SUPPLEMENT TO "EVALUATION OF TRANSPLANT BENEFITS WITH THE U.S. SCIENTIFIC REGISTRY OF TRANSPLANT RECIPIENTS BY SEMIPARAMETRIC REGRESSION OF MEAN RESIDUAL LIFE"

BY GE ZHAO[1,a], YANYUAN MA[2,b] HUAZHEN LIN[3,c] AND YI LI[4,d]

[1]*Department of Mathematics and Statistics, Portland State University,* [a]*gzhao@pdx.edu*

[2]*Department of Statistics, Pennsylvania State University,* [b]*yzm63@psu.edu*

[3]*Center of Statistical Research, Southwestern University of Finance and Economics,* [c]*linhz@swufe.edu.cn*

[4]*Department of Biostatistics, University of Michigan, Ann Arbor,* [d]*yili@umich.edu*

SUPPLEMENTARY MATERIAL

The Supplement is structured as follows. Section 1 presents the derivation of the mean residual life function based on our hazard model (1). In Section 2, we derive the estimators for both the finite-dimensional parameters and nonparametric functions. The asymptotic results of these estimators are discussed in Section 3, where Section 4 provides the proofs for prerequisite lemmas and the main theorems related to the asymptotic results. Additionally, Section 5 explores a potential relaxation of an assumption concerning event times. To facilitate implementation, Section 6 offers the code for estimating both the mean residual life function and the associated coefficients.

1. Derivations of the mean residual life function (2)**.** As the left hand side of (2) is a function of $t, \mathbf{X}, I(W \leqslant t)$ and $WI(W \leqslant t)$, we discuss it separately when $W \leqslant t$ holds and when it does not. First, when $W \leqslant t$ does hold, it means that W is observed by time t, and, therefore, the left hand side of (2) will depend on $t, \mathbf{X}, W$, or, specifically,

$$
\begin{aligned}
E\{T-t \mid T \geqslant t, \mathbf{X}, I(W \leqslant t)=1, W\} &= \frac{\int_t^\infty S(s \mid \mathbf{X}, W \leqslant t, W)ds}{S(t \mid \mathbf{X}, W \leqslant t, W)} \\
&= e^{\Lambda_T(t-W,\mathbf{X},W)} \int_t^\infty e^{-\Lambda_T(s-W,\mathbf{X},W)} ds \\
&= e^{\Lambda_T(t-W,\mathbf{X},W)} \int_{t-W}^\infty e^{-\Lambda_T(s,\mathbf{X},W)} ds
\end{aligned}
$$

based on our model, which is equal to the right hand side of (2) in this case. Second, when $W \leqslant t$ does not hold, i.e., no transplant by time t, we would effectively treat $W > t$ as $W = \infty$ when we project the future survival at time t, given the information available up to time t. Hence, the left hand side of (2) should only depend on $t, \mathbf{X}$, and

$$
\begin{aligned}
E\{T-t \mid T \geqslant t, \mathbf{X}, I(W \leqslant t)=0\} &= E\{T-t \mid T \geqslant t, \mathbf{X}, W=\infty\} \\
&= \frac{\int_t^\infty S(s \mid \mathbf{X}, W=\infty)ds}{S(t \mid \mathbf{X}, W=\infty)} \\
&= e^{\Lambda_N(t,\mathbf{X})} \int_t^\infty e^{-\Lambda_N(s,\mathbf{X})} ds,
\end{aligned}
$$

which is equal to the right hand side of (2) in this case.

2. Derivations of the estimators for the finite-dimensional parameters and nonparametric functions.

2.1. *Derivation of an efficient score function.* Let $Y(t) = I(Z \geqslant t)$ be the at risk process and $N(t) = I(Z \leqslant t)\Delta$ be the counting process. Define the filtration $\mathcal{F}_t$ to be $\sigma\{N(u), Y(u), \mathbf{X}, I(W \leqslant u), WI(W \leqslant u), 0 \leqslant u < t\}$, and, thus, $M(t) = N(t) - \int_0^t Y(s)\lambda\{s, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\}ds$ is a martingale with respect to $\mathcal{F}_t$. The tangent space for the nuisance parameters, which will be utilized in deriving our estimator, is obtained through the following process.

PROPOSITION 1. *The nuisance tangent space is $\mathcal{T} = \mathcal{T}_1 \oplus \mathcal{T}_2 \oplus \mathcal{T}_3$, where each component corresponds to $f_{\mathbf{X},W}$, m defined in (2), and λ_c, respectively. Specifically,*

$$\mathcal{T}_1 = [\mathbf{a}(\mathbf{X}, W) : E\{\mathbf{a}(\mathbf{X}, W)\} = \mathbf{0}, \mathbf{a}(\mathbf{X}, W) \in \mathcal{R}^{(p-d)d}, var\{\mathbf{a}(\mathbf{X}, W)\} < \infty],$$

$$\mathcal{T}_2 = \Bigg(\int_0^\infty \left[\frac{\mathbf{h}_1\{s, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\}}{m_1\{s, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\} + 1} - \frac{\mathbf{h}\{s, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\}}{m\{s, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\}}\right]$$
$$\times dM\{s, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\} : \forall \mathbf{h}\{z, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, I(W \leqslant z), WI(W \leqslant z)\} \in \mathcal{R}^{(p-d)d},$$
$$var\{z, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, I(W \leqslant z), WI(W \leqslant z)\} < \infty\Bigg),$$

$$\mathcal{T}_3 = \left[\int_0^\infty \mathbf{h}(s, \mathbf{X})dM_c(s, \mathbf{X}) : \forall \mathbf{h}(z, \mathbf{X}) \in \mathcal{R}^{(p-d)d}, var\{\mathbf{h}(z, \mathbf{X})\} < \infty\right].$$

The derivation of Proposition 1 is provided in Supplement 2.1.1. Taking the derivative of the logarithm of (5) with respect to $\boldsymbol{\beta}$, we obtain the score function

$$\mathbf{S}_{\boldsymbol{\beta}}\{\Delta, Z, \mathbf{X}, I(W \leqslant Z), WI(W \leqslant Z)\}$$
$$= \int_0^\infty \left[\frac{\mathbf{m}_{12}\{s, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\}}{m_1\{s, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\} + 1} - \frac{\mathbf{m}_2\{s, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\}}{m\{s, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\}}\right]$$
$$\otimes \mathbf{X}_l dM\{s, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\},$$

where

$$m_1\{s, \mathbf{v}, \cdot, \cdot)\} \equiv \partial m\{s, \mathbf{v}, \cdot, \cdot\}/\partial s = \partial m_T(s - W, \mathbf{v})/\partial s I(W \leqslant s) + \partial m_N(s, \mathbf{v})/\partial s\{1 - I(W \leqslant s)\},$$

$$\mathbf{m}_2\{s, \mathbf{v}, \cdot, \cdot\} \equiv \partial m\{s, \mathbf{v}, \cdot, \cdot\}/\partial \mathbf{v} = \partial m_T(s - W, \mathbf{v})/\partial \mathbf{v} I(W \leqslant s) + \partial m_N(s, \mathbf{v})/\partial \mathbf{v}\{1 - I(W \leqslant s)\},$$

and

$$\mathbf{m}_{12}\{s, \mathbf{v}, \cdot, \cdot\} \equiv \partial \mathbf{m}_2\{s, \mathbf{v}, \cdot, \cdot\}/\partial s = \partial^2 m_T(s - W, \mathbf{v})/\partial s \partial \mathbf{v} I(W \leqslant s) + \partial^2 m_N(s, \mathbf{v})/\partial s \partial \mathbf{v}\{1 - I(W \leqslant s)\}.$$

We can verify that, at $\boldsymbol{\beta}_0$, $\mathbf{S}_{\boldsymbol{\beta}}\{\Delta, Z, \mathbf{X}, I(W \leqslant Z), WI(W \leqslant Z)\} \perp \mathcal{T}_1$ and $\mathbf{S}_{\boldsymbol{\beta}}\{\Delta, Z, \mathbf{X}, I(W \leqslant Z), WI(W \leqslant Z)\} \perp \mathcal{T}_3$ due to the martingale properties. The next steps are to look for an efficient score. To proceed, we first project $\mathbf{S}_{\boldsymbol{\beta}}\{\Delta, Z, \mathbf{X}, I(W \leqslant Z), WI(W \leqslant Z)\}$ at $\boldsymbol{\beta}_0$ to $\mathcal{T}_2$, and we search for $\mathbf{h}^*\{s, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\}$ such that

$$\mathbf{S}_{\text{eff}}\{\Delta, Z, \mathbf{X}, I(W \leqslant Z), WI(W \leqslant Z)\}$$
$$= \mathbf{S}_{\boldsymbol{\beta}}\{\Delta, Z, \mathbf{X}, I(W \leqslant Z), WI(W \leqslant Z)\} - \int_0^\infty \left[\frac{\mathbf{h}_1^*\{s, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\}}{m_1\{s, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\} + 1}\right.$$

$$- \frac{\mathbf{h}^*\{s, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\}}{m\{s, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\}} \Bigg] dM\{s, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\}$$

$$= \int_0^\infty \Bigg[\frac{\mathbf{m}_{12}\{s, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\} \otimes \mathbf{X}_l - \mathbf{h}_1^*\{s, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\}}{m_1\{s, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\} + 1}$$

$$- \frac{\mathbf{m}_2\{s, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\} \otimes \mathbf{X}_l - \mathbf{h}^*\{s, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\}}{m\{s, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\}} \Bigg]$$

$$\times dM\{s, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\},$$

where $\mathbf{h}_1^*\{s, \mathbf{v}, \cdot, \cdot\} = \partial \mathbf{h}^*\{s, \mathbf{v}, \cdot, \cdot\}/\partial s$, is orthogonal to $\mathcal{T}_2$. It implies that, for any $\mathbf{h}\{s, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\}$, it holds that

(S1)

$$0 = E\left(\int_0^\infty \mathbf{a}\{s, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\}^{\mathrm{T}} \Bigg[\frac{\mathbf{h}_1\{s, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\}}{m_1\{s, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\} + 1} - \frac{\mathbf{h}\{s, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\}}{m\{s, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\}} \Bigg] ds \right),$$

where

(S2)

$$\mathbf{a}\{s, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\}$$

$$\equiv E\Bigg(\Bigg[\frac{\mathbf{m}_{12}\{s, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\} \otimes \mathbf{X}_l - \mathbf{h}_1^*\{s, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\}}{m_1\{s, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\} + 1}$$

$$- \frac{\mathbf{m}_2\{s, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\} \otimes \mathbf{X}_l - \mathbf{h}^*\{s, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\}}{m\{s, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\}} \Bigg]$$

$$\times S_c(s, \mathbf{X}) \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X} \Bigg) S\{s, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\}$$

$$\times \frac{m_1\{s, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\} + 1}{m\{s, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\}},$$

and $\mathbf{h}_1\{s, \mathbf{v}, \cdot, \cdot\} = \partial \mathbf{h}\{s, \mathbf{v}, \cdot, \cdot\}/\partial s$. We can choose any $\mathbf{h}\{s, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\}$ function. Specifically, by letting $\mathbf{h}\{s, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\} = 0$ for $s < t$ and $\mathbf{h}\{s, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\} = \mathbf{c}(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X})$ for $s \geqslant t$ with an arbitrary function $\mathbf{c}(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X})$, we obtain that

$$\frac{\mathbf{a}\{t, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}, I(W \leqslant t), WI(W \leqslant t)\}}{m_1\{t, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}, I(W \leqslant t), WI(W \leqslant t)\} + 1} - \int_t^\infty \frac{\mathbf{a}\{s, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\}}{m\{s, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\}} ds = \mathbf{0}.$$

Solving this integral equation leads to

$$\mathbf{a}\{t, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}, I(W \leqslant t), WI(W \leqslant t)\}$$

$$= \{m_1\{t, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}, I(W \leqslant t), WI(W \leqslant t)\} + 1\}$$

$$\times \exp\left[- \int_0^t \frac{m_1\{s, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\} + 1}{m\{s, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\}} ds \right] \mathbf{c}(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}),$$

for any function $\mathbf{c}(\cdot)$. Thus, reusing (S2), we require that for all $\mathbf{h}\{t,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X},I(W\leqslant t),WI(W\leqslant t)\}$,

$$
\begin{aligned}
0 = E\Bigg(\int_0^\infty & [m_1\{t,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X},I(W\leqslant t),WI(W\leqslant t)\}+1] \\
&\times \exp\left[-\int_0^t \frac{m_1\{s,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X},I(W\leqslant s),WI(W\leqslant s)\}+1}{m\{s,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X},I(W\leqslant s),WI(W\leqslant s)\}}ds\right]\mathbf{c}(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X})^{\mathrm{T}} \\
&\times\left[\frac{\mathbf{h}_1\{t,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X},I(W\leqslant t),WI(W\leqslant t)\}}{m_1\{t,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X},I(W\leqslant t),WI(W\leqslant t)\}+1}-\frac{\mathbf{h}\{t,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X},I(W\leqslant t),WI(W\leqslant t)\}}{m\{t,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X},I(W\leqslant t),WI(W\leqslant t)\}}\right]dt\Bigg) \\
= -E&\left[\mathbf{c}(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X})^{\mathrm{T}}\mathbf{h}\{0,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X},I(W\leqslant 0),WI(W\leqslant 0)\}\right].
\end{aligned}
$$

