; & m \sup_{\mathscr{D}} \|\eta_{ni}\{\boldsymbol{\gamma}, \boldsymbol{\beta}, g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}), \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*, \widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*, \boldsymbol{\gamma}^*)\}\|^2 \\
= \; & O_p[m\{h_1^4 + (nh_1)^{-1/2}\}]. \qquad \text{(S5.25)}
\end{aligned}$$

Based on (S5.23), (S5.24), (S5.25) and Lemma 3.3 of He and Shao (2000), we get

$$\begin{aligned}
& \sup_{\mathscr{D}} \|\sum_{i\in I_2} \eta_{mi}\{\boldsymbol{\gamma}, \boldsymbol{\beta}, g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}), \boldsymbol{\gamma}^*, \boldsymbol{\beta}^*, \widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}^*, \boldsymbol{\gamma}^*)\}\| \\
= \; & O_p[m^{1/2}\{h_1^2 + (nh_1)^{-1/4}\}(\log m)^{1/2}].
\end{aligned}$$

By Condition C4, Lemma 6 is shown.

Lemma 7. Under the Regularity conditions C1, C2 and C4, we have

$$\frac{1}{m}\sum_{i\in I_2}[G_{pi}\{\widetilde{\boldsymbol{\gamma}}, \widetilde{\boldsymbol{\beta}}, \widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}, \widetilde{\boldsymbol{\gamma}})\} - \widetilde{G}_{pi}\{\widetilde{\boldsymbol{\gamma}}, \widetilde{\boldsymbol{\beta}}, \widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}, \widetilde{\boldsymbol{\gamma}})\}] = o_p(m^{-1/2}).$$

Lemma 7 can be proven by combining a second-order Taylor series expansion and several inequalities hence is omitted.

Lemma 8. Let

$$W_{ns} \equiv \frac{1}{n}\sum_{j\in I_1}\frac{1}{m}\sum_{i\in I_2}\frac{f_\epsilon^*(0,\mathbf{x}_i,\mathbf{z}_i)f_\epsilon(0,\mathbf{x}_i,\mathbf{z}_i)P_s(\mathbf{x}_i,\mathbf{z}_i)K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta} - \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})}{\tau(1-\tau)f_{\epsilon|\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}}(0,\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})f_{\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}}(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})}$$

$$([I\{\epsilon_j < g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})\} - I(\epsilon_j < 0)] -$$

$$[\mathrm{pr}\{\epsilon_j < g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})|\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}, \mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}\} - \mathrm{pr}(\epsilon_j < 0|\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})]),$$

where $P_s(\mathbf{x}_i, \mathbf{z}_i)$ is the sth element of $P(\mathbf{x}_i, \mathbf{z}_i)$. Under the Regularity Conditions C1-C5 and C9, we have $W_{ns} = o_p(n^{-1/2})$, for $s = 1, 2, \cdots, d_{\boldsymbol{\gamma}} + d_{\boldsymbol{\beta}} - 1$.

Proof. For notational brevity, for $i \in I_2$, let

$$\begin{aligned}
& L_{njs}(\mathbf{x}_i, \mathbf{z}_i, \boldsymbol{\beta}) \\
\equiv\ & \frac{f_\epsilon^*(0, \mathbf{x}_i, \mathbf{z}_i) f_\epsilon(0, \mathbf{x}_i, \mathbf{z}_i) P_s(\mathbf{x}_i, \mathbf{z}_i) K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta} - \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})}{\tau(1-\tau) f_{\epsilon|\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}}(0, \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) f_{\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}}(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})} \\
& ([I\{\epsilon_j < g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})\} - I(\epsilon_j < 0)] \\
& -[\mathrm{pr}\{\epsilon_j < g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})|\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}, \mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}\} - \mathrm{pr}(\epsilon_j < 0|\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})]).
\end{aligned}$$

Then,

$$W_{ns} = \frac{1}{n}\sum_{j \in I_1}\frac{1}{m}\sum_{i \in I_2} L_{njs}(\mathbf{x}_i, \mathbf{z}_i, \boldsymbol{\beta}).$$

Let $\mathscr{A} = \sigma(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}, j \in I_1, \mathbf{x}_i, \mathbf{z}_i, i \in I_2)$ and $c = 1/(2\alpha_1) + 1/4$. Note that under Condition C4, $c > 2$. By Bernstein inequality on page 95 of Serfling (1980), we have

$$\mathrm{pr}(|W_{ns}| \geqslant h_1^c \mathrm{log}n|\mathscr{A}) \leqslant 2\exp\{-\frac{(nh_1^{c+1}\mathrm{log}n)^2}{2\sum_{j \in I_1} var(L_{njs}|\mathscr{A}) + 4C_1 n h_1^{c+1}\mathrm{log}n/3}\},$$

where $L_{njs} = m^{-1}\sum_{i \in I_2} h_1 L_{njs}(\mathbf{x}_i, \mathbf{z}_i, \boldsymbol{\beta}), j \in I_1$ and $C_1 = \sup|h_1 L_{njs}(\mathbf{x}_i, \mathbf{z}_i, \boldsymbol{\beta})|$, which is a finite constant under Conditions C1, C2, C5 and C9. By simple derivation, we get

$$var(L_{njs}|\mathscr{A}) = E(L_{njs}^2|\mathscr{A})$$

$$
\begin{aligned}
&= \frac{1}{m^2}\sum_{i\in I_2} h_1^2 E\{L_{njs}^2(\mathbf{x}_i,\mathbf{z}_i,\boldsymbol{\beta})|\mathscr{A}\} \\
&\quad + \frac{2}{m^2}\sum_{p\neq q, p,q\in I_2} h_1^2 E\{L_{njs}(\mathbf{x}_p,\mathbf{z}_p,\boldsymbol{\beta})L_{njs}(\mathbf{x}_q,\mathbf{z}_q,\boldsymbol{\beta})|\mathscr{A}\}. \qquad \text{(S5.26)}
\end{aligned}
$$

Note that

$$
\begin{aligned}
& E\{([I\{\epsilon_j < g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})\} - I(\epsilon_j < 0)] \\
& -[\mathrm{pr}\{\epsilon_j < g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})|\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta},\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}\} - \mathrm{pr}(\epsilon_j < 0|\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})])^2|\mathscr{A}\} \\
= \; & |\mathrm{pr}\{\epsilon_j < g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})|\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta},\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}\} - \mathrm{pr}(\epsilon_j < 0|\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})| \\
& \times[1 - |\mathrm{pr}\{\epsilon_j < g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})|\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta},\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}\} - \mathrm{pr}(\epsilon_j < 0|\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})|].
\end{aligned}
$$