Letting $\mathbf{h}\{0,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X},I(W\leqslant 0),WI(W\leqslant 0)\}=\mathbf{c}(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X})$ yields the only possibility of $\mathbf{c}(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X})=\mathbf{0}$, hence $\mathbf{a}\{t,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X},I(W\leqslant t),WI(W\leqslant t)\}=\mathbf{0}$. Inserting the expression of $\mathbf{a}\{t,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X},I(W\leqslant t),WI(W\leqslant t)\}$ into (S3), we have

$$
\begin{aligned}
&\frac{\mathbf{h}_1^*\{t,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X},I(W\leqslant t),WI(W\leqslant t)\}}{m_1\{t,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X},I(W\leqslant t),WI(W\leqslant t)\}+1}-\frac{\mathbf{h}^*\{t,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X},I(W\leqslant t),WI(W\leqslant t)\}}{m\{t,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X},I(W\leqslant t),WI(W\leqslant t)\}} \\
&= \mathbf{b}\{t,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X},I(W\leqslant t),WI(W\leqslant t)\},
\end{aligned}
$$

where

$$
\begin{aligned}
\mathbf{b}\{t,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X},I(W\leqslant t),WI(W\leqslant t)\} &= \Bigg\{\frac{\mathbf{m}_{12}\{t,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X},I(W\leqslant t),WI(W\leqslant t)\}}{m_1\{t,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X},I(W\leqslant t),WI(W\leqslant t)\}+1} \\
&-\frac{\mathbf{m}_2\{t,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X},I(W\leqslant t),WI(W\leqslant t)\}}{m\{t,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X},I(W\leqslant t),WI(W\leqslant t)\}}\Bigg\}\otimes\frac{E\left\{\mathbf{X}_l S_c(t,\mathbf{X})\mid\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X},W\right\}}{E\left\{S_c(t,\mathbf{X})\mid\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X},W\right\}}.
\end{aligned}
$$

Thus, we have obtained an efficient score as follows:

$$
\begin{aligned}
&\mathbf{S}_{\mathrm{eff}}\{\Delta,Z,\mathbf{X},I(W\leqslant Z),WI(W\leqslant Z)\} \\
&= \int_0^\infty\left[\frac{\mathbf{m}_{12}\{s,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X},I(W\leqslant s),WI(W\leqslant s)\}}{m_1\{s,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X},I(W\leqslant s),WI(W\leqslant s)\}+1}-\frac{\mathbf{m}_2\{s,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X},I(W\leqslant s),WI(W\leqslant s)\}}{m\{s,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X},I(W\leqslant s),WI(W\leqslant s)\}}\right] \\
&\otimes\left[\mathbf{X}_l-\frac{E\left\{\mathbf{X}_l S_c(s,\mathbf{X})\mid\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X},W\right\}}{E\left\{S_c(s,\mathbf{X})\mid\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X},W\right\}}\right]dM\{s,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X},I(W\leqslant s),WI(W\leqslant s)\}.
\end{aligned}
$$

2.1.1. *Proof of Proposition 1* . Proof: Let $\mathcal{T}_1$, $\mathcal{T}_2$ and $\mathcal{T}_3$ be the nuisance tangent spaces corresponding to $f_{\mathbf{X},W}$, m and λ_c respectively. The result of $\mathcal{T}_1$ follows obviously. To obtain $\mathcal{T}_2$, let $m\{z,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},I(W\leqslant z),WI(W\leqslant z)\}+\boldsymbol{\gamma}^{\mathrm{T}}\mathbf{h}\{z,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},I(W\leqslant z),WI(W\leqslant z)\}$ be a submodel of $m\{z,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},I(W\leqslant z),WI(W\leqslant z)\}$, where

$$
\mathbf{h}\{z,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},I(W\leqslant z),WI(W\leqslant z)\}\in\mathcal{R}^{(p-d)d}
$$

with var$[\mathbf{h}\{z,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},I(W\leqslant z),WI(W\leqslant z)\}]<\infty$. We then differentiate the log of (5) with respect to $\boldsymbol{\gamma}$ and evaluate it at $\boldsymbol{\gamma}=0$. Then, $\mathcal{T}_2$ is

$$
\begin{aligned}
&\Delta\left\{\frac{\mathbf{h}_1\{z,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},I(W\leqslant z),WI(W\leqslant z)\}}{m_1\{z,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},I(W\leqslant z),WI(W\leqslant z)\}+1}-\frac{\mathbf{h}\{z,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},I(W\leqslant z),WI(W\leqslant z)\}}{m\{z,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},I(W\leqslant z),WI(W\leqslant z)\}}\right\} \\
&-\int_0^z\frac{\mathbf{h}_1\{s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},I(W\leqslant s),WI(W\leqslant s)\}m\{s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},I(W\leqslant s),WI(W\leqslant s)\}}{m^2\{s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},I(W\leqslant s),WI(W\leqslant s)\}}
\end{aligned}
$$

$$
\begin{aligned}
&-\frac{\{m_1\{s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},I(W\leqslant s),WI(W\leqslant s)\}+1\}\mathbf{h}\{s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},I(W\leqslant s),WI(W\leqslant s)\}}{m^2\{s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},I(W\leqslant s),WI(W\leqslant s)\}}dt\\
&=\Delta\left\{\frac{\mathbf{h}_1\{z,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},I(W\leqslant z),WI(W\leqslant z)\}}{m_1\{z,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},I(W\leqslant z),WI(W\leqslant z)\}+1}-\frac{\mathbf{h}\{z,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},I(W\leqslant z),WI(W\leqslant z)\}}{m\{z,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},I(W\leqslant z),WI(W\leqslant z)\}}\right\}\\
&-\int_0^z\left[\frac{\mathbf{h}_1\{s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},I(W\leqslant s),WI(W\leqslant s)\}}{m_1\{s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},I(W\leqslant s),WI(W\leqslant s)\}+1}-\frac{\mathbf{h}\{s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},I(W\leqslant s),WI(W\leqslant s)\}}{m\{s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},I(W\leqslant s),WI(W\leqslant s)\}}\right]\\
&\quad\lambda\{s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},I(W\leqslant s),WI(W\leqslant s)\}ds\\
&=\int_0^\infty\left[\frac{\mathbf{h}_1\{s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},I(W\leqslant s),WI(W\leqslant s)\}}{m_1\{s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},I(W\leqslant s),WI(W\leqslant s)\}+1}-\frac{\mathbf{h}\{s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},I(W\leqslant s),WI(W\leqslant s)\}}{m\{s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},I(W\leqslant s),WI(W\leqslant s)\}}\right]\\
&\quad dM\{s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},I(W\leqslant s),WI(W\leqslant s)\}.
\end{aligned}
$$

To obtain $\mathcal{T}_3$, we let $\lambda_c(t,\mathbf{X})\{1+\boldsymbol{\gamma}^{\mathrm{T}}\mathbf{h}(t,\mathbf{X})\}$ be a submodel of $\lambda_c(t,\mathbf{X})$, where $\mathbf{h}(z,\mathbf{X})\in\mathcal{R}^{(p-d)d}$, with var$\{\mathbf{h}(z,\mathbf{X})\}<\infty$. We then obtain $\mathcal{T}_3$ as follows

$$
\begin{aligned}
\frac{\partial\log f(\mathbf{X},Z,\Delta)}{\partial\boldsymbol{\gamma}}|_{\boldsymbol{\gamma}=0}&=(1-\Delta)\mathbf{h}(Z,\mathbf{X})-\int_0^Z\mathbf{h}(s,\mathbf{X})\lambda_c(s,\mathbf{X})ds\\
&=\int_0^\infty\mathbf{h}(s,\mathbf{X})dM_c(s,\mathbf{X}),
\end{aligned}
$$

where $M_c(t,\mathbf{X})=N_c(t)-\int_0^t I(Z\geqslant s)\lambda_c(s,\mathbf{X})ds$ is a martingale by Theorem 1.3.2 in Fleming and Harrington (1991). Because $\lambda_c(t,\mathbf{X})$ can be any positive function, so can be $\mathbf{h}(s,\mathbf{X})$, which leads to the form of $\mathcal{T}_3$.

By taking conditional expectations given $\mathbf{X}$ and W, it follows that $\mathcal{T}_1\perp\mathcal{T}_2$ and $\mathcal{T}_1\perp\mathcal{T}_3$. Further, $\mathcal{T}_2\perp\mathcal{T}_3$ because the martingale integrals associated with $M\{s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},I(W\leqslant s),WI(W\leqslant s)\}$ and $M_c(s,\mathbf{X})$ are independent conditional on $\mathbf{X}$ due to $T\perp\!\!\!\perp C\mid W,\mathbf{X}$. □

2.1.2. *Proof of Equations (7) and (8).* Proof: First, we note that, when $t\leqslant\tau$,

$$
\begin{aligned}
E\left\{\mathbf{X}_lY(t)\mid\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},W\right\}&=E[E\left\{\mathbf{X}_lI(T\geqslant t)I(C\geqslant t)\mid\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},\mathbf{X},W\right\}\mid\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},W]\\
&=E[E\left\{\mathbf{X}_lI(T\geqslant t)I(C\geqslant t)\mid\mathbf{X},W\right\}\mid\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},W]\\
&=E\left[\mathbf{X}_lS\{t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},I(W\leqslant t),WI(W\leqslant t)\}S_c(t,\mathbf{X})\mid\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},W\right]\\
&=S\{t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},I(W\leqslant t),WI(W\leqslant t)\}E\left\{\mathbf{X}_lS_c(t,\mathbf{X})\mid\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\right\},
\end{aligned}
$$

where the second to last equality holds because of $T\perp\!\!\!\perp C\mid\mathbf{X},W$ and that pr$(T\geqslant t|\mathbf{X})$ is a function of $\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}$ and W. Similarly, for $t\leqslant\tau$, we obtain that

$$
E\left\{Y(t)\mid\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},W\right\}=S\{t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},I(W\leqslant t),WI(W\leqslant t)\}E\left\{S_c(t,\mathbf{X})\mid\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\right\}.
$$

Hence, when $t\leqslant\tau$,

$$
\frac{E\left\{\mathbf{X}_lY(t)\mid\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},W\right\}}{E\left\{Y(t)\mid\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},W\right\}}=\frac{E\left\{\mathbf{X}_lS_c(t,\mathbf{X})\mid\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\right\}}{E\left\{S_c(t,\mathbf{X})\mid\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\right\}}. \tag{S3}
$$

Second, when $t>\tau$, $Y(t)=0$ and $S_c(t,\mathbf{X})=0$. Hence, we have a $0/0$ scenario in (S3), in which case, we define the left hand side of (S3) to be

$$
\frac{E\left\{\mathbf{X}_lS_c(\tau,\mathbf{X})\mid\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\right\}}{E\left\{S_c(\tau,\mathbf{X})\mid\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\right\}}=\frac{E\left\{\mathbf{X}_lp(\mathbf{X})\mid\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\right\}}{E\left\{p(\mathbf{X})\mid\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\right\}},
$$

a time-invariant constant, where $p(\mathbf{X})$ is as defined in Section 3. Therefore, (6) and (7) hold over $[0,\infty)$, when $\boldsymbol{\beta}=\boldsymbol{\beta}_0$, the truth.

With a generic $\boldsymbol{\beta}$, (S3) leads to

$$E\left(\int_0^\infty \mathbf{g}\{s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},I(W\leqslant s),WI(W\leqslant s)\}\otimes\left[\mathbf{X}_l-\frac{E\left\{\mathbf{X}_lS_c(s,\mathbf{X})\mid\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\right\}}{E\left\{S_c(s,\mathbf{X})\mid\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\right\}}\right]\right.$$
$$\left.\times\, Y(s)\lambda_0\{s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},I(W\leqslant s),WI(W\leqslant s)\}ds\right)$$
$$=E\left(\int_0^\infty \mathbf{g}\{s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},I(W\leqslant s),WI(W\leqslant s)\}\otimes\Big[E\{\mathbf{X}_lY(s)\mid\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},W\}\right.$$
$$\left.-\frac{E\left\{\mathbf{X}_lS_c(s,\mathbf{X})\mid\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\right\}}{E\left\{S_c(s,\mathbf{X})\mid\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\right\}}E\{Y(s)\mid\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},I(W\leqslant s),WI(W\leqslant s)\}\Big]\lambda_0(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})ds\right)$$
$$=\mathbf{0},$$

as the quantity inside the square bracket is zero. In addition,

$$E\left(\int_0^\infty \mathbf{g}\{s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},I(W\leqslant s),WI(W\leqslant s)\}\right.$$
$$\left.\otimes\left[\mathbf{X}_l-\frac{E\left\{\mathbf{X}_lS_c(s,\mathbf{X})\mid\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\right\}}{E\left\{S_c(s,\mathbf{X})\mid\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\right\}}\right]dM\{s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},I(W\leqslant s),WI(W\leqslant s)\}\right)=\mathbf{0},$$

because $dM\{s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},I(W\leqslant s),WI(W\leqslant s)\}=dN(s)-Y(s)\lambda_0\{s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},I(W\leqslant s),W I(W\leqslant s)\}ds$ is a martingale. Therefore, we have

$$E\left(\int_0^\infty \mathbf{g}\{s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},I(W\leqslant s),WI(W\leqslant s)\}\otimes\left[\mathbf{X}_l-\frac{E\left\{\mathbf{X}_lS_c(s,\mathbf{X})\mid\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\right\}}{E\left\{S_c(s,\mathbf{X})\mid\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\right\}}\right]dN(s)\right)=\mathbf{0}.$$