We first analyze the first term of (S5.26). We have

$$
\begin{aligned}
& \frac{1}{m^2}\sum_{i\in I_2} h_1^2 E\{L_{njs}^2(\mathbf{x}_i,\mathbf{z}_i,\boldsymbol{\beta})|\mathscr{A}\} \\
\leqslant\; & \frac{1}{m^2}\sum_{i\in I_2} h_1^2 K_{h_1}^2(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta} - \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})\left\{\frac{f_\epsilon^*(0,\mathbf{x}_i,\mathbf{z}_i)f_\epsilon(0,\mathbf{x}_i,\mathbf{z}_i)P_s(\mathbf{x}_i,\mathbf{z}_i)}{\tau(1-\tau)f_{\epsilon|\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}}(0,\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})f_{\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}}(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})}\right\}^2 \\
& \times|\mathrm{pr}\{\epsilon_j < g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})|\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta},\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}\} - \mathrm{pr}(\epsilon_j < 0|\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})| \\
\leqslant\; & \frac{C_2'}{m^2}\sum_{i\in I_2} h_1^2 K_{h_1}^2(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta} - \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})|\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta} - \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}| \\
\leqslant\; & \frac{C_2''}{m^2}\sum_{i\in I_2} h_1^3 K_{h_1}^2(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta} - \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) \\
=\; & \frac{C_2''h_1^2}{m}\{\frac{1}{mh_1}\sum_{i\in I_2} h_1^2 K_{h_1}^2(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta} - \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})\} \\
=\; & \frac{C_2''h_1^2}{m}[\int k^2(u)\mathrm{d}u f_{\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}) + O_p\{k(n,h_1)\}] \\
\leqslant\; & \frac{C_2h_1^2}{m} \qquad \text{(S5.27)}
\end{aligned}
$$

in probability, where C_2', C_2'', and C_2 are constants. The second inequality holds under

Conditions C1, C2, C3 and C9, the third inequality holds under Condition C5, the second last equality holds under Condition C4 and C5, and the last inequality holds under Conditions C2 and C5. Next we analyze the second term of (S5.26). Similar as the first term,

$$\begin{aligned}
& \frac{2}{m^2}\sum_{p\neq q,p,q\in I_2} h_1^2 E\{L_{njs}(\mathbf{x}_p,\mathbf{z}_p,\boldsymbol{\beta})L_{njs}(\mathbf{x}_q,\mathbf{z}_q,\boldsymbol{\beta})|\mathscr{A}\} \\
\leqslant\ & \frac{C_3}{m^2}\sum_{p\neq q,p,q\in I_2} h_1^2 |K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}-\mathbf{z}_p^{\mathrm{T}}\boldsymbol{\beta})K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}-\mathbf{z}_q^{\mathrm{T}}\boldsymbol{\beta})| \\
& \times[|\mathrm{pr}\{\epsilon_j<g(\mathbf{z}_p^{\mathrm{T}}\boldsymbol{\beta})-g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})|\mathbf{z}_p^{\mathrm{T}}\boldsymbol{\beta},\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}\}-\mathrm{pr}(\epsilon_j<0|\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})| \\
& +|\mathrm{pr}\{\epsilon_j<g(\mathbf{z}_q^{\mathrm{T}}\boldsymbol{\beta})-g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})|\mathbf{z}_q^{\mathrm{T}}\boldsymbol{\beta},\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}\}-\mathrm{pr}(\epsilon_j<0|\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})|] \\
\leqslant\ & \frac{C_4}{m^2}\{\sum_{i\in I_2} I(|\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}-\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}|\leqslant h_1)\}^2 h_1 \\
=\ & C_4 h_1^3\{\frac{1}{mh_1}\sum_{i\in I_2} I(|\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}-\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}|\leqslant h_1)\}^2 \\
=\ & C_4 h_1^3[f_{\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})+O_p\{k(n,h_1)\}]^2 \\
\leqslant\ & C_5 h_1^3
\end{aligned}$$

in probability. The first inequality holds under Conditions C1, C2, C3, C5 and C9. The second inequality holds under Conditions C3, C5 and C9, the second last equality and the last inequality hold under Conditions C2 and C4. Then combining(S5.26), (S5.27) and (S5.28), we obtain

$$\sum_{j\in I_1} var(L_{njs}|\mathscr{A}) \leqslant \sum_{j\in I_1}(\frac{C_2h_1^2}{m}+C_5h_1^3) \leqslant C_6(h_1^2+nh_1^3) \ \leqslant\ C_7nh_1^3\mathrm{log}n \tag{S5.28}$$

in probability. Then we can get

$$\begin{aligned}
\frac{(nh_1^{c+1}\mathrm{log}n)^2}{2\sum_{j\in I_1}var(L_{njs}|\mathscr{A})+4C_1nh_1^{c+1}\mathrm{log}n/3} \ &\geqslant\ C_8(nh_1^{c+1}\mathrm{log}n)^2/nh_1^3\mathrm{log}n \\
&=\ C_8nh_1^{2c-1}\mathrm{log}n
\end{aligned} \tag{S5.29}$$

in probability. Also we have

$$\mathrm{pr}(|W_{ns}| \geqslant h_1^c \log n) = E\{\mathrm{pr}(|W_{ns}| \geqslant h_1^c \log n | \mathscr{A})\}. \tag{S5.30}$$

Then combining (S5.26), (S5.29), and (S5.30), we get

$$\mathrm{pr}(|W_{ns}| \geqslant h_1^c \log n) \leqslant \exp(-C_8 n h_1^{2c-1} \log n)$$

which goes to zero under Condition C4. This means $W_{ns} = O_p(h_1^c \log n) = O_p(n^{-1/2-\alpha_1/4} \log n) = o_p(n^{-1/2})$. Lemma 8 is proved.

<u>Lemma</u> 9. Under the Regularity Conditions C1-C5 and C9, we have

$$\begin{aligned}
& \frac{1}{m}\sum_{i\in I_2} \frac{f_\epsilon^*(0,\mathbf{x}_i,\mathbf{z}_i) f_\epsilon(0,\mathbf{x}_i,\mathbf{z}_i)}{\tau(1-\tau) f_{\epsilon|\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}}(0,\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})} P(\mathbf{x}_i,\mathbf{z}_i) \widetilde{F}\{g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}),\boldsymbol{\gamma},\boldsymbol{\beta}\} \\
= & \frac{1}{m}\sum_{i\in I_2} \frac{f_\epsilon^*(0,\mathbf{x}_i,\mathbf{z}_i) f_\epsilon(0,\mathbf{x}_i,\mathbf{z}_i)}{\tau(1-\tau) f_{\epsilon|\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}}(0,\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) n^{-1}\sum_{k\in I_1} K_{h_1}(\mathbf{z}_k^{\mathrm{T}}\boldsymbol{\beta} - \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})} P(\mathbf{x}_i,\mathbf{z}_i) \\
& \frac{1}{n}\sum_{j\in I_1} K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta} - \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})(I(\epsilon_j<0) + [\mathrm{pr}\{\epsilon_j < g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}) | \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}, \mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}\} \\
& -\mathrm{pr}(\epsilon_j < 0 | \mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})]) + o_p(n^{-1/2}).
\end{aligned}$$

Lemma 9 can be proven by combining Lemma 8 and standard kernel regression analysis and is omitted.

<u>Lemma</u> 10. For notational brevity, let

$$\begin{aligned}
W_n^* \equiv & \frac{1}{m}\sum_{i\in I_2} \frac{f_\epsilon^*(0,\mathbf{x}_i,\mathbf{z}_i) f_\epsilon(0,\mathbf{x}_i,\mathbf{z}_i)}{\tau(1-\tau) f_{\epsilon|\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}}(0,\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) n^{-1}\sum_{k\in I_1} K_{h_1}(\mathbf{z}_k^{\mathrm{T}}\boldsymbol{\beta} - \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})} P(\mathbf{x}_i,\mathbf{z}_i) \\
& \frac{\sqrt{m}}{n}\sum_{j\in I_1} K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta} - \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})(\{\tau - I(\epsilon_j<0)\} \\
& -[\mathrm{pr}\{\epsilon_j < g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}) | \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}, \mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}\} - \mathrm{pr}(\epsilon_j < 0 | \mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})]).
\end{aligned}$$

Under the Regularity Conditions C1, C2, C4, C5,C9 and C10, we have

$$W_n^* = \frac{\sqrt{m}}{n}\sum_{j\in I_1}\frac{E\{f_\epsilon^*(0,\mathbf{x}_j,\mathbf{z}_j)f_\epsilon(0,\mathbf{x}_j,\mathbf{z}_j)P(\mathbf{x}_j,\mathbf{z}_j)|\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}\}}{\tau(1-\tau)f_\epsilon(0,\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})}\psi_j\{\boldsymbol{\gamma},g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})\}+o_p(1),$$

where $\psi_j\{\boldsymbol{\gamma},g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})\}\equiv\tau-I\{Y_j-\mathbf{x}_j^{\mathrm{T}}\boldsymbol{\gamma}-g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})<0\}=\tau-I(\epsilon_j<0)$.

Proof.

$$\begin{aligned}
W_n^* &= \frac{1}{m}\sum_{i\in I_2}\frac{f_\epsilon^*(0,\mathbf{x}_i,\mathbf{z}_i)f_\epsilon(0,\mathbf{x}_i,\mathbf{z}_i)}{\tau(1-\tau)f_{\epsilon|\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}}(0,\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})f_{\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}}(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})}P(\mathbf{x}_i,\mathbf{z}_i)\\
&\quad \frac{\sqrt{m}}{n}\sum_{j\in I_1}K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}-\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})(\{\tau-I(\epsilon_j<0)\}\\
&\quad -[\mathrm{pr}\{\epsilon_j<g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})-g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})|\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta},\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}\}-\mathrm{pr}(\epsilon_j<0|\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})])\\
&\quad +\frac{1}{m}\sum_{i\in I_2}\frac{f_\epsilon^*(0,\mathbf{x}_i,\mathbf{z}_i)f_\epsilon(0,\mathbf{x}_i,\mathbf{z}_i)}{\tau(1-\tau)f_{\epsilon|\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}}(0,\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})f_{\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}}(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})}P(\mathbf{x}_i,\mathbf{z}_i)\\
&\quad \frac{\sqrt{m}}{n}\sum_{j\in I_1}K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}-\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})\frac{f_{\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}}(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})-n^{-1}\sum_{k\in I_1}K_{h_1}(\mathbf{z}_k^{\mathrm{T}}\boldsymbol{\beta}-\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})}{n^{-1}\sum_{k\in I_1}K_{h_1}(\mathbf{z}_k^{\mathrm{T}}\boldsymbol{\beta}-\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})}\\
&\quad \times(\{\tau-I(\epsilon_j<0)\}-[\mathrm{pr}\{\epsilon_j<g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})-g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})|\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta},\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}\}\\
&\quad -\mathrm{pr}(\epsilon_j<0|\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})])=B_1+B_2.
\end{aligned}\tag{S5.31}$$

Note that

$$\begin{aligned}
&\frac{1}{m}\sum_{i\in I_2}\frac{f_\epsilon^*(0,\mathbf{x}_i,\mathbf{z}_i)f_\epsilon(0,\mathbf{x}_i,\mathbf{z}_i)}{\tau(1-\tau)f_{\epsilon|\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}}(0,\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})f_{\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}}(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})}P(\mathbf{x}_i,\mathbf{z}_i)\\
&\frac{\sqrt{m}}{n}\sum_{j\in I_1}K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}-\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})\psi_j\{\boldsymbol{\gamma},g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})\}\\
=\ &\frac{\sqrt{m}}{n}\sum_{j\in I_1}\frac{E\{f_\epsilon^*(0,\mathbf{x}_j,\mathbf{z}_j)f_\epsilon(0,\mathbf{x}_j,\mathbf{z}_j)P(\mathbf{x}_j,\mathbf{z}_j)|\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}\}}{\tau(1-\tau)f_\epsilon(0,\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})}\psi_j\{\boldsymbol{\gamma},g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})\}\\
&+o_p(1).
\end{aligned}\tag{S5.32}$$

The above equality is obtained by using standard kernel regression analysis and by noting that $E[\psi_j\{\boldsymbol{\gamma}, g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})\} \mid \mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}] = 0$. For notational brevity, let

$$S_{is}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}) \equiv \frac{K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta} - \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) f_\epsilon^*(0, \mathbf{x}_i, \mathbf{z}_i) f_\epsilon(0, \mathbf{x}_i, \mathbf{z}_i) P_s(\mathbf{x}_i, \mathbf{z}_i)}{f_{\epsilon|\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}}(0, \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) f_{\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}}(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})}\{g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})\},$$

where $i \in I_2$, $j \in I_1$, and $s = 1, 2, \cdots, d_{\boldsymbol{\gamma}} + d_{\boldsymbol{\beta}} - 1$. Let $\mathbf{S}_i(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}) = \{S_{i1}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}), \ldots, S_{id_{\boldsymbol{\gamma}}+d_{\boldsymbol{\beta}}-1}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})\}^{\mathrm{T}}$.

Then we also have

$$\begin{aligned}
&\frac{1}{m}\sum_{i\in I_2}\frac{f_\epsilon^*(0, \mathbf{x}_i, \mathbf{z}_i) f_\epsilon(0, \mathbf{x}_i, \mathbf{z}_i)}{\tau(1-\tau) f_{\epsilon|\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}}(0, \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) f_{\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}}(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})} P(\mathbf{x}_i, \mathbf{z}_i)\\
&\frac{\sqrt{m}}{n}\sum_{j\in I_1} K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta} - \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})[\mathrm{pr}\{Y_j - \mathbf{x}_j^{\mathrm{T}}\boldsymbol{\gamma} - g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) < 0 | \mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}\}\\
&-\mathrm{pr}\{Y_j - \mathbf{x}_j^{\mathrm{T}}\boldsymbol{\gamma} - g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}) < 0 | \mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}\}]\\
= \;&\frac{1}{m}\sum_{i\in I_2}\frac{f_\epsilon^*(0, \mathbf{x}_i, \mathbf{z}_i) f_\epsilon(0, \mathbf{x}_i, \mathbf{z}_i)}{\tau(1-\tau) f_{\epsilon|\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}}(0, \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) f_{\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}}(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})} P(\mathbf{x}_i, \mathbf{z}_i)\\
&\frac{\sqrt{m}}{n}\sum_{j\in I_1} K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta} - \mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})[f_{\epsilon|\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}}(0, \mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})\{g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})\}\\
&+2^{-1} f'_{\epsilon|\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}}(g_{ij}^*, \mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})\{g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})\}^2]\\
= \;&\frac{\sqrt{m}}{n}\sum_{j\in I_1}\frac{f_{\epsilon|\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}}(0, \mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})}{\tau(1-\tau)}\frac{1}{m}\sum_{i\in I_2}\mathbf{S}_i(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})\\
&+\frac{\sqrt{m}}{2n}\sum_{j\in I_1}\frac{1}{m}\sum_{i\in I_2}\{g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})\}\mathbf{S}_i(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})\frac{f'_{\epsilon|\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}}(g_{ij}^*, \mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})}{\tau(1-\tau)},
\end{aligned} \tag{S5.33}$$

where g_{ij}^* is between $g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}) - g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})$ and 0. By standard kernel regression analysis, under Conditions C5 and C10, we have

$$E\{S_{is}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}) | \mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}\} = O_p(h_1^4). \tag{S5.34}$$