Hence, (8) holds when $\boldsymbol{\beta}=\boldsymbol{\beta}_0$, the truth. □

2.2. *Explicit Expressions of Nonparametric Estimators.*

2.2.1. *Nonparametric Estimators of Hazard and Mean Residual Life Functions and Their Derivatives.*

The estimators of Λ_T and Λ_N have a similar form, differing only in the samples utilized and an additional argument of W in the estimator of Λ_T. Also, the nonparametric estimators for (9) are given by

$$\widehat{E}\left\{Y_i(Z_i)\mid\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\right\}$$
$$=\frac{\sum_{j=1}^n K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_j-\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i,W_j-W_i)I(Z_j\geqslant Z_i)I(W_j\leqslant Z_j)}{\sum_{j=1}^n K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_j-\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i,W_j-W_i)I(W_j\leqslant Z_j)}I(W_i\leqslant Z_i)$$
$$+\frac{\sum_{j=1}^n K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_j-\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i)I(Z_j\geqslant Z_i)\{1-I(W_j\leqslant Z_j)\}}{\sum_{j=1}^n K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_j-\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i)\{1-I(W_j\leqslant Z_j)\}}\{1-I(W_i\leqslant Z_i)\},\tag{S4}$$

$$\widehat{E}\left\{\mathbf{X}_{li}Y_i(Z_i)\mid\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\right\}$$
$$=\frac{\sum_{j=1}^n K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_j-\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i,W_j-W_i)\mathbf{X}_{lj}I(Z_j\geqslant Z_i)I(W_j\leqslant Z_j)}{\sum_{j=1}^n K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_j-\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i,W_j-W_i)I(W_j\leqslant Z_j)}I(W_i\leqslant Z_i)$$
$$+\frac{\sum_{j=1}^n K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_j-\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i)\mathbf{X}_{lj}I(Z_j\geqslant Z_i)\{1-I(W_j\leqslant Z_j)\}}{\sum_{j=1}^n K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_j-\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i)\{1-I(W_j\leqslant Z_j)\}}\{1-I(W_i\leqslant Z_i)\}.\tag{S5}$$

We propose a general form of the nonparametric estimators of $\boldsymbol{\Lambda}_2(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})$, $\lambda(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})$, $\boldsymbol{\lambda}_2(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})$, where $\boldsymbol{\Lambda}_2(t,\mathbf{v})\equiv\partial\Lambda(t,\mathbf{v})/\partial\mathbf{v}$ and $\boldsymbol{\lambda}_2(t,\mathbf{v})\equiv\partial\lambda(t,\mathbf{v})/\partial\mathbf{v}$.

$$\begin{aligned}
&\widehat{\boldsymbol{\Lambda}}_2(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})\\
&=-\sum_{i=1}^{n}\frac{I(Z_i\leqslant t)\Delta_i\mathbf{K}_h'(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i-\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})}{\sum_{j=1}^{n}I(Z_j\geqslant Z_i)K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_j-\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})}\\
&+\sum_{i=1}^{n}I(Z_i\leqslant t)\Delta_iK_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i-\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})\frac{\sum_{j=1}^{n}I(Z_j\geqslant Z_i)\mathbf{K}_h'(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_j-\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})}{\{\sum_{j=1}^{n}I(Z_j\geqslant Z_i)K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_j-\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})\}^2},
\end{aligned}\tag{S6}$$

$$\begin{aligned}
&\widehat{\lambda}(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})\\
&=\sum_{i=1}^{n}\frac{K_b(Z_i-t)\Delta_iK_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i-\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})}{\sum_{j=1}^{n}I(Z_j\geqslant Z_i)K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_j-\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})},
\end{aligned}\tag{S7}$$

$$\begin{aligned}
&\widehat{\boldsymbol{\lambda}}_2(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})\\
&=-\sum_{i=1}^{n}\frac{K_b(Z_i-t)\Delta_i\mathbf{K}_h'(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i-\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})}{\sum_{j=1}^{n}I(Z_j\geqslant Z_i)K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_j-\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})}\\
&+\sum_{i=1}^{n}K_b(Z_i-t)\Delta_iK_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i-\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})\frac{\sum_{j=1}^{n}I(Z_j\geqslant Z_i)\mathbf{K}_h'(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_j-\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})}{\{\sum_{j=1}^{n}I(Z_j\geqslant Z_i)K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_j-\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})\}^2},
\end{aligned}\tag{S8}$$

The estimators of $m_N(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x})$, $m_T(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x})$, $m_1\{t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},I(W\leqslant t),WI(W\leqslant t)\}$, $\mathbf{m}_2\{t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, I(W\leqslant t),WI(W\leqslant t)\}$, and $\mathbf{m}_{12}\{t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},I(W\leqslant t),WI(W\leqslant t)\}$ are

$$\begin{aligned}
&\widehat{m}_N(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x})=e^{\widehat{\Lambda}_N(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x})}\int_t^\tau e^{-\widehat{\Lambda}_N(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x})}ds,\\
&\widehat{m}_T(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x})=e^{\widehat{\Lambda}_T(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x})}\int_t^\tau e^{-\widehat{\Lambda}_T(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x})}ds,\\
&\widehat{m}_1\{t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},I(W\leqslant t),WI(W\leqslant t)\}\\
&=\widehat{\lambda}_T(t-W,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})e^{\widehat{\Lambda}_T(t-W,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})}\int_t^\tau e^{-\widehat{\Lambda}_T(s-W,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})}dsI(W\leqslant t)\\
&+\widehat{\lambda}_N(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})e^{\widehat{\Lambda}_N(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})}\int_t^\tau e^{-\widehat{\Lambda}_N(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})}ds\{1-I(W\leqslant t)\}-1,\\
&\widehat{\mathbf{m}}_2\{t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},I(W\leqslant t),WI(W\leqslant t)\}\\
&=\Bigg\{\widehat{\boldsymbol{\Lambda}}_{T2}(t-W,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})e^{\widehat{\Lambda}_T(t-W,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})}\int_t^\tau e^{-\widehat{\Lambda}_T(s-W,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})}ds\\
&\quad-e^{\widehat{\Lambda}_T(t-W,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})}\int_t^\tau\widehat{\boldsymbol{\Lambda}}_{T2}(s-W,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})e^{-\widehat{\Lambda}_T(s-W,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})}ds\Bigg\}I(W\leqslant t)\\
&+\Bigg\{\widehat{\boldsymbol{\Lambda}}_{N2}(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})e^{\widehat{\Lambda}_N(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})}\int_t^\tau e^{-\widehat{\Lambda}_N(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})}ds\\
&\quad-e^{\widehat{\Lambda}_N(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})}\int_t^\tau\widehat{\boldsymbol{\Lambda}}_{N2}(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})e^{-\widehat{\Lambda}_N(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})}ds\Bigg\}\{1-I(W\leqslant t)\},\\
&\widehat{\mathbf{m}}_{12}\{t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},I(W\leqslant t),WI(W\leqslant t)\}
\end{aligned}$$

$$
\begin{aligned}
&= \left\{ \widehat{\boldsymbol{\lambda}}_{T2}(t-W,\boldsymbol{\beta}^{\rm T}\mathbf{X}) e^{\widehat{\Lambda}_T(t-W,\boldsymbol{\beta}^{\rm T}\mathbf{X})} \int_t^\tau e^{-\widehat{\Lambda}_T(s-W,\boldsymbol{\beta}^{\rm T}\mathbf{X})} ds \right.\\
&\quad - \widehat{\lambda}_T(t-W,\boldsymbol{\beta}^{\rm T}\mathbf{X}) e^{\widehat{\Lambda}_T(t-W,\boldsymbol{\beta}^{\rm T}\mathbf{X})} \int_t^\tau \widehat{\boldsymbol{\Lambda}}_{T2}(s-W,\boldsymbol{\beta}^{\rm T}\mathbf{X}) e^{-\widehat{\Lambda}_T(s-W,\boldsymbol{\beta}^{\rm T}\mathbf{X})} ds \\
&\quad \left. + \widehat{\lambda}_T(t-W,\boldsymbol{\beta}^{\rm T}\mathbf{X}) \widehat{\boldsymbol{\Lambda}}_{T2}(t-W,\boldsymbol{\beta}^{\rm T}\mathbf{X}) e^{\widehat{\Lambda}_T(t-W,\boldsymbol{\beta}^{\rm T}\mathbf{X})} \int_t^\tau e^{-\widehat{\Lambda}_T(s-W,\boldsymbol{\beta}^{\rm T}\mathbf{X})} ds \right\} I(W \leqslant t) \\
&+ \left\{ \widehat{\boldsymbol{\lambda}}_{N2}(t,\boldsymbol{\beta}^{\rm T}\mathbf{X}) e^{\widehat{\Lambda}_N(t,\boldsymbol{\beta}^{\rm T}\mathbf{X})} \int_t^\tau e^{-\widehat{\Lambda}_N(s,\boldsymbol{\beta}^{\rm T}\mathbf{X})} ds \right.\\
&\quad - \widehat{\lambda}_N\{t,\boldsymbol{\beta}^{\rm T}\mathbf{X}, I(W\leqslant t), WI(W\leqslant t)\} e^{\widehat{\Lambda}_N(t,\boldsymbol{\beta}^{\rm T}\mathbf{X})} \int_t^\tau \widehat{\boldsymbol{\Lambda}}_{N2}(s,\boldsymbol{\beta}^{\rm T}\mathbf{X}) e^{-\widehat{\Lambda}_N(s,\boldsymbol{\beta}^{\rm T}\mathbf{X})} ds. \\
&\quad \left. + \widehat{\lambda}_N(t,\boldsymbol{\beta}^{\rm T}\mathbf{X}) \widehat{\boldsymbol{\Lambda}}_{N2}(t,\boldsymbol{\beta}^{\rm T}\mathbf{X}) e^{\widehat{\Lambda}_N(t,\boldsymbol{\beta}^{\rm T}\mathbf{X})} \int_t^\tau e^{-\widehat{\Lambda}_N(s,\boldsymbol{\beta}^{\rm T}\mathbf{X})} ds \right\} \{1 - I(W\leqslant t)\}.
\end{aligned}
$$

3. Discussion of asymptotic properties and semiparametric efficiency. We briefly summarize the implications and utility of the three major theorems. First, Theorems 1 and 2 establish the root-n consistency and asymptotic normality, respectively, of the profile parameter estimator $\widehat{\boldsymbol{\beta}}$. The asymptotic variance of $\widehat{\boldsymbol{\beta}}$ can be estimated as follows. With $\mathbf{S}_{\rm eff}$ being a martingale,

$$
\begin{aligned}
& E[\mathbf{S}_{\rm eff}^{\otimes 2}\{Z,\boldsymbol{\beta}^{\rm T}\mathbf{X}, I(W\leqslant Z), WI(W\leqslant Z)\}] \\
&= E\left\{ \int_0^\infty \left(\left[\frac{\mathbf{m}_{12}\{s,\boldsymbol{\beta}^{\rm T}\mathbf{X}, I(W\leqslant s), WI(W\leqslant s)\}}{m_1\{s,\boldsymbol{\beta}^{\rm T}\mathbf{X}, I(W\leqslant s), WI(W\leqslant s)\}+1} \right.\right.\right.\\
&\quad \left. - \frac{\mathbf{m}_2\{s,\boldsymbol{\beta}^{\rm T}\mathbf{X}, I(W\leqslant s), WI(W\leqslant s)\}}{m\{s,\boldsymbol{\beta}^{\rm T}\mathbf{X}, I(W\leqslant s), WI(W\leqslant s)\}} \right] \\
&\quad \left.\left. \otimes \left[\mathbf{X}_l - \frac{E\{\mathbf{X}_l S_c(s,\mathbf{X}) \mid \boldsymbol{\beta}^{\rm T}\mathbf{X}\}}{E\{S_c(s,\mathbf{X}) \mid \boldsymbol{\beta}^{\rm T}\mathbf{X}\}} \right] \right)^{\otimes 2} dN(s) \right\},
\end{aligned}
$$

which leads to a consistent estimator of $E[\mathbf{S}_{\rm eff}^{\otimes 2}\{s,\boldsymbol{\beta}^{\rm T}\mathbf{X}, I(W\leqslant s), WI(W\leqslant s)\}]$ as follows

$$
\begin{aligned}
& \frac{1}{n}\sum_{i=1}^n \delta_i \left(\frac{\widehat{\boldsymbol{\lambda}}_2\{z_i,\widehat{\boldsymbol{\beta}}^{\rm T}\mathbf{x}_i, I(w_i\leqslant z_i), w_iI(w_i\leqslant z_i)\}}{\widehat{\lambda}\{z_i,\widehat{\boldsymbol{\beta}}^{\rm T}\mathbf{x}_i, I(w_i\leqslant z_i), w_iI(w_i\leqslant z_i)\}} \right.\\
& \left. \otimes \left[\mathbf{x}_{il} - \frac{\widehat{E}\left\{\mathbf{X}_l Y(z_i) \mid \widehat{\boldsymbol{\beta}}^{\rm T}\mathbf{x}_i, I(w_i\leqslant z_i), w_iI(w_i\leqslant z_i)\right\}}{\widehat{E}\left\{Y(z_i) \mid \widehat{\boldsymbol{\beta}}^{\rm T}\mathbf{x}_i, I(w_i\leqslant z_i), w_iI(w_i\leqslant z_i)\right\}} \right] \right)^{\otimes 2}.
\end{aligned}
$$

Here $\widehat{E}\left\{Y(z_i) \mid \widehat{\boldsymbol{\beta}}^{\rm T}\mathbf{x}_i, I(w_i\leqslant z_i), w_iI(w_i\leqslant z_i)\right\}$, $\widehat{E}\left\{\mathbf{X}_l Y(z_i) \mid \widehat{\boldsymbol{\beta}}^{\rm T}\mathbf{x}_i, I(w_i\leqslant z_i), w_iI(w_i\leqslant z_i)\right\}$, $\widehat{\lambda}\{z_i,\widehat{\boldsymbol{\beta}}^{\rm T}\mathbf{x}_i, I(w_i\leqslant z_i), w_iI(w_i\leqslant z_i)\}$ and $\widehat{\boldsymbol{\lambda}}_2\{z_i,\widehat{\boldsymbol{\beta}}^{\rm T}\mathbf{x}_i, I(w_i\leqslant z_i), w_iI(w_i\leqslant z_i)\}$ are given in (S4), (S5), (S7) and (S8) respectively.