We can also get

$$\begin{aligned}
E\{S_{is}^2(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})|\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}\} &\leqslant CE\left(\frac{K_{h_1}^2(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}-\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})}{f_{\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}}(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})}(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}-\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})^2|\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}\right) \\
&= C\int\frac{K_{h_1}^2(t-\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})}{f_{\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}}(t)}(t-\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})^2 f_{\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}}(t)\mathrm{d}t \\
&= C\int K^2(u)/h_1^2(h_1u)^2h_1\mathrm{d}u \leqslant C'h_1. \qquad \text{(S5.35)}
\end{aligned}$$

The first inequality holds under Conditions C1, C2, andC9, and the last inequality holds under Condition C5. By Chebyshev's inequality, we can get

$$\mathrm{pr}\{|S_{is}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})| \geqslant 1\} \leqslant E\{S_{is}^2(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})\} \leqslant C'h_1. \qquad \text{(S5.36)}$$

Let $S_s(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}) = m^{-1}\sum_{i\in I_2} S_{is}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}), j \in I_1$. Under Condition C4, (S5.35) and (S5.36), together with Lemma 18 (page 20), Lemma 25 (page 27), Lemma 28 (page 30), the proof of Lemma 33 (page 31-33) and theorem 37 (page 34) in Pollard (1984), yield

$$\sup_{\mathbf{z}_j}|S_s(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}) - ES_s(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})| = O_p(m^{-1/2}h_1^{1/2}\mathrm{log}m). \qquad \text{(S5.37)}$$

By (S5.34) and (S5.37), we get

$$\begin{aligned}
\sup_{\mathbf{z}_j} S_s(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}) &\leqslant \sup_{\mathbf{z}_j}|S(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}) - ES(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})| + \sup_{\mathbf{z}_j}|E\{S_s(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})\}| \\
&= O_p(m^{-1/2}h_1^{1/2}\mathrm{log}m) + O_p(h_1^4). \qquad \text{(S5.38)}
\end{aligned}$$

By (S5.38), we get the first term of (S5.33) is $O_p(h_1^{1/2}\mathrm{log}m) + O_p(m^{1/2}h_1^4) = o_p(1)$ under Conditions C2 and C4. Similarly we can prove the second term of (S5.33) is $o_p(1)$. Combining these results and (S5.33), we get

$$\begin{aligned}
&\frac{1}{m}\sum_{i\in I_2}\frac{f_\epsilon^*(0,\mathbf{x}_i,\mathbf{z}_i)f_\epsilon(0,\mathbf{x}_i,\mathbf{z}_i)}{\tau(1-\tau)f_{\epsilon|\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}}(0,\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})f_{\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}}(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})}P(\mathbf{x}_i,\mathbf{z}_i)\\
&\frac{\sqrt{m}}{n}\sum_{j\in I_1}K_{h_1}(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}-\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})[\mathrm{pr}\{Y_j-\mathbf{x}_j^{\mathrm{T}}\boldsymbol{\gamma}-g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})<0|\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}\}\\
&-\mathrm{pr}\{Y_j-\mathbf{x}_j^{\mathrm{T}}\boldsymbol{\gamma}-g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})<0|\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}\}]=o_p(1). \qquad \text{(S5.39)}
\end{aligned}$$

By (S5.32) and (S5.39), we get

$$\begin{aligned}
B_1 &= \frac{\sqrt{m}}{n}\sum_{j\in I_1}\frac{E\{f_\epsilon^*(0,\mathbf{x}_j,\mathbf{z}_j)f_\epsilon(0,\mathbf{x}_j,\mathbf{z}_j)P(\mathbf{x}_j,\mathbf{z}_j)|\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}\}}{\tau(1-\tau)f_\epsilon(0,\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})}\psi_j\{\boldsymbol{\gamma},g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})\}\\
&\quad +o_p(1). \qquad \text{(S5.40)}
\end{aligned}$$

Using the results of (S5.40) and similar standard kernel regression analysis, we can get $B_2=o_p(1)$. Combining these results, (S5.31) and (S5.40), Lemma 10 is proved.

S6 Proof of Theorem 3

<u>Lemma</u> 11. Under the Regularity Conditions C1-C5,C9 and C10,we have

$$\begin{aligned}
&\frac{1}{\sqrt{m}}\sum_{i\in I_2}G_i\{\widetilde{\boldsymbol{\gamma}},\widetilde{\boldsymbol{\beta}},\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}},\widetilde{\boldsymbol{\gamma}})\}\\
=\;&\frac{1}{\sqrt{m}}\sum_{i\in I_2}G_i\{\boldsymbol{\gamma},\boldsymbol{\beta},g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})\}-V\begin{pmatrix}\sqrt{m}(\widetilde{\boldsymbol{\gamma}}-\boldsymbol{\gamma})\\ \sqrt{m}(\widetilde{\boldsymbol{\beta}}_{-1}-\boldsymbol{\beta}_{-1})\end{pmatrix}\\
&-\frac{\sqrt{m}}{n}\sum_{j\in I_1}\frac{E\{f_\epsilon^*(0,\mathbf{x}_j,\mathbf{z}_j)f_\epsilon(0,\mathbf{x}_j,\mathbf{z}_j)P(\mathbf{x}_j,\mathbf{z}_j)|\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}\}}{\tau(1-\tau)f_{\epsilon|\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}}(0,\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})}\psi_j\{\boldsymbol{\gamma},g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})\}\\
&+o_p(1).
\end{aligned}$$

Lemma 11 can be proven by using Lemma 6, Lemma 7, Lemma 9, Lemma 10, the convergence rate of $\widetilde{\boldsymbol{\gamma}}$ and $\widetilde{\boldsymbol{\beta}}$ and the law of large numbers.

Lemma 12. Under the Regularity Conditions C1, C3, C4, C5-C8, we have

$$\sup_{i\in I_2}\{\widehat{P}(\mathbf{x}_i,\mathbf{z}_i)-P(\mathbf{x}_i,\mathbf{z}_i)\}=o_p(1).$$

Lemma 12 can be proven by combining the definitions of $\widehat{P}(\mathbf{x}_i,\mathbf{z}_i)$ and $P(\mathbf{x}_i,\mathbf{z}_i)$, Proposition 2 and the standard kernel regression properties.

Lemma 13. Under the Regularity Conditions C1-C10, we have

$$\begin{aligned}
& (\widehat{V}_0)^{-1}\frac{1}{\sqrt{m}}\sum_{i\in I_2}\frac{f_\epsilon^*(0,\mathbf{x}_i,\mathbf{z}_i)}{\tau(1-\tau)}[\tau-I\{Y_i-\mathbf{x}_i^{\mathrm{T}}\widetilde{\boldsymbol{\gamma}}-\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}},\widetilde{\boldsymbol{\gamma}})<0\}]\widehat{P}(\mathbf{x}_i,\mathbf{z}_i)\\
= & V_0^{-1}\frac{1}{\sqrt{m}}\sum_{i\in I_2}G_i\{\widetilde{\boldsymbol{\gamma}},\widetilde{\boldsymbol{\beta}},\widetilde{g}(\mathbf{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}},\widetilde{\boldsymbol{\gamma}})\}+o_p(1),
\end{aligned}$$

where V_0 is defined in Section S2.

Lemma 13 can be proven by using Lemma 11, Lemma 12 and the similar proof of Lemma 12.