Second, Theorem 3 implies that we can estimate the variance $\sigma_T^2(t,\boldsymbol{\beta}^{\rm T}\mathbf{x},w)$, $\sigma_N^2(t,\boldsymbol{\beta}^{\rm T}\mathbf{x})$, without estimating λ or $f_{\boldsymbol{\beta}^{\rm T}\mathbf{X}}(\boldsymbol{\beta}^{\rm T}\mathbf{x})$. Specifically, we have the following estimators:

$$
\widehat{\sigma}_T^2(t,\widehat{\boldsymbol{\beta}}^{\rm T}\mathbf{x},w)
$$

$$= I(w \leqslant t) e^{2\widehat{\Lambda}_T(t, \widehat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{x}, w)} \int K^2(u) du \sum_{i=1}^{n} I(w_i \leqslant t)$$

$$\left(\frac{\widehat{\Lambda}_T(t_{(i)}, \widehat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{x}, w) - \widehat{\Lambda}_T(t_{(i-1)}, \widehat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{x}, w)}{1/n \sum_{j=1}^{n} Y_j(t_{(i-1)}) K_h(\widehat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{x}_j - \widehat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{x}, w_j - w)} \right.$$

$$\left[I(t_{(i-1)} < t) \sum_{j=1}^{n} I(t_{(j)} > t) e^{-\widehat{\Lambda}_T(t_{(j-1)}, \widehat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{x}, w)} \{t_{(j)} - \max(t, t_{(j-1)})\} \right.$$

$$\left.\left. + \sum_{j=1}^{n} I\{t_{(j)} > \max(t, t_{(i-1)})\} e^{-\widehat{\Lambda}_T(t_{(j-1)}, \widehat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{x}, w)} \{t_{(j)} - \max(t_{(i-1)}, t_{(j-1)})\} \right]^2 \right),$$

$$\widehat{\sigma}_N^2(t, \widehat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{x})$$

$$= \{1 - I(w \leqslant t)\} e^{2\widehat{\Lambda}_N(t, \widehat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{x})} \int K^2(u) du \sum_{i=1}^{n} \{1 - I(w_i \leqslant t)\}$$

$$\times \left(\frac{\widehat{\Lambda}_N(t_{(i)}, \widehat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{x}) - \widehat{\Lambda}_N(t_{(i-1)}, \widehat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{x})}{1/n \sum_{j=1}^{n} Y_j(t_{(i-1)}) K_h(\widehat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{x}_j - \widehat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{x})} \right.$$

$$\left[I(t_{(i-1)} < t) \sum_{j=1}^{n} I(t_{(j)} > t) e^{-\widehat{\Lambda}_N(t_{(j-1)}, \widehat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{x})} \{t_{(j)} - \max(t, t_{(j-1)})\} \right.$$

$$\left.\left. + \sum_{j=1}^{n} I\{t_{(j)} > \max(t, t_{(i-1)})\} e^{-\widehat{\Lambda}_N(t_{(j-1)}, \widehat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{x})} \{t_{(j)} - \max(t_{(i-1)}, t_{(j-1)})\} \right]^2 \right),$$

where $t_{(1)} < t_{(2)} < \cdots < t_{(n_T)}$ are the observed event times and $t_{(0)} = 0$.

4. Proofs of Theorems 1–3. This section includes the regularity conditions, useful lemmas with their proofs, and the proofs of Theorems 1–3.

4.1. *Regularity conditions.*

C1 (*kernel function*) The kernel function $K_h(\cdot) = h^{-d}K(\cdot/h)$ where $K(\mathbf{a}) = \prod_{j=1}^{d} K(a_j)$ for $\mathbf{a} = (a_1, ..., a_d)^{\mathrm{T}}$ is symmetric on each individual entry and $K(x)$ is differentiable, decreasing when $x \geqslant 0$, and $\int_{-\infty}^{\infty} K(x)dx = 1$, $\int_{-\infty}^{\infty} x^j K(x)dx = 0$, for $1 \leqslant j < \nu$, $0 < \int_{-\infty}^{\infty} x^\nu K(x)dx < \infty$, and $\int_{-\infty}^{\infty} K^2(x)dx$, $\int_{-\infty}^{\infty} x^2 K^2(x)dx$, $\int_{-\infty}^{\infty} K'^2(x)dx$, $\int_{-\infty}^{\infty} x^2 K'^2(x)dx$, $\int_{-\infty}^{\infty} K''^2(x)dx$, $\int_{-\infty}^{\infty} x^2 K''^2(x)dx$ are all bounded. When there is no confusion, we use the same K for both univariate and multivariate kernel functions for simplicity.

C2 (*bandwidths*) The bandwidths h and b satisfy $h \to 0$, $nh^{2\nu} \to 0$, $b \to 0$ and $nh^{d+2}b \to \infty$, where $2\nu > d + 1$.

C3 (*density functions of covariates*) For all $\boldsymbol{\beta} \in \mathcal{B}$, the parameter space, the probability density function of $\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}$, $f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x})$, has a compact support and is bounded away from zero and ∞. Furthermore, the first and second derivatives of $f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x})$ exist and are Lipschitz continuous.

C4 (*smoothness*) For all $\boldsymbol{\beta} \in \mathcal{B}$ and $t > 0$, $E\{\mathbf{X}_j I(Z_j \geqslant t) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_j = \boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}\}$ is bounded and its first derivative is a Lipschitz continuous function of $\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}$. $E\{\mathbf{X}_j\mathbf{X}_j^{\mathrm{T}} I(Z_j \geqslant t) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_j = \boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}\}$ is a bounded and Lipschitz continuous function of $\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}$.

C5 (*survival function*) For all $\boldsymbol{\beta} \in \mathcal{B}$ and $t \in (0, \tau)$, $E\{S_c(t, \mathbf{X}) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}\}$ and $f(t, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{x})$ are bounded away from zero. Their first derivatives with respect to t and first and second derivatives with respect to $\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}$ exist and are Lipschitz continuous. In addition, $S_c(\tau, \mathbf{X})$ is bounded way from zero.

C6 (*boundedness*) The true parameter $\boldsymbol{\beta}_0$ is an interior point in $\mathcal{B}$, and $\mathcal{B}$ is bounded.

C7 (*uniqueness*) The equation of

$$E\left(\Delta\left[\frac{\mathbf{m}_{12}\{Z, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, I(W \leqslant Z), WI(W \leqslant Z)\}}{m_1\{Z, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, I(W \leqslant Z), WI(W \leqslant Z)\} + 1} - \frac{\mathbf{m}_2\{Z, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, I(W \leqslant Z), WI(W \leqslant Z)\}}{m\{Z, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, I(W \leqslant Z), WI(W \leqslant Z)\}}\right] \otimes \left[\mathbf{X}_l - \frac{E\left\{\mathbf{X}_l Y(Z) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, I(W \leqslant Z), WI(W \leqslant Z)\right\}}{E\left\{Y(Z) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, I(W \leqslant Z), WI(W \leqslant Z)\right\}}\right]\right) = \mathbf{0}$$

has a unique solution in $\mathcal{B}$.

Conditions C1 and C2 are standard assumptions in kernel regression analysis (Silverman, 1986; Ma and Zhu, 2013). Conditions C3 and C4 assume the boundedness of covariates and their expectations, while Condition C5 assumes the boundedness of the density functions of event and censoring times. These assumptions are generally applicable to real datasets. Additionally, the smoothness of several functions in Condition C4 is ensured by constraining their derivatives, a common practice in literature (Silverman, 1978). The boundedness assumption on the parameter space $\mathcal{B}$ in Condition C6 is reasonable in practical problems (Härdle et al., 1997). Furthermore, Condition C7 prevents the estimating equation from being degenerate.

4.2. *Two useful lemmas.* We present two lemmas and their proofs that are useful for proving Theorems 1–3. Specifically, Lemma 1 will be needed for proving Theorems 1–3, while Lemma 2 is also needed by Theorem 3.

4.2.1. *Lemma 1.*

Lemma 1. *Under the regularity conditions C1-C5 listed above,*

$$\widehat{E}\left\{Y(Z) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\right\} = E\{Y(Z) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\} + O_p\{(nh)^{-1/2} + h^2\}, \tag{S9}$$

$$\widehat{E}\left\{\mathbf{X}Y(Z) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\right\} = E\{\mathbf{X}Y(Z) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\} + O_p\{(nh)^{-1/2} + h^2\}, \tag{S10}$$

$$\widehat{\lambda}(z, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) = \lambda(z, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) + O_p\{(nhb)^{-1/2} + h^2 + b^2\} \tag{S11}$$

$$\widehat{\boldsymbol{\lambda}}_2(z, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) = \boldsymbol{\lambda}_2(z, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) + O_p\{(nbh^3)^{-1/2} + h^2 + b^2\} \tag{S12}$$

$$\widehat{\Lambda}(z, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) = \Lambda(z, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) + O_p\{(nh)^{-1/2} + h^2\} \tag{S13}$$

$$\widehat{\boldsymbol{\Lambda}}_2(z, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) = \boldsymbol{\Lambda}_2(z, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) + O_p\{(nh^3)^{-1/2} + h^2\} \tag{S14}$$

uniformly for all $z, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}$.

Proof: For notation convenience, we prove the results for $d = 1$. We prove

$$\widehat{E}\left\{\mathbf{X}Y(Z) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\right\} = E\{\mathbf{X}Y(Z) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\} + O_p\{(nh)^{-1/2} + h^2\}$$

and

$$\widehat{\mathbf{\Lambda}}_2(z,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) = \mathbf{\Lambda}_2(z,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) + O_p\{(nh^3)^{-1/2} + h^2\}.$$

and skip the remaining results because of the similar arguments.

First, for any $\mathbf{X}$ and $\boldsymbol{\beta}$ in a local neighborhood of $\boldsymbol{\beta}_0$,

$$\text{(S15)} \qquad \frac{1}{n}\sum_{j=1}^{n} K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_j - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) = f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) + O_p(n^{-1/2}h^{-1/2} + h^2),$$

To see this, the absolute bias of the left hand side of (S15) is

$$\begin{aligned}
&\left| E\left\{\frac{1}{n}\sum_{j=1}^{n} K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_j - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})\right\} - f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})\right| \\
&= \left|\int \frac{1}{h} K\left(\frac{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}_j - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}{h}\right) f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}_j) d\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}_j - f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})\right| \\
&= \left|\int K(u) f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X} + hu) du - f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})\right| \\
&= \left|\int K(u)\left\{f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) + f'_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})hu + O(h^2)u^2\right\} du - f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})\right| \\
&= O(h^2)
\end{aligned}$$

under Conditions C1-C3. The variance is

$$\begin{aligned}
&\operatorname{var}\left\{\frac{1}{n}\sum_{j=1}^{n} K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_j - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})\right\} \\
&= \frac{1}{n}\operatorname{var} K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_j - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) \\
&= \frac{1}{n}\left[E K_h^2(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_j - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) - \left\{E K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_j - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})\right\}^2\right] \\
&= \frac{1}{n}\left[\int \frac{1}{h^2} K^2\{(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}_j - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})/h\} f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}_j) d\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}_j - f^2_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) + O(h^2)\right] \\
&= \frac{1}{nh}\int K^2(u) f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X} + hu) du - \frac{1}{n} f^2_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) + O(h^2/n) \\
&\leqslant \frac{1}{nh} f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) \int K^2(u) du + O(h/n)\int u^2 K^2(u) du + \frac{1}{n}|f^2_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})| + O(h^2/n).
\end{aligned}$$

Therefore, applying the central limit theorem, we have that

$$\frac{1}{n}\sum_{j=1}^{n} K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_j - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) = f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) + O_p(n^{-1/2}h^{-1/2} + h^2)$$

for all $\boldsymbol{\beta}$ under Conditions C1-C3. Condition C3 also holds for any $\boldsymbol{\beta}$ in a local neighborhood of $\boldsymbol{\beta}_0$ due to the continuity. Similarly, We have

$$\text{(S16)} \qquad -\frac{1}{n}\sum_{j=1}^{n} K'_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_j - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) = f'_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) + O_p(n^{-1/2}h^{-3/2} + h^2).$$

To show (S10), the absolute bias is

$$\left| E\left\{ \frac{1}{n}\sum_{j=1}^{n} \mathbf{X}_j I(Z_j \geqslant z) K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_j - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) \right\} - f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) E\{\mathbf{X}_j(Z_j \geqslant z) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\} \right|$$

$$= \left| \int f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) E\left\{ \mathbf{X}_j I(Z_j \geqslant z) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X} \right\} K(u) du - f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) E\{\mathbf{X}_j(Z_j \geqslant z) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\} \right.$$

$$\left. + h \int \frac{\partial}{\partial(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})} f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) E\left\{ \mathbf{X}_j I(Z_j \geqslant z) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X} \right\} u K(u) du + O(h^2) \right|$$

$$= O(h^2)$$

under Conditions C1-C4. The variance is

$$\operatorname{var}\left\{ -\frac{1}{n}\sum_{j=1}^{n} \mathbf{X}_j I(Z_j \geqslant z) K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_j - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) \right\}$$

$$= \frac{1}{n^2}\sum_{j=1}^{n} \left(E\left\{ \mathbf{X}_j \mathbf{X}_j^{\mathrm{T}} I(Z_j \geqslant z) K_h^2(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_j - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) \right\} \right.$$

$$\left. - \left[E\left\{ \mathbf{X}_j I(Z_j \geqslant z) K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_j - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) \right\} \right]^2 \right)$$

$$\leqslant \frac{1}{nh} \sup_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}} \left| f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) E\{\mathbf{X}_j \mathbf{X}_j^{\mathrm{T}} I(Z_j \geqslant z) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\} \right| \int K^2(u) du + O(1/n)$$

under Conditions C1-C4. So

$$\frac{1}{n}\sum_{j=1}^{n} \mathbf{X}_j I(Z_j \geqslant z) K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_j - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) = f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) E\{\mathbf{X}_j I(Z_j \geqslant z) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\}$$

$$+ O_p(n^{-1/2}h^{-1/2} + h^2) \tag{S17}$$

under Conditions C1-C4.

To show (S14), let

$$\widehat{\boldsymbol{\Lambda}}_{21}(z, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) = -\sum_{i=1}^{n} \frac{I(Z_i \leqslant z)\Delta_i K_h'(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})}{\sum_{j=1}^{n} I(Z_j \geqslant Z_i) K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_j - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})}$$

$$\widehat{\boldsymbol{\Lambda}}_{22}(z, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) = \sum_{i=1}^{n} I(Z_i \leqslant z)\Delta_i K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) \frac{\sum_{j=1}^{n} I(Z_j \geqslant Z_i) K_h'(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_j - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})}{\{\sum_{j=1}^{n} I(Z_j \geqslant Z_i) K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_j - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})\}^2}.$$

Then $\widehat{\boldsymbol{\Lambda}}_2(z, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) = \widehat{\boldsymbol{\Lambda}}_{21}(z, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) + \widehat{\boldsymbol{\Lambda}}_{22}(z, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})$. To analyze $\widehat{\boldsymbol{\Lambda}}_{21}$,

$$E\widehat{\boldsymbol{\Lambda}}_{21}(z, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})$$

$$= E\left[\frac{-I(Z_i \leqslant z)\Delta_i K_h'(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})}{f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) S(Z_i, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) E\{S_c(Z_i, \mathbf{X}_j) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_j = \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, Z_i\}} \right]$$

$$+ E\left[\frac{1}{n}\sum_{i=1}^{n} \frac{-I(Z_i \leqslant z)\Delta_i K_h'(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})}{f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) S(Z_i, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) E\{S_c(Z_i, \mathbf{X}_j) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_j = \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, Z_i\}} O_p(A) \right].$$

The first term is

$$E\left[\frac{-I(Z_i \leqslant z)\Delta_i K_h'(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})}{f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) S(Z_i, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) E\{S_c(Z_i, \mathbf{X}_j) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_j = \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, Z_i\}} \right]$$

$$= \int I(z_i \leqslant z) \frac{\partial \left[f(z_i, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) E\{S_c(z_i, \mathbf{X}_i) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, z_i\} f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) \right] / \partial \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}{f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) S(z_i, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) E\{S_c(z_i, \mathbf{X}_j) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, z_i\}} dz_i$$

$$+ \iint I(z_i \leqslant z) O(h^2) u^3 K'(u) du dz_i \tag{S18}$$

under Condition C1 and C3-C5. Hence

$$\begin{aligned} &\Bigg| E\left[\frac{-I(Z_i \leqslant z)\Delta_i K_h'(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})}{f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) S(Z_i, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) E\{S_c(Z_i, \mathbf{X}_j) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_j = \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, Z_i\}} \right] \\ &- \int I(z_i \leqslant z) \frac{\partial \left[f(z_i, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) E\{S_c(z_i, \mathbf{X}_i) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, z_i\} f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) \right] / \partial \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}{f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) S(z_i, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) E\{S_c(z_i, \mathbf{X}_j) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, z_i\}} dz_i \Bigg| \\ &= O(h^2) \end{aligned}$$

under Condition C1 and C3-C5. Similarly, we conclude that

$$\frac{1}{n} \sum_{i=1}^{n} \frac{-I(Z_i \leqslant z)\Delta_i K_h'(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})}{f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) S(Z_i, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) E\{S_c(Z_i, \mathbf{X}_j) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_j = \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, Z_i\}} O_p(A) = O_p\{h^2 + (nh)^{-1/2}\}$$

under Conditions C1-C5 due to $A = O_p\{h^2 + (nh)^{-1/2}\}$. Therefore

$$\begin{aligned} E\widehat{\boldsymbol{\Lambda}}_{21}(z, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) = &\int I(z_i \leqslant z) \frac{\partial \left[f(z_i, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) E\{S_c(z_i, \mathbf{X}_i) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, z_i\} f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) \right] / \partial \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}{f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) S(z_i, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) E\{S_c(z_i, \mathbf{X}_j) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, z_i\}} dz_i \\ &+ O\{(nh)^{-1/2} + h^2\}. \end{aligned}$$

For $\widehat{\boldsymbol{\Lambda}}_{22}$, let $B = -1/n \sum_{j=1}^{n} I(Z_j \geqslant Z_i) K_h'(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_j - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) - \partial f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) E\{I(Z_j \geqslant Z_i) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\} / \partial \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}$, then

$$\begin{aligned} &\widehat{\boldsymbol{\Lambda}}_{22}(z, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) \\ &= -\frac{1}{n} \sum_{i=1}^{n} I(Z_i \leqslant z)\Delta_i K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) \\ &\quad \times \frac{\partial \left[f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) S(z_i, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) E\{S_c(z_i, \mathbf{X}_j) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, z_i\} \right] / \partial \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}{f^2_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) S^2(z_i, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) E^2\{S_c(z_i, \mathbf{X}_j) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, z_i\}} \\ &\quad \times \left\{ 1 + O_p(B) + O_p(A) \right\}. \end{aligned}$$

We have

$$\begin{aligned} &E\Bigg[-\frac{1}{n} \sum_{i=1}^{n} I(Z_i \leqslant z)\Delta_i K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) \\ &\quad \frac{\partial \left[f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) E\{I(Z_j \geqslant Z_i) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, Z_i\} \right] / \partial \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}{f^2_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) E^2\{I(Z_j \geqslant Z_i) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, Z_i\}} \Bigg] \\ &= -\iint I(z_i \leqslant z) K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) \\ &\quad \times \frac{\partial \left[f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) S(z_i, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) E\{S_c(z_i, \mathbf{X}_j) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, z_i\} \right] / \partial \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}{f^2_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) S^2(z_i, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) E^2\{S_c(z_i, \mathbf{X}_j) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, z_i\}} \end{aligned}$$

$$\times f(z_i,\boldsymbol\beta^{\rm T}\mathbf{X}_i)E\{S_c(z_i,\mathbf{X}_i)\mid\boldsymbol\beta^{\rm T}\mathbf{X}_i,z_i\}f_{\boldsymbol\beta^{\rm T}\mathbf{X}}(\boldsymbol\beta^{\rm T}\mathbf{X}_i)dz_id\boldsymbol\beta^{\rm T}\mathbf{X}_i$$
$$=-\int I(z_i\leqslant z)\frac{\partial\left[f_{\boldsymbol\beta^{\rm T}\mathbf{X}}(\boldsymbol\beta^{\rm T}\mathbf{X})S(z_i,\boldsymbol\beta^{\rm T}\mathbf{X})E\{S_c(z_i,\mathbf{X}_j)\mid\boldsymbol\beta^{\rm T}\mathbf{X},z_i\}\right]/\partial\boldsymbol\beta^{\rm T}\mathbf{X}}{f_{\boldsymbol\beta^{\rm T}\mathbf{X}}(\boldsymbol\beta^{\rm T}\mathbf{X})S^2(z_i,\boldsymbol\beta^{\rm T}\mathbf{X})E\{S_c(z_i,\mathbf{X}_j)\mid\boldsymbol\beta^{\rm T}\mathbf{X},z_i\}}f(z_i,\boldsymbol\beta^{\rm T}\mathbf{X})dz_i$$
$$-\iint I(z_i\leqslant z)\frac{\partial\left[f_{\boldsymbol\beta^{\rm T}\mathbf{X}}(\boldsymbol\beta^{\rm T}\mathbf{X})S(z_i,\boldsymbol\beta^{\rm T}\mathbf{X})E\{S_c(z_i,\mathbf{X}_j)\mid\boldsymbol\beta^{\rm T}\mathbf{X},z_i\}\right]/\partial\boldsymbol\beta^{\rm T}\mathbf{X}}{f^2_{\boldsymbol\beta^{\rm T}\mathbf{X}}(\boldsymbol\beta^{\rm T}\mathbf{X})S^2(z_i,\boldsymbol\beta^{\rm T}\mathbf{X})E^2\{S_c(z_i,\mathbf{X}_j)\mid\boldsymbol\beta^{\rm T}\mathbf{X},z_i\}}$$
$$\times O(h^2)u^2K(u)dz_idu,$$
(S19)

therefore

$$\Bigg|E\left[-\frac{1}{n}\sum_{i=1}^n I(Z_i\leqslant z)\Delta_iK_h(\boldsymbol\beta^{\rm T}\mathbf{X}_i-\boldsymbol\beta^{\rm T}\mathbf{X})\right.$$
$$\left.\frac{\partial\left[f_{\boldsymbol\beta^{\rm T}\mathbf{X}}(\boldsymbol\beta^{\rm T}\mathbf{z})E\{I(Z_j\geqslant Z_i)\mid\boldsymbol\beta^{\rm T}\mathbf{X},Z_i\}\right]/\partial\boldsymbol\beta^{\rm T}\mathbf{X}}{f^2_{\boldsymbol\beta^{\rm T}\mathbf{X}}(\boldsymbol\beta^{\rm T}\mathbf{X})E^2\{I(Z_j\geqslant Z_i)\mid\boldsymbol\beta^{\rm T}\mathbf{X},Z_i\}}\right]$$
$$+\int I(z_i\leqslant z)\frac{\partial\left[f_{\boldsymbol\beta^{\rm T}\mathbf{X}}(\boldsymbol\beta^{\rm T}\mathbf{X})S(z_i,\boldsymbol\beta^{\rm T}\mathbf{X})E\{S_c(z_i,\mathbf{X}_j)\mid\boldsymbol\beta^{\rm T}\mathbf{X},z_i\}\right]/\partial\boldsymbol\beta^{\rm T}\mathbf{X}}{f_{\boldsymbol\beta^{\rm T}\mathbf{X}}(\boldsymbol\beta^{\rm T}\mathbf{X})S^2(z_i,\boldsymbol\beta^{\rm T}\mathbf{X})E\{S_c(z_i,\mathbf{X}_j)\mid\boldsymbol\beta^{\rm T}\mathbf{X},z_i\}}f(z,\boldsymbol\beta^{\rm T}\mathbf{X})dz_i\Bigg|$$
$$=O(h^2)$$

under Conditions C1-C5. Recall $B=O_p(n^{-1/2}h^{-3/2}+h^2)$, then

$$\begin{aligned}&-\frac{1}{n}\sum_{i=1}^n I(Z_i\leqslant z)\Delta_iK_h(\boldsymbol\beta^{\rm T}\mathbf{X}_i-\boldsymbol\beta^{\rm T}\mathbf{X})\\&\quad\times\frac{\partial\left[f_{\boldsymbol\beta^{\rm T}\mathbf{X}}(\boldsymbol\beta^{\rm T}\mathbf{X})E\{I(Z_j\geqslant Z_i)\mid\boldsymbol\beta^{\rm T}\mathbf{X},Z_i\}\right]/\partial\boldsymbol\beta^{\rm T}\mathbf{X}}{f^2_{\boldsymbol\beta^{\rm T}\mathbf{X}}(\boldsymbol\beta^{\rm T}\mathbf{X})E^2\{I(Z_j\geqslant Z_i)\mid\boldsymbol\beta^{\rm T}\mathbf{X},Z_i\}}O_p(B)\\=&\ O_p(n^{-1/2}h^{-3/2}+h^2),\\&-\frac{1}{n}\sum_{i=1}^n I(Z_i\leqslant z)\Delta_iK_h(\boldsymbol\beta^{\rm T}\mathbf{X}_i-\boldsymbol\beta^{\rm T}\mathbf{X})\\&\quad\times\frac{\partial\left[f_{\boldsymbol\beta^{\rm T}\mathbf{X}}(\boldsymbol\beta^{\rm T}\mathbf{X})E\{I(Z_j\geqslant Z_i)\mid\boldsymbol\beta^{\rm T}\mathbf{X},Z_i\}\right]/\partial\boldsymbol\beta^{\rm T}\mathbf{X}}{f^2_{\boldsymbol\beta^{\rm T}\mathbf{X}}(\boldsymbol\beta^{\rm T}\mathbf{X})E^2\{I(Z_j\geqslant Z_i)\mid\boldsymbol\beta^{\rm T}\mathbf{X},Z_i\}}O_p(A)\\=&\ O_p(n^{-1/2}h^{-1/2}+h^2)\end{aligned}$$

under condition C1-C5. Therefore

$$E\widehat{\boldsymbol\Lambda}_{22}=-\int\frac{\partial\left[f_{\boldsymbol\beta^{\rm T}\mathbf{X}}(\boldsymbol\beta^{\rm T}\mathbf{X})S(z_i,\boldsymbol\beta^{\rm T}\mathbf{X})E\{S_c(z_i,\mathbf{X}_j)\mid\boldsymbol\beta^{\rm T}\mathbf{X},z_i\}\right]/\partial\boldsymbol\beta^{\rm T}\mathbf{X}}{f_{\boldsymbol\beta^{\rm T}\mathbf{X}}(\boldsymbol\beta^{\rm T}\mathbf{X})S^2(z_i,\boldsymbol\beta^{\rm T}\mathbf{X})E\{S_c(z_i,\mathbf{X}_j)\mid\boldsymbol\beta^{\rm T}\mathbf{X},z_i\}}$$
$$\times I(z_i\leqslant z)f(z_i,\boldsymbol\beta^{\rm T}\mathbf{X})dz_i+O(n^{-1/2}h^{-3/2}+h^2).$$