Lemma 14. Under the Regularity Conditions, C1-C10, we have

$$\begin{aligned}
& \begin{pmatrix}\sqrt{m}(\widehat{\boldsymbol{\gamma}}-\boldsymbol{\gamma})\\ \sqrt{m}(\widehat{\boldsymbol{\beta}}_{-1}-\boldsymbol{\beta}_{-1})\end{pmatrix}\\
= & V_0^{-1}(V_0-V)\begin{pmatrix}\sqrt{m}(\widetilde{\boldsymbol{\gamma}}-\boldsymbol{\gamma})\\ \sqrt{m}(\widetilde{\boldsymbol{\beta}}_{-1}-\boldsymbol{\beta}_{-1})\end{pmatrix}\\
& +V_0^{-1}\frac{1}{\sqrt{m}}\sum_{i\in I_2}\frac{f_\epsilon^*(0,\mathbf{x}_i,\mathbf{z}_i)}{\tau(1-\tau)}P(\mathbf{x}_i,\mathbf{z}_i)\psi_i\{\boldsymbol{\gamma},g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})\}\\
& -V_0^{-1}\frac{\sqrt{m}}{n}\sum_{j\in I_1}\frac{E\{f_\epsilon^*(0,\mathbf{x}_j,\mathbf{z}_j)f_\epsilon(0,\mathbf{x}_j,\mathbf{z}_j)P(\mathbf{x}_j,\mathbf{z}_j)|\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta}\}}{\tau(1-\tau)f_{\epsilon|\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}}(0,\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})}\psi_j\{\boldsymbol{\gamma},g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})\}
\end{aligned}$$

$+o_p(1).$

Lemma 14 can be proven by Lemma 11 and Lemma 13.

Proof of Theorem 3. Repeating all the analysis for $\widehat{\boldsymbol{\beta}}_{-1}$ and $\widehat{\boldsymbol{\gamma}}$, we can get

$$
\begin{aligned}
&\begin{pmatrix} \sqrt{n}(\widehat{\boldsymbol{\gamma}}' - \boldsymbol{\gamma}) \\ \sqrt{n}(\widehat{\boldsymbol{\beta}}'_{-1} - \boldsymbol{\beta}_{-1}) \end{pmatrix} \\
= \; & V_0^{-1}(V_0 - V) \begin{pmatrix} \sqrt{n}(\widetilde{\boldsymbol{\gamma}}' - \boldsymbol{\gamma}) \\ \sqrt{n}(\widetilde{\boldsymbol{\beta}}'_{-1} - \boldsymbol{\beta}_{-1}) \end{pmatrix} \\
& + V_0^{-1} \frac{1}{\sqrt{n}} \sum_{i \in I_1} \frac{f_\epsilon^*(0, \mathbf{x}_i, \mathbf{z}_i)}{\tau(1-\tau)} P(\mathbf{x}_i, \mathbf{z}_i) \psi_i\{\boldsymbol{\gamma}, g(\mathbf{z}_i^{\mathrm{T}} \boldsymbol{\beta})\} \\
& - V_0^{-1} \frac{\sqrt{n}}{m} \sum_{j \in I_2} \frac{E\{f_\epsilon^*(0, \mathbf{x}_j, \mathbf{z}_j) f_\epsilon(0, \mathbf{x}_j, \mathbf{z}_j) P(\mathbf{x}_j, \mathbf{z}_j) | \mathbf{z}_j^{\mathrm{T}} \boldsymbol{\beta}\}}{\tau(1-\tau) f_{\epsilon|\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}}(0, \mathbf{z}_j^{\mathrm{T}} \boldsymbol{\beta})} \psi_j\{\boldsymbol{\gamma}, g(\mathbf{z}_j^{\mathrm{T}} \boldsymbol{\beta})\} \\
& + o_p(1), \qquad \text{(S6.41)}
\end{aligned}
$$

where $\widehat{\boldsymbol{\gamma}}'$ and $\widehat{\boldsymbol{\beta}}'_{-1}$ are the initial estimators of $\boldsymbol{\gamma}$ and $\boldsymbol{\beta}_{-1}$ based on the data from the group I_2. Combining Lemma 14 and (S6.41), we have

$$
\begin{aligned}
&\sqrt{N} \begin{pmatrix} (\widehat{\boldsymbol{\gamma}} + \widehat{\boldsymbol{\gamma}}')/2 - \boldsymbol{\gamma} \\ (\widehat{\boldsymbol{\beta}}_{-1} + \widehat{\boldsymbol{\beta}}'_{-1})/2 - \boldsymbol{\beta}_{-1}) \end{pmatrix} \\
= \; & V_0^{-1}(V_0 - V) \frac{\sqrt{N}}{2} \begin{pmatrix} (\widetilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma}) + (\widetilde{\boldsymbol{\gamma}}' - \boldsymbol{\gamma}) \\ (\widetilde{\boldsymbol{\beta}}_{-1} - \boldsymbol{\beta}_{-1}) + (\widetilde{\boldsymbol{\beta}}'_{-1} - \boldsymbol{\beta}_{-1}) \end{pmatrix} \\
& + V_0^{-1} \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \frac{f_\epsilon^*(0, \mathbf{x}_i, \mathbf{z}_i)}{\tau(1-\tau)} P(\mathbf{x}_i, \mathbf{z}_i) \psi_i\{\boldsymbol{\gamma}, g(\mathbf{z}_i^{\mathrm{T}} \boldsymbol{\beta})\} \\
& - V_0^{-1} \frac{1}{\sqrt{N}} \sum_{j=1}^{N} \frac{E\{f_\epsilon^*(0, \mathbf{x}_j, \mathbf{z}_j) f_\epsilon(0, \mathbf{x}_j, \mathbf{z}_j) P(\mathbf{x}_j, \mathbf{z}_j) | \mathbf{z}_j^{\mathrm{T}} \boldsymbol{\beta}\}}{\tau(1-\tau) f_{\epsilon|\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}}(0, \mathbf{z}_j^{\mathrm{T}} \boldsymbol{\beta})} \psi_j\{\boldsymbol{\gamma}, g(\mathbf{z}_j^{\mathrm{T}} \boldsymbol{\beta})\} \\
& + o_p(1)
\end{aligned}
$$

$$= \quad V_0^{-1}\{(V_0 - V)M_1 + M_2 - M_3)\} + o_p(1),$$

where M_1 satisfies

$$\frac{\sqrt{N}}{2}\begin{pmatrix} (\widetilde{\boldsymbol\gamma} - \boldsymbol\gamma) + (\widetilde{\boldsymbol\gamma}' - \boldsymbol\gamma) \\ (\widetilde{\boldsymbol\beta}_{-1} - \boldsymbol\beta_{-1}) + (\widetilde{\boldsymbol\beta}'_{-1} - \boldsymbol\beta_{-1}) \end{pmatrix} = M_1 + o_p(1),$$

$$M_2 = \frac{1}{\sqrt{N}}\sum_{i=1}^N \frac{f_\epsilon^*(0, \mathbf{x}_i, \mathbf{z}_i)}{\tau(1-\tau)} P(\mathbf{x}_i, \mathbf{z}_i)\psi_i\{\boldsymbol\gamma, g(\mathbf{z}_i^{\mathrm{T}}\boldsymbol\beta)\},$$

$$M_3 = \frac{1}{\sqrt{N}}\sum_{j=1}^N \frac{E\{f_\epsilon^*(0, \mathbf{x}_j, \mathbf{z}_j) f_\epsilon(0, \mathbf{x}_j, \mathbf{z}_j) P(\mathbf{x}_j, \mathbf{z}_j)|\mathbf{z}_j^{\mathrm{T}}\boldsymbol\beta\}}{\tau(1-\tau) f_{\epsilon|\mathbf{z}^{\mathrm{T}}\boldsymbol\beta}(0, \mathbf{z}_j^{\mathrm{T}}\boldsymbol\beta)} \psi_j\{\boldsymbol\gamma, g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol\beta)\}.$$