In addition

$$\int\Bigg(I(z_i\leqslant z)\frac{\partial\left[f(z_i,\boldsymbol\beta^{\rm T}\mathbf{X})E\{S_c(z_i,\mathbf{X}_j)\mid\boldsymbol\beta^{\rm T}\mathbf{X},z_i\}f_{\boldsymbol\beta^{\rm T}\mathbf{X}}(\boldsymbol\beta^{\rm T}\mathbf{X})\right]/\partial\boldsymbol\beta^{\rm T}\mathbf{X}}{f_{\boldsymbol\beta^{\rm T}\mathbf{X}}(\boldsymbol\beta^{\rm T}\mathbf{X})S(z_i,\boldsymbol\beta^{\rm T}\mathbf{X})E\{S_c(z_i,\mathbf{X}_j)\mid\boldsymbol\beta^{\rm T}\mathbf{X},z_i\}}$$

$$- I(z_i \leqslant z)\frac{\partial\left[f_{\boldsymbol{\beta}^{\mathrm T}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm T}\mathbf{X})S(z_i,\boldsymbol{\beta}^{\mathrm T}\mathbf{X})E\{S_c(z_i,\mathbf{X}_j)\mid\boldsymbol{\beta}^{\mathrm T}\mathbf{X},z_i\}\right]/\partial\boldsymbol{\beta}^{\mathrm T}\mathbf{X}}{f_{\boldsymbol{\beta}^{\mathrm T}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm T}\mathbf{X})S^2(z_i,\boldsymbol{\beta}^{\mathrm T}\mathbf{X})E\{S_c(z_i,\mathbf{X}_j)\mid\boldsymbol{\beta}^{\mathrm T}\mathbf{X},z_i\}}f(z_i,\boldsymbol{\beta}^{\mathrm T}\mathbf{X})\Bigg) dz_i$$

$$= \boldsymbol{\Lambda}_2(z,\boldsymbol{\beta}^{\mathrm T}\mathbf{X}).$$

Combining $E\widehat{\boldsymbol{\Lambda}}_{21}(z,\boldsymbol{\beta}^{\mathrm T}\mathbf{X})$ and $E\widehat{\boldsymbol{\Lambda}}_{22}(z,\boldsymbol{\beta}^{\mathrm T}\mathbf{X})$ gives

$$\left|E\widehat{\boldsymbol{\Lambda}}_2(z,\boldsymbol{\beta}^{\mathrm T}\mathbf{X}) - \boldsymbol{\Lambda}_2(z,\boldsymbol{\beta}^{\mathrm T}\mathbf{X})\right| = O(n^{-1/2}h^{-3/2}+h^2)$$

under Conditions C1-C5.

The variance of $\widehat{\boldsymbol{\Lambda}}_2(z,\boldsymbol{\beta}^{\mathrm T}\mathbf{X})$ is

$$\begin{aligned}\mathrm{var}\left\{\widehat{\boldsymbol{\Lambda}}_2(z,\boldsymbol{\beta}^{\mathrm T}\mathbf{X})\right\} &= \mathrm{var}\left\{\widehat{\boldsymbol{\Lambda}}_{21}(z,\boldsymbol{\beta}^{\mathrm T}\mathbf{X}) + \widehat{\boldsymbol{\Lambda}}_{22}(z,\boldsymbol{\beta}^{\mathrm T}\mathbf{X})\right\}\\ &\leqslant 2\mathrm{var}\{\widehat{\boldsymbol{\Lambda}}_{21}(z,\boldsymbol{\beta}^{\mathrm T}\mathbf{X})\} + 2\mathrm{var}\{\widehat{\boldsymbol{\Lambda}}_{22}(z,\boldsymbol{\beta}^{\mathrm T}\mathbf{X})\}.\end{aligned}$$

The first term

$$\begin{aligned}&\mathrm{var}\{\widehat{\boldsymbol{\Lambda}}_{21}(z,\boldsymbol{\beta}^{\mathrm T}\mathbf{X})\}\\ &\leqslant \frac{2}{n}\mathrm{var}\left[\frac{-I(Z_i\leqslant z)\Delta_i K_h'(\boldsymbol{\beta}^{\mathrm T}\mathbf{X}_i-\boldsymbol{\beta}^{\mathrm T}\mathbf{X})}{f_{\boldsymbol{\beta}^{\mathrm T}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm T}\mathbf{X})S(Z_i,\boldsymbol{\beta}^{\mathrm T}\mathbf{X})E\{S_c(Z_i,\mathbf{X}_j)\mid\boldsymbol{\beta}^{\mathrm T}\mathbf{X}_i=\boldsymbol{\beta}^{\mathrm T}\mathbf{X},Z_i\}}\right]\\ &+2\mathrm{var}\left[\frac{1}{n}\sum_{i=1}^n\frac{-I(Z_i\leqslant z)\Delta_i K_h'(\boldsymbol{\beta}^{\mathrm T}\mathbf{X}_i-\boldsymbol{\beta}^{\mathrm T}\mathbf{X})}{f_{\boldsymbol{\beta}^{\mathrm T}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm T}\mathbf{X})S(Z_i,\boldsymbol{\beta}^{\mathrm T}\mathbf{X})E\{S_c(Z_i,\mathbf{X}_j)\mid\boldsymbol{\beta}^{\mathrm T}\mathbf{X}_i=\boldsymbol{\beta}^{\mathrm T}\mathbf{X},Z_i\}}O_p(A)\right]\end{aligned}$$

The first part is

$$\begin{aligned}&\frac{2}{n}\mathrm{var}\left[\frac{-I(Z_i\leqslant z)\Delta_i K_h'(\boldsymbol{\beta}^{\mathrm T}\mathbf{X}_i-\boldsymbol{\beta}^{\mathrm T}\mathbf{X})}{f_{\boldsymbol{\beta}^{\mathrm T}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm T}\mathbf{X})S(Z_i,\boldsymbol{\beta}^{\mathrm T}\mathbf{X})E\{S_c(Z_i,\mathbf{X}_j)\mid\boldsymbol{\beta}^{\mathrm T}\mathbf{X}_i=\boldsymbol{\beta}^{\mathrm T}\mathbf{X},Z_i\}}\right]\\ &=\frac{2}{n}E\left[\frac{-I(Z_i\leqslant z)\Delta_i K_h'(\boldsymbol{\beta}^{\mathrm T}\mathbf{X}_i-\boldsymbol{\beta}^{\mathrm T}\mathbf{X})}{f_{\boldsymbol{\beta}^{\mathrm T}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm T}\mathbf{X})S(Z_i,\boldsymbol{\beta}^{\mathrm T}\mathbf{X})E\{S_c(Z_i,\mathbf{X}_j)\mid\boldsymbol{\beta}^{\mathrm T}\mathbf{X}_i=\boldsymbol{\beta}^{\mathrm T}\mathbf{X},Z_i\}}\right]^2\\ &-\frac{2}{n}\left(E\left[\frac{-I(Z_i\leqslant z)\Delta_i K_h'(\boldsymbol{\beta}^{\mathrm T}\mathbf{X}_i-\boldsymbol{\beta}^{\mathrm T}\mathbf{X})}{f_{\boldsymbol{\beta}^{\mathrm T}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm T}\mathbf{X})S(Z_i,\boldsymbol{\beta}^{\mathrm T}\mathbf{X})E\{S_c(Z_i,\mathbf{X}_j)\mid\boldsymbol{\beta}^{\mathrm T}\mathbf{X}_i=\boldsymbol{\beta}^{\mathrm T}\mathbf{X},Z_i\}}\right]\right)^2\\ &\leqslant\frac{2}{nh^3}\int\frac{I(z_i\leqslant z)f(z_i,\boldsymbol{\beta}^{\mathrm T}\mathbf{X})}{f_{\boldsymbol{\beta}^{\mathrm T}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm T}\mathbf{X})S^2(z_i,\boldsymbol{\beta}^{\mathrm T}\mathbf{X})E\{S_c(z_i,\mathbf{X}_j)\mid\boldsymbol{\beta}^{\mathrm T}\mathbf{X},z_i\}}dz_i\left\{\int K'^2(u)du\right\}\\ &+\frac{1}{nh^3}\int O(h^2)u^2K'^2(u)du+O(1/n)\end{aligned}$$

$$\text{(S20)}\quad = O\{1/(nh^3)\}$$

under Conditions C1-C5. The second part is

$$\begin{aligned}&2\mathrm{var}\left[\frac{1}{n}\sum_{i=1}^n\frac{-I(Z_i\leqslant z)\Delta_i K_h'(\boldsymbol{\beta}^{\mathrm T}\mathbf{X}_i-\boldsymbol{\beta}^{\mathrm T}\mathbf{X})}{f_{\boldsymbol{\beta}^{\mathrm T}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm T}\mathbf{X})S(Z_i,\boldsymbol{\beta}^{\mathrm T}\mathbf{X})E\{S_c(Z_i,\mathbf{X}_j)\mid\boldsymbol{\beta}^{\mathrm T}\mathbf{X}_i=\boldsymbol{\beta}^{\mathrm T}\mathbf{X},Z_i\}}O_p(A)\right]\\ &\leqslant 2E\left[\frac{1}{n}\sum_{i=1}^n\frac{-I(Z_i\leqslant z)\Delta_i K_h'(\boldsymbol{\beta}^{\mathrm T}\mathbf{X}_i-\boldsymbol{\beta}^{\mathrm T}\mathbf{X})}{f_{\boldsymbol{\beta}^{\mathrm T}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm T}\mathbf{X})S(Z_i,\boldsymbol{\beta}^{\mathrm T}\mathbf{X})E\{S_c(Z_i,\mathbf{X}_j)\mid\boldsymbol{\beta}^{\mathrm T}\mathbf{X}_i=\boldsymbol{\beta}^{\mathrm T}\mathbf{X},Z_i\}}O_p(A)\right]^2\end{aligned}$$

$$= \left\{ \left(E\left[\frac{-I(Z_i \leqslant z)\Delta_i K_h'(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})}{f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})S(Z_i, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})E\{S_c(Z_i, \mathbf{X}_j) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i = \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, Z_i\}} \right] \right)^2 \right.$$

$$\left. + \frac{1}{n}\mathrm{var}\left[\frac{-I(Z_i \leqslant z)\Delta_i K_h'(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})}{f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})S(Z_i, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})E\{S_c(Z_i, \mathbf{X}_j) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i = \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, Z_i\}} \right] \right\} O\{1/(nh) + h^4\}$$

$$= O\{1/(nh) + h^4\}$$

under Conditions C1-C5, where the second to last equation holds because of (S18) and (S20). Therefore $\mathrm{var}\{\widehat{\boldsymbol{\Lambda}}_{21}(z, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})\} = O\{1/(nh^3)\}$ under Conditions C1-C5.