Write $J \equiv (V_0 - V)$. Then we get

$$\begin{pmatrix} \sqrt{N}(\overline{\boldsymbol\gamma} - \boldsymbol\gamma) \\ \sqrt{N}(\overline{\boldsymbol\beta}_{-1} - \boldsymbol\beta_{-1}) \end{pmatrix} \to N(0, \boldsymbol\Sigma)$$

in distribution with

$$\begin{aligned} \boldsymbol\Sigma \quad = \quad & V_0^{-1}\{\mathrm{cov}(JM_1) + \mathrm{cov}(M_2) + \mathrm{cov}(M_3) + 2\mathrm{cov}(JM_1, M_2) \\ & -2\mathrm{cov}(JM_1, M_3) - 2\mathrm{cov}(M_2, M_3)\}V_0^{-1}, \end{aligned} \tag{S6.42}$$

where

$$\begin{aligned} \mathrm{cov}(M_2) \quad &= \quad V_0, \\ \mathrm{cov}(M_3) \quad &= \quad E\left(\frac{[E\{f_\epsilon^*(0, \mathbf{x}_i, \mathbf{z}_i) f_\epsilon(0, \mathbf{x}_i, \mathbf{z}_i) P(\mathbf{x}_i, \mathbf{z}_i)|\mathbf{z}_i^{\mathrm{T}}\boldsymbol\beta\}]^{\otimes 2}}{\tau(1-\tau) f^2_{\epsilon|\mathbf{z}^{\mathrm{T}}\boldsymbol\beta}(0, \mathbf{z}_i^{\mathrm{T}}\boldsymbol\beta)}\right), \\ \mathrm{cov}(M_2, M_3) \quad &= \quad E\left[\frac{f_\epsilon^*(0, \mathbf{x}_i, \mathbf{z}_i)}{\tau(1-\tau) f_{\epsilon|\mathbf{z}^{\mathrm{T}}\boldsymbol\beta}(0, \mathbf{z}_i^{\mathrm{T}}\boldsymbol\beta)} P(\mathbf{x}_i, \mathbf{z}_i) E\{f_\epsilon^*(0, \mathbf{x}_i, \mathbf{z}_i)\right. \end{aligned}$$

$$f_\epsilon(0, \mathbf{x}_i, \mathbf{z}_i) P(\mathbf{x}_i, \mathbf{z}_i) | \mathbf{z}_i^{\mathrm{T}} \boldsymbol{\beta}\}^{\mathrm{T}} \Bigg].$$

Theorem 3 is proved.

Remark 1. We now provide the detailed form of M_1 and related quantities if the initial estimators of $\boldsymbol{\gamma}$ and $\boldsymbol{\beta}_{-1}$ are obtained by the method proposed in Ma and He (2016). Theorem 2 in Ma and He (2016) provided the asymptotic results of the initial estimators of $\boldsymbol{\gamma}$ and $\boldsymbol{\beta}$ under their parameterization. Using the delta method, we can obtain the corresponding results for $\widetilde{\boldsymbol{\gamma}}$ and $\widetilde{\boldsymbol{\beta}}_{-1}$. Specifically, let $\mathbf{D} \equiv \text{diag}\left(I_{d_{\boldsymbol{\gamma}} \times d_{\boldsymbol{\gamma}}}, \mathbf{A}\right)$, $\mathbf{A} \equiv \|\widetilde{\boldsymbol{\beta}}\|_2(-\widetilde{\boldsymbol{\beta}}_{-1}, I_{(d_{\boldsymbol{\beta}}-1)\times(d_{\boldsymbol{\beta}}-1)})$, where $I_{(d_{\boldsymbol{\beta}}-1)\times(d_{\boldsymbol{\beta}}-1)}$ and $I_{d_{\boldsymbol{\gamma}} \times d_{\boldsymbol{\gamma}}}$ are identity matrices, and let

$$\Omega \equiv E\{f_\epsilon(0, \mathbf{x}, \mathbf{z}) P^0(\mathbf{x}, \mathbf{z})^{\otimes 2}\}.$$

Note that D is matrix with full row rank and Ω is a positive definite matrix. Let D^+ satisfy $DD^+ = \mathbf{I}_{d_{\boldsymbol{\beta}}+d_{\boldsymbol{\gamma}}-1}$. Then, following Ma and He (2016),

$$\begin{aligned}
&\begin{pmatrix} \sqrt{n}(\widetilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma}) \\ \sqrt{n}(\widetilde{\boldsymbol{\beta}}_{-1} - \boldsymbol{\beta}_{-1}) \end{pmatrix} \\
&= D(D^{\mathrm{T}}\Omega D)^+ D^{\mathrm{T}} \frac{1}{\sqrt{n}} \sum_{j \in I_1} \psi_j\{\boldsymbol{\gamma}, g(\mathbf{z}_j^{\mathrm{T}} \boldsymbol{\beta})\} P^0(\mathbf{x}_j, \mathbf{z}_j) + o_p(1) \\
&= \Omega^{-1} \frac{1}{\sqrt{n}} \sum_{j \in I_1} \psi_j\{\boldsymbol{\gamma}, g(\mathbf{z}_j^{\mathrm{T}} \boldsymbol{\beta})\} P^0(\mathbf{x}_j, \mathbf{z}_j) + o_p(1).
\end{aligned} \tag{S6.43}$$

By the same derivation, the corresponding results for $\widetilde{\boldsymbol{\gamma}}'$ and $\widetilde{\boldsymbol{\beta}}'_{-1}$ are

$$\begin{pmatrix} \sqrt{n}(\widetilde{\boldsymbol{\gamma}}' - \boldsymbol{\gamma}) \\ \sqrt{n}(\widetilde{\boldsymbol{\beta}}'_{-1} - \boldsymbol{\beta}_{-1}) \end{pmatrix}$$

$$= \Omega^{-1}\frac{1}{\sqrt{m}}\sum_{j\in I_2}\psi_j\{\boldsymbol{\gamma}, g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})\}P^0(\mathbf{x}_j,\mathbf{z}_j) + o_p(1). \tag{S6.44}$$

Based on (S6.43) and (S6.44), we can get

$$\begin{aligned}
&\frac{\sqrt{N}}{2}\begin{pmatrix} (\widetilde{\boldsymbol{\gamma}}-\boldsymbol{\gamma}) + (\widetilde{\boldsymbol{\gamma}}'-\boldsymbol{\gamma}) \\ (\widetilde{\boldsymbol{\beta}}_{-1}-\boldsymbol{\beta}_{-1}) + (\widetilde{\boldsymbol{\beta}}'_{-1}-\boldsymbol{\beta}_{-1}) \end{pmatrix} \\
= & \frac{1}{\sqrt{N}}\Omega^{-1}\sum_{j=1}^{N}\psi_j\{\boldsymbol{\gamma}, g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})\}P^0(\mathbf{x}_j,\mathbf{z}_j) + o_p(1).
\end{aligned}$$

This means $M_1 = \Omega^{-1}\frac{1}{\sqrt{N}}\sum_{j=1}^{N}\psi_j\{\boldsymbol{\gamma}, g(\mathbf{z}_j^{\mathrm{T}}\boldsymbol{\beta})\}P^0(\mathbf{x}_j,\mathbf{z}_j)$. Then $\mathrm{cov}(JM_1) = J\Omega' J^{\mathrm{T}}$, where $\Omega' = \tau(1-\tau)\Omega^{-1}E\{P^0(\mathbf{x}_j,\mathbf{z}_j)^{\otimes 2}\}\Omega^{-1}$.

$$\begin{aligned}
\mathrm{cov}(JM_1, M_2) &= J\Omega^{-1}E\left[f_\epsilon^*(0,\mathbf{x}_i,\mathbf{z}_i)P^0(\mathbf{x}_i,\mathbf{z}_i)P(\mathbf{x}_i,\mathbf{z}_i)^{\mathrm{T}}\right], \\
\mathrm{cov}(JM_1, M_3) &= J\Omega^{-1}E\Bigg[P^0(\mathbf{x}_i,\mathbf{z}_i) \\
&\qquad \frac{E\{f_\epsilon^*(0,\mathbf{x}_i,\mathbf{z}_i)f_\epsilon(0,\mathbf{x}_i,\mathbf{z}_i)P(\mathbf{x}_i,\mathbf{z}_i)|\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta}\}^{\mathrm{T}}}{f_{\epsilon|\mathbf{z}^{\mathrm{T}}\boldsymbol{\beta}}(0,\mathbf{z}_i^{\mathrm{T}}\boldsymbol{\beta})}\Bigg].
\end{aligned}$$

S7 Figures for Setting 2 of Simulation 2

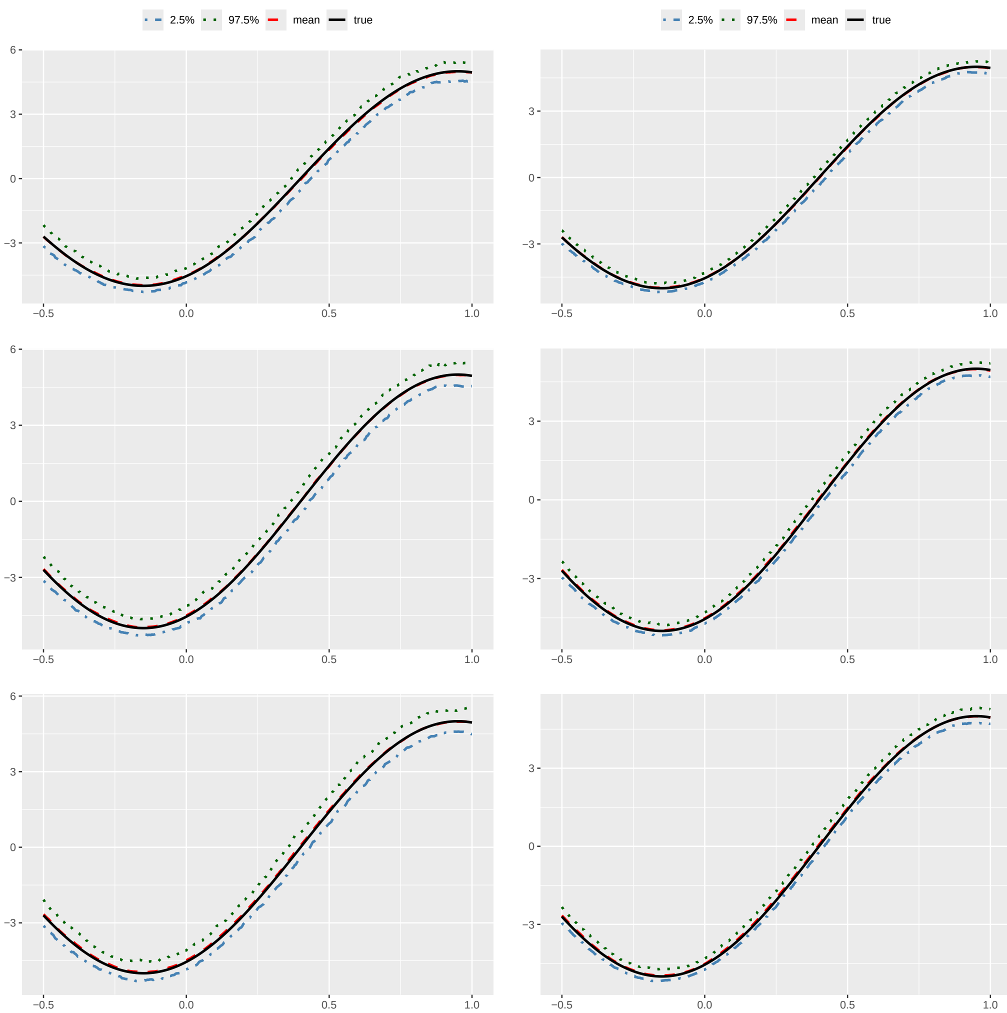

Figure S.1: Plots of estimators for $g(\cdot)$ in Simulation 2, Setting 2: The true error $\epsilon_i = \sigma(\mathbf{x}_i, \mathbf{z}_i)\{\widetilde{\epsilon}_i - F^{-1}(\tau)\}$, where $\widetilde{\epsilon}_i = F(\cdot)$ and $\sigma(\mathbf{x}_i, \mathbf{z}_i) = \sqrt{|\widetilde{\mathbf{x}}_{i2}|}$. Estimation is based on our proposed method with the correct model $\widetilde{\epsilon}_i \sim SN(0, 1, 3)$. Top to bottom panels: $\tau = 0.4, 0.5, 0.6$. Left panels: $n = 400$, right panels: $n = 1000$. The black sold lines are the true, the dashed red lines are the mean curve, the blue dot-dashed lines are 2.5% quantile, and the darkgreen dotted lines are 97.5% quantile.