For $\widehat{\boldsymbol{\Lambda}}_{22}(z, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})$,

$$\mathrm{var}\{\widehat{\boldsymbol{\Lambda}}_{22}(z, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})\}$$

$$\leqslant \frac{2}{n}\mathrm{var}\left[I(Z_i \leqslant z)\Delta_i K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) \frac{\partial\left[f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})E\{I(Z_j \geqslant Z_i) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\}\right]/\partial\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}{f^2_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})E^2\{I(Z_j \geqslant Z_i) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\}} \right]$$

$$+ 4\mathrm{var}\left[\frac{1}{n}\sum_{i=1}^n I(Z_i \leqslant z)\Delta_i K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) \right.$$

$$\left. \times \frac{\partial\left[f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})E\{I(Z_j \geqslant Z_i) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\}\right]/\partial\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}{f^2_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})E^2\{I(Z_j \geqslant Z_i) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\}} O_p(B) \right]$$

$$+ 4\mathrm{var}\left[\frac{1}{n}\sum_{i=1}^n I(Z_i \leqslant z)\Delta_i K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) \right.$$

$$\left. \times \frac{\partial\left[f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})E\{I(Z_j \geqslant Z_i) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\}\right]/\partial\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}{f^2_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})E^2\{I(Z_j \geqslant Z_i) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\}} O_p(A) \right].$$

The first part is

$$\frac{2}{n}\mathrm{var}\left[I(Z_i \leqslant z)\Delta_i K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) \frac{\partial\left[f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})E\{I(Z_j \geqslant Z_i) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, Z_i\}\right]/\partial\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}{f^2_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})E^2\{I(Z_j \geqslant Z_i) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, Z_i\}} \right]$$

$$\leqslant \frac{2}{nh}\int I(z_i \leqslant z)f(z_i, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) \frac{\left(\partial\left[f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})S(z_i, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})E\{S_c(z_i, \mathbf{X}_j) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\}\right]/\partial\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\right)^2}{f^3_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})S^4(z_i, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})E^3\{S_c(z_i, \mathbf{X}_j) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, z_i\}} dz_i$$

$$\times \left\{ \int K^2(u)du \right\}$$

$$+ \frac{1}{nh}\sup_{z, z_i, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}} \left| \frac{\left(\partial\left[f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})S(z_i, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})E\{S_c(z_i, \mathbf{X}_i) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\}\right]/\partial\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\right)^2}{f^4_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})S^4(z_i, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})E^4\{S_c(z_i, \mathbf{X}_i) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, z_i\}} \right|$$

$$\times \left\{ \int O(h^2)u^2K^2(u)du \right\} + O(1/n)$$

$$= O\{1/(nh)\}$$

(S21)

under Conditions C1-C5. The second part is

$$
\begin{aligned}
&4\mathrm{var}\Bigg[\frac{1}{n}\sum_{i=1}^{n} I(Z_i \leqslant z)\Delta_i K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) \\
&\quad \frac{\partial\left[f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})E\{I(Z_j \geqslant Z_i) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\}\right]/\partial\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}{f^2_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})E^2\{I(Z_j \geqslant Z_i) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, Z_i\}} O_p(B)\Bigg] \\
\leqslant\ & 4E\left(\left[\frac{1}{n}\sum_{i=1}^{n} I(Z_i \leqslant z)\Delta_i K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})\right.\right. \\
&\quad \left.\left.\times\ \frac{\partial\left[f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})E\{I(Z_j \geqslant Z_i) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\}\right]/\partial\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}{f^2_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})E^2\{I(Z_j \geqslant Z_i) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, Z_i\}}\right]^2\right) \\
&\quad \times O\{1/(nh^3) + h^4\} \\
=\ & O\{1/(nh^3) + h^4\}
\end{aligned}
$$

under Conditions C1-C5, where the second to last equation holds because of (S19) and (S21). The last part is

$$
\begin{aligned}
&4\mathrm{var}\Bigg[\frac{1}{n}\sum_{i=1}^{n} I(Z_i \leqslant z)\Delta_i K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) \\
&\quad \times\ \frac{\partial\left[f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})E\{I(Z_j \geqslant Z_i) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\}\right]/\partial\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}{f^2_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})E^2\{I(Z_j \geqslant Z_i) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, Z_i\}} O_p(A)\Bigg] \\
\leqslant\ & 4E\left(\left[\frac{1}{n}\sum_{i=1}^{n} I(Z_i \leqslant z)\Delta_i K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})\right.\right. \\
&\quad \left.\left.\times\ \frac{\partial\left[f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})E\{I(Z_j \geqslant Z_i) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\}\right]/\partial\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}{f^2_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})E^2\{I(Z_j \geqslant Z_i) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, Z_i\}}\right]^2\right) \\
&\quad \times O\{(nh)^{-1} + h^4\} \\
=\ & O\{(nh)^{-1} + h^4\}
\end{aligned}
$$

under Conditions C1-C5 where the second to last equation is due to (S19) and (S21). Therefore $\mathrm{var}\{\widehat{\boldsymbol{\Lambda}}_{22}(z, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})\} = O\{1/(nh^3)\}$ under Conditions C1-C5.

Combining all these results, we have that $\mathrm{var}\{\widehat{\boldsymbol{\Lambda}}_2(z, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})\} = O\{1/(nh^3)\}$. Hence, the estimator $\widehat{\boldsymbol{\Lambda}}_2(z, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})$ satisfies

$$\widehat{\boldsymbol{\Lambda}}_2(z, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) = \boldsymbol{\Lambda}_2(z, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) + O_p\{(nh^3)^{-1/2} + h^2\}$$

under Conditions C1-C5.

We provide a detailed proof for the uniform result of (S10) only. Since the domain of $\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}$ is compact, we divide it into rectangular regions. Within each region, the distance between a point $\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}$ and the nearest grid point is less than n^{-2}. We need only $N \leqslant Cn^2$ grid points, where C is a constant. Let the grid points be $\boldsymbol{\kappa}_1, \ldots, \boldsymbol{\kappa}_N$. Let $\widehat{\boldsymbol{\rho}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) = \widehat{E}\{\mathbf{X}Y(Z) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\}$ and $\boldsymbol{\rho}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) = E\{\mathbf{X}Y(Z) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\}$. Then for any $(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})$, there exists a $\boldsymbol{\kappa}_i$, $1 \leqslant i \leqslant N$, such that

$$|\widehat{\boldsymbol{\rho}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) - \boldsymbol{\rho}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})| \leqslant |\widehat{\boldsymbol{\rho}}(\boldsymbol{\kappa}_i) - \boldsymbol{\rho}(\boldsymbol{\kappa}_i)| + |\widehat{\boldsymbol{\rho}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) - \widehat{\boldsymbol{\rho}}(\boldsymbol{\kappa}_i)| + |\boldsymbol{\rho}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) - \boldsymbol{\rho}(\boldsymbol{\kappa}_i)|$$

$$\leqslant |\widehat{\boldsymbol{\rho}}(\boldsymbol{\kappa}_i) - \boldsymbol{\rho}(\boldsymbol{\kappa}_i)| + D_1 n^{-2},$$

for an absolute constant D_1 under Conditions C1 and C5. Thus, for any $D \geqslant D_1$,

$$\begin{aligned}
&\text{pr}(\sup_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}} |\widehat{\boldsymbol{\rho}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) - \boldsymbol{\rho}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})| > 2D[h^2 + \{\log n (nh)^{-1}\}^{1/2}]) \\
&\leqslant \text{pr}(\sup_{\boldsymbol{\kappa}_i} |\widehat{\boldsymbol{\rho}}(\boldsymbol{\kappa}_i) - \boldsymbol{\rho}(\boldsymbol{\kappa}_i)| > 2D[h^2 + \{\log n (nh)^{-1}\}^{1/2}] - D_1 n^{-2}) \\
&\leqslant \text{pr}(\sup_{\boldsymbol{\kappa}_i} |\widehat{\boldsymbol{\rho}}(\boldsymbol{\kappa}_i) - \boldsymbol{\rho}(\boldsymbol{\kappa}_i)| > D[h^2 + \{\log n (nh)^{-1}\}^{1/2}])
\end{aligned}$$

under Condition C2. Using Bernstein's inequality on $\widehat{\boldsymbol{\rho}}(\boldsymbol{\kappa}_i)$, under Conditions C1-C5, we have that

$$\begin{aligned}
\text{pr}[|\widehat{\boldsymbol{\rho}}(\boldsymbol{\kappa}_i) - \boldsymbol{\rho}(\boldsymbol{\kappa}_i)| \geqslant A\{\log n/(nh)\}^{1/2}] &\leqslant 2\exp\left\{\frac{-nA^2 \log n/(nh)}{2D_2 h^{-1} + 2/3AD_3(\log n)^{1/2}(nh)^{-1/2}}\right\} \\
&= 2\exp\left\{\frac{-A^2\log n}{2D_2 + 2/3AD_3(\log nh/n)^{1/2}}\right\} \\
&\leqslant 2\exp\left(\frac{-A^2\log n}{2D_2 + AD_3}\right)
\end{aligned}$$

for all $A > D_3 + \sqrt{D_3^2 + 4D_2}$, where D_2 and D_3 are constants satisfying

$$\text{var}\{\widehat{\boldsymbol{\rho}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) - \boldsymbol{\rho}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})\} \leqslant \frac{D_2}{nh},$$

$$\left|\frac{K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})\mathbf{X}_i I(Z_i \geqslant Z)}{1/n\sum_{j=1}^n K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_j - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})} - \boldsymbol{\rho}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})\right| \leqslant D_3 \text{ with probability 1.}$$

This leads to

$$\begin{aligned}
\text{pr}[\sup_{\boldsymbol{\kappa}_i} |\widehat{\boldsymbol{\rho}}(\boldsymbol{\kappa}_i) - E\widehat{\boldsymbol{\rho}}(\boldsymbol{\kappa}_i)| \geqslant A\{\log n/(nh)\}^{1/2}] &\leqslant 2Cn^2 \exp\left(\frac{-A^2\log n}{2D_2 + AD_3}\right) \\
&= 2C\exp\left[\{2 - A^2/(2D_2 + AD_3)\}\log n\right] \to 0
\end{aligned}$$

because $A > D_3 + \sqrt{D_3^2 + 4D_2}$. Combining the above results, for $A_1 = \max(A, D)$,

$$\begin{aligned}
&\text{pr}(\sup_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}} |\widehat{\boldsymbol{\rho}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, \boldsymbol{\beta}) - \boldsymbol{\rho}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})| > 2A_1[h^2 + \{\log n (nh)^{-1}\}^{1/2}]) \\
&\leqslant \text{pr}(\sup_{\boldsymbol{\kappa}_i} |\widehat{\boldsymbol{\rho}}(\boldsymbol{\kappa}_i) - \boldsymbol{\rho}(\boldsymbol{\kappa}_i)| > A_1[h^2 + \{\log n (nh)^{-1}\}^{1/2}]) \\
&\leqslant \text{pr}\{\sup_{\boldsymbol{\kappa}_i} |\widehat{\boldsymbol{\rho}}(\boldsymbol{\kappa}_i) - \boldsymbol{\rho}(\boldsymbol{\kappa}_i)| > A_1 h^2\} + \text{pr}([\sup_{\boldsymbol{\kappa}_i} |\widehat{\boldsymbol{\rho}}(\boldsymbol{\kappa}_i) - \boldsymbol{\rho}(\boldsymbol{\kappa}_i)| \geqslant A_1\{\log n (nh)^{-1}\}^{1/2}]) \\
&\to 0.
\end{aligned}$$

The uniform convergence results for (S11)-(S14) exhibit slight variations due to the additional component Z in these functions. Nonetheless, under Condition C5, the support of $(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i, Z_i)$ or $(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i, Z_j)$ is also bounded. Consequently, we can divide the region using $N \leqslant Cn^{2+2}$ grid points, ensuring that the distance of a point to the nearest grid point is less than n^{-2}. The subsequent analysis follows a similar approach as outlined above, leading to the establishment of uniform convergence. □

4.2.2. *Lemma 2.*

Lemma 2. *The estimator $\widehat{\Lambda}(t, \widehat{\boldsymbol\beta}^{\mathrm T}\mathbf{X})$ has the expansion*

$$\begin{aligned}&\sqrt{nh}\left\{\widehat{\Lambda}(t,\widehat{\boldsymbol\beta}^{\mathrm T}\mathbf{X})-\Lambda(t,\boldsymbol\beta^{\mathrm T}\mathbf{X})\right\}\\ &=\sqrt{\frac{h}{n}}\sum_{i=1}^{n}\int_0^t\frac{I\left\{\sum_{j=1}^n Y_j(s)K_h(\boldsymbol\beta^{\mathrm T}\mathbf{X}_j-\boldsymbol\beta^{\mathrm T}\mathbf{X})>0\right\}}{f_{\boldsymbol\beta^{\mathrm T}\mathbf{X}}(\boldsymbol\beta^{\mathrm T}\mathbf{X})E\{I(z\geqslant s)\mid\boldsymbol\beta^{\mathrm T}\mathbf{X}\}}K_h(\boldsymbol\beta^{\mathrm T}\mathbf{X}_i-\boldsymbol\beta^{\mathrm T}\mathbf{X})\\ &\times dM_i\{s,\boldsymbol\beta^{\mathrm T}\mathbf{X},I(W\leqslant s),WI(W\leqslant s)\}+o_p(1),\end{aligned}$$

and satisfies

$$\sqrt{nh}\left\{\widehat{\Lambda}(t,\widehat{\boldsymbol\beta}^{\mathrm T}\mathbf{X})-\Lambda(t,\boldsymbol\beta^{\mathrm T}\mathbf{X})\right\}\to N\{0,\sigma^2(t,\boldsymbol\beta^{\mathrm T}\mathbf{X})\}$$

in distribution when $n\to\infty$ for all $t,\boldsymbol\beta^{\mathrm T}\mathbf{X}$ under Conditions C1-C5, where

$$\sigma^2(t,\boldsymbol\beta^{\mathrm T}\mathbf{X})=\int K^2(u)du\int_0^t\frac{\lambda(s,\boldsymbol\beta^{\mathrm T}\mathbf{X})}{f_{\boldsymbol\beta^{\mathrm T}\mathbf{X}}(\boldsymbol\beta^{\mathrm T}\mathbf{X})E\{I(Z\geqslant s)\mid\boldsymbol\beta^{\mathrm T}\mathbf{X}\}}ds.$$