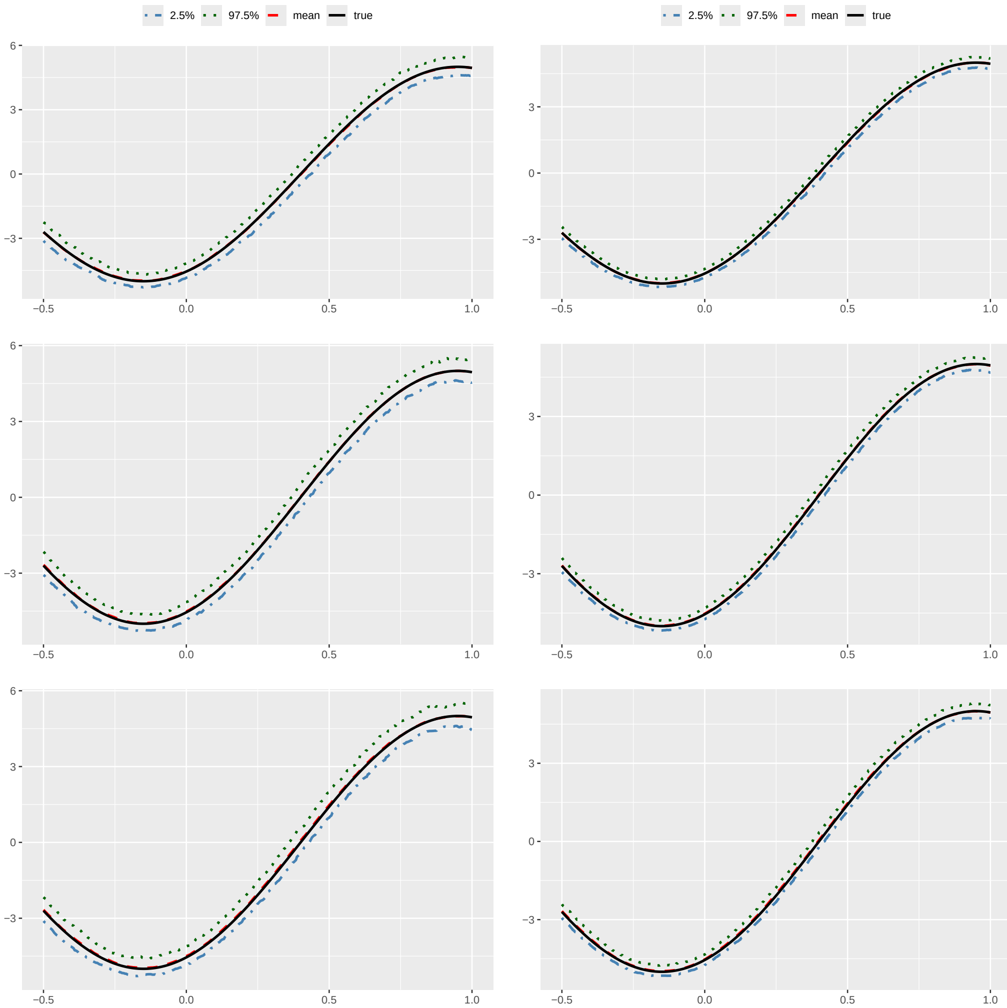

Figure S.2: Plots of estimators for $g(\cdot)$ in Simulation 2, Setting 2: The true error $\epsilon_i = \sigma(\mathbf{x}_i, \mathbf{z}_i)\{\widetilde{\epsilon}_i - F^{-1}(\tau)\}$, where $\widetilde{\epsilon}_i = F(\cdot)$ and $\sigma(\mathbf{x}_i, \mathbf{z}_i) = \sqrt{|\widetilde{\mathbf{x}}_{i2}|}$. Estimation is based on our proposed method with the first misspecified error model $\widetilde{\epsilon}_i^* \sim N(0, 1)$. Top to bottom panels: $\tau = 0.4, 0.5, 0.6$. Left panels: $n = 400$, right panels: $n = 1000$. The black sold lines are the true, the dashed red lines are the mean curve, the blue dot-dashed lines are 2.5% quantile, and the darkgreen dotted lines are 97.5% quantile.

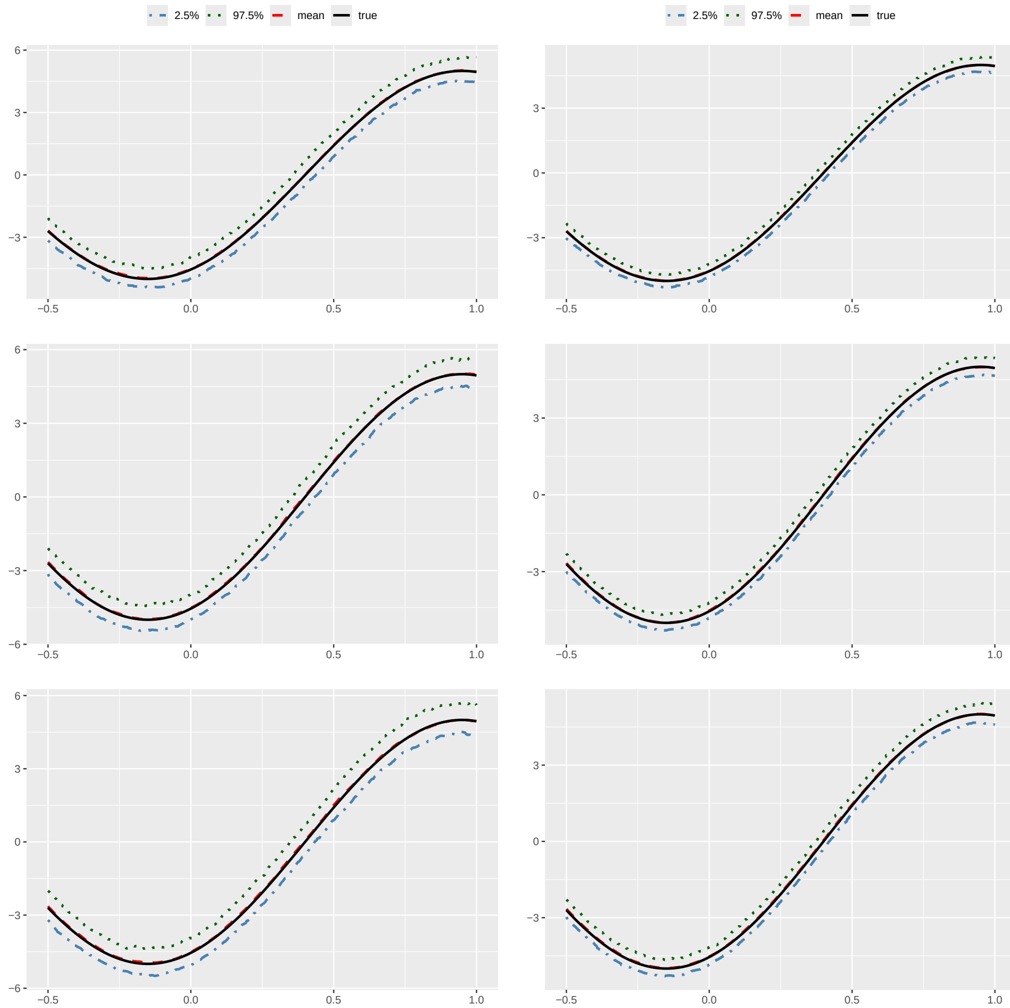

Figure S.3: Plots of estimators for $g(\cdot)$ in Simulation 2, Setting 2: The true error $\epsilon_i = \sigma(\mathbf{x}_i, \mathbf{z}_i)\{\widetilde{\epsilon}_i - F^{-1}(\tau)\}$, where $\widetilde{\epsilon}_i \sim F(\cdot)$ and $\sigma(\mathbf{x}_i, \mathbf{z}_i) = \sqrt{|\widetilde{\mathbf{x}}_{i2}|}$. Estimation is based on our proposed method with the second misspecified error model $\widetilde{\epsilon}_i^* \sim t_{(3)}$. Top to bottom panels: $\tau = 0.4, 0.5, 0.6$. Left panels: $n = 400$, right panels: $n = 1000$. The black sold lines are the true, the dashed red lines are the mean curve, the blue dot-dashed lines are 2.5% quantile, and the darkgreen dotted lines are 97.5% quantile.

Bibliography

He, X. M. and Shao, Q. M. (2000). On parameters of increasing dimensions. *Journal of Multivariate Analysis* **73**, 120–135.

Ma, S. J. and He, X. M. (2016). Inference for single-index quantile regression models with profile optimization. *Annals of Statistics* **44**, 1234–1268.

Pollard, D. (1984). *Convergence of Stochastic Processes*. Springer, New York.

Serfling, R. J. (1980). *Approximation theorems of mathematical statistics.* Wiley, New York.