Proof: We only need to prove $\sqrt{nh}\left\{\widehat{\Lambda}(t,\boldsymbol\beta^{\mathrm T}\mathbf{X})-\Lambda(t,\boldsymbol\beta^{\mathrm T}\mathbf{X})\right\}\to N\{0,\sigma^2(t,\boldsymbol\beta^{\mathrm T}\mathbf{X})\}$ in distribution because $\sqrt{nh}\left\{\widehat{\Lambda}(t,\widehat{\boldsymbol\beta}^{\mathrm T}\mathbf{X})-\widehat{\Lambda}(t,\boldsymbol\beta^{\mathrm T}\mathbf{X})\right\}=O_p(\sqrt{h})=o_p(1)$ due to Theorems 1 and 2. For notational convenience, let $d=1$ and $\nu=2$. For any t and $\boldsymbol\beta^{\mathrm T}\mathbf{X}$, define

$$\begin{aligned}\phi_n(s,\boldsymbol\beta^{\mathrm T}\mathbf{X})&\equiv\sum_{j=1}^n Y_j(s)K_h(\boldsymbol\beta^{\mathrm T}\mathbf{X}_j-\boldsymbol\beta^{\mathrm T}\mathbf{X}),\\ Q_n(t,\boldsymbol\beta^{\mathrm T}\mathbf{X})&\equiv\widehat{\boldsymbol\Lambda}(t,\boldsymbol\beta^{\mathrm T}\mathbf{X})-\int_0^t I\left\{\phi_n(s,\boldsymbol\beta^{\mathrm T}\mathbf{X})>0\right\}\lambda(s,\boldsymbol\beta^{\mathrm T}\mathbf{X})ds,\\ D_n(t,\boldsymbol\beta^{\mathrm T}\mathbf{X})&\equiv\int_0^t\lambda(s,\boldsymbol\beta^{\mathrm T}\mathbf{X})\left[1-I\left\{\phi_n(s,\boldsymbol\beta^{\mathrm T}\mathbf{X})>0\right\}\right]ds.\end{aligned}$$

Then

$$\sqrt{nh}\left\{\widehat{\Lambda}(t,\boldsymbol\beta^{\mathrm T}\mathbf{X})-\Lambda(t,\boldsymbol\beta^{\mathrm T}\mathbf{X})\right\}=\sqrt{nh}Q_n(t,\boldsymbol\beta^{\mathrm T}\mathbf{X})-\sqrt{nh}D_n(t,\boldsymbol\beta^{\mathrm T}\mathbf{X}).$$

We first show that $\sqrt{nh}D_n(t,\boldsymbol\beta^{\mathrm T}\mathbf{X})\to 0$ in probability uniformly. It suffices to show that

$$\sqrt{nh}\left[1-I\left\{\phi_n(s,\boldsymbol\beta^{\mathrm T}\mathbf{X})>0\right\}\right]\xrightarrow{p}0,$$

which is equivalent to show that for any $\epsilon>0$,

$$\Pr\left(\sqrt{nh}\left[1-I\left\{\phi_n(s,\boldsymbol\beta^{\mathrm T}\mathbf{X})>0\right\}\right]>\epsilon\right)\to 0.$$

Now for $\epsilon\geqslant\sqrt{nh}$, the above holds trivially. For $\epsilon<\sqrt{nh}$, this is equivalent to show

$$\Pr\left\{n^{-1}\phi_n(s,\boldsymbol\beta^{\mathrm T}\mathbf{X})\leqslant 0\right\}\to 0. \tag{S22}$$

Because $n^{-1}\phi_n(s,\boldsymbol\beta^{\mathrm T}\mathbf{X})=f_{\boldsymbol\beta^{\mathrm T}\mathbf{X}}(\boldsymbol\beta^{\mathrm T}\mathbf{X})E\{I(Z_j\geqslant s)\mid\boldsymbol\beta^{\mathrm T}\mathbf{X}\}+O_p\{h^2+(nh)^{-1/2}\}$, (S22) is equivalent to

$$\Pr\left[f_{\boldsymbol\beta^{\mathrm T}\mathbf{X}}(\boldsymbol\beta^{\mathrm T}\mathbf{X})E\{I(Z_j\geqslant s)\mid\boldsymbol\beta^{\mathrm T}\mathbf{X}\}+O_p\{h^2+(nh)^{-1/2}\}\leqslant 0\right]\to 0,$$

which automatically holds under Conditions C3 and C5. Hence $\sqrt{nh}D_n(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) \to 0$ in probability uniformly. Second we inspect the asymptotic property of $\sqrt{nh}Q_n(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})$. Recall $M_i(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})$ is the martingale corresponding to the counting process $N_i(s)$ and satisfies $dM_i(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) = dN_i(s) - Y_i(s)\lambda(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})ds$.

$$
\begin{aligned}
&\sqrt{nh}Q_n(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) \\
=& \int_0^t \sqrt{nh}\frac{1}{\sum_{j=1}^n Y_j(s)K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_j - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})}\sum_{i=1}^n K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})dN_i(s) \\
&- \int_0^t \sqrt{nh} I\left\{\phi_n(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) > 0\right\}\lambda(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})ds \\
=& \int_0^t \sqrt{nh}\frac{I\{\phi_n(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) \leqslant 0\}}{\sum_{j=1}^n Y_j(s)K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_j - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})}\sum_{i=1}^n K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})dN_i(s) \\
&+ \int_0^t \sqrt{nh}\frac{I\{\phi_n(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) > 0\}}{\sum_{j=1}^n Y_j(s)K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_j - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})}\sum_{i=1}^n K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})dN_i(s) \\
&- \int_0^t \sqrt{nh}\frac{I\{\phi_n(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) > 0\}}{\sum_{j=1}^n Y_j(s)K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_j - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})}\sum_{i=1}^n K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})Y_i(s)\lambda(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})ds \\
=& \int_0^t \sqrt{nh}\frac{I\{\phi_n(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) > 0\}}{\sum_{j=1}^n Y_j(s)K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_j - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})} \\
&\times \sum_{i=1}^n K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})dM_i\{s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\} \\
&\text{(S23)} \\
&\text{(S24)}+ \int_0^t \sqrt{nh}\frac{I\{\phi_n(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) \leqslant 0\}}{\sum_{j=1}^n Y_j(s)K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_j - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})}\sum_{i=1}^n K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})dN_i(s).
\end{aligned}
$$

In (S23), it leads to the same martingale $dM_i\{s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\} = dN_i(t) - Y_i(s)[\lambda_T(t-W,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})I(W \leqslant t) + \lambda_N(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})\{1 - I(W \leqslant t)\}]ds$ in either group of transplant or nontransplant because one of $\lambda_T(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})$ and $\lambda_N(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})$ is zero.

We decompose (S23) as

$$
\begin{aligned}
&\int_0^t \sqrt{nh}\frac{I\{\phi_n(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) > 0\}}{\sum_{j=1}^n Y_j(s)K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_j - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})} \\
&\quad \sum_{i=1}^n K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})dM_i\{s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\} \\
=& \sqrt{\frac{h}{n}}\sum_{i=1}^n \int_0^t \frac{I\{\phi_n(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) > 0\}}{1/n\sum_{j=1}^n Y_j(s)K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_j - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})} \\
&\quad K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})dM_i\{s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\} \\
=& \sqrt{\frac{h}{n}}\sum_{i=1}^n \int_0^t I\left\{\phi_n(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) > 0\right\}\Bigg[\frac{1}{f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})E\{I(Z_j \geqslant s) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\}}
\end{aligned}
$$

$$-\frac{1/n\sum_{j=1}^{n}Y_j(s)K_h(\boldsymbol\beta^{\mathrm T}\mathbf X_j-\boldsymbol\beta^{\mathrm T}\mathbf X)}{f^2_{\boldsymbol\beta^{\mathrm T}\mathbf X}(\boldsymbol\beta^{\mathrm T}\mathbf X)E^2\{I(Z_j\geqslant s)\mid\boldsymbol\beta^{\mathrm T}\mathbf X\}}+\frac{1}{f_{\boldsymbol\beta^{\mathrm T}\mathbf X}(\boldsymbol\beta^{\mathrm T}\mathbf X)E\{I(Z_j\geqslant s)\mid\boldsymbol\beta^{\mathrm T}\mathbf X\}}$$

$$+O_p\{h^4+(nh)^{-1}\}\Big]\times K_h(\boldsymbol\beta^{\mathrm T}\mathbf X_i-\boldsymbol\beta^{\mathrm T}\mathbf X)dM_i\{s,\boldsymbol\beta^{\mathrm T}\mathbf X,I(W\leqslant s),WI(W\leqslant s)\}$$

$$= Q_{n1}-Q_{n2}+o_p(1), \tag{S25}$$

where

$$Q_{n1}=\sqrt{\frac{h}{n}}\sum_{i=1}^{n}\int_0^t\frac{I\left\{\phi_n(s,\boldsymbol\beta^{\mathrm T}\mathbf X)>0\right\}K_h(\boldsymbol\beta^{\mathrm T}\mathbf X_i-\boldsymbol\beta^{\mathrm T}\mathbf X)}{f_{\boldsymbol\beta^{\mathrm T}\mathbf X}(\boldsymbol\beta^{\mathrm T}\mathbf X)E\{I(Z\geqslant s)\mid\boldsymbol\beta^{\mathrm T}\mathbf X\}}$$
$$\times dM_i\{s,\boldsymbol\beta^{\mathrm T}\mathbf X,I(W\leqslant s),WI(W\leqslant s)\},$$

$$Q_{n2}=\sqrt{\frac{h}{n}}\sum_{i=1}^{n}\int_0^t\frac{I\left\{\phi_n(s,\boldsymbol\beta^{\mathrm T}\mathbf X)>0\right\}}{f^2_{\boldsymbol\beta^{\mathrm T}\mathbf X}(\boldsymbol\beta^{\mathrm T}\mathbf X)E^2\{I(Z_j\geqslant s)\mid\boldsymbol\beta^{\mathrm T}\mathbf X\}}$$
$$\times\left[\frac{1}{n}\sum_{j=1}^{n}Y_j(s)K_h(\boldsymbol\beta^{\mathrm T}\mathbf X_j-\boldsymbol\beta^{\mathrm T}\mathbf X)-f_{\boldsymbol\beta^{\mathrm T}\mathbf X}(\boldsymbol\beta^{\mathrm T}\mathbf X)E\{I(Z\geqslant s)\mid\boldsymbol\beta^{\mathrm T}\mathbf X\}\right]$$
$$\times K_h(\boldsymbol\beta^{\mathrm T}\mathbf X_i-\boldsymbol\beta^{\mathrm T}\mathbf X)dM_i\{s,\boldsymbol\beta^{\mathrm T}\mathbf X,I(W\leqslant s),WI(W\leqslant s)\}$$
$$=\sqrt{\frac{h}{n^3}}\sum_{i=1}^{n}\sum_{j=1}^{n}\int_0^t\frac{I\left\{\phi_n(s,\boldsymbol\beta^{\mathrm T}\mathbf X)>0\right\}}{f^2_{\boldsymbol\beta^{\mathrm T}\mathbf X}(\boldsymbol\beta^{\mathrm T}\mathbf X)E^2\{I(Z_j\geqslant s)\mid\boldsymbol\beta^{\mathrm T}\mathbf X\}}$$
$$\times\left[Y_j(s)K_h(\boldsymbol\beta^{\mathrm T}\mathbf X_j-\boldsymbol\beta^{\mathrm T}\mathbf X)-f_{\boldsymbol\beta^{\mathrm T}\mathbf X}(\boldsymbol\beta^{\mathrm T}\mathbf X)E\{I(Z\geqslant s)\mid\boldsymbol\beta^{\mathrm T}\mathbf X\}\right]$$
$$\times K_h(\boldsymbol\beta^{\mathrm T}\mathbf X_i-\boldsymbol\beta^{\mathrm T}\mathbf X)dM_i\{s,\boldsymbol\beta^{\mathrm T}\mathbf X,I(W\leqslant s),WI(W\leqslant s)\},$$

and the remaining term in (S25) is $o_p(1)$ because $\sqrt{n/h}O_p\{h^4+(nh)^{-1}\}=O_p\{n^{1/2}h^{7/2}+(nh^3)^{-1/2}\}=o_p(1)$ by Condition C2.

Using the U-statistic property, Q_{n2} has leading order terms $Q_{n21}+Q_{n22}-Q_{n23}$, where

$$Q_{n21}=\sqrt{\frac{h}{n}}E\Bigg(\sum_{i=1}^{n}\int_0^t\frac{I\left\{\phi_n(s,\boldsymbol\beta^{\mathrm T}\mathbf X)>0\right\}}{f^2_{\boldsymbol\beta^{\mathrm T}\mathbf X}(\boldsymbol\beta^{\mathrm T}\mathbf X)E^2\{I(Z_j\geqslant s)\mid\boldsymbol\beta^{\mathrm T}\mathbf X\}}$$
$$\times\left[Y_j(s)K_h(\boldsymbol\beta^{\mathrm T}\mathbf X_j-\boldsymbol\beta^{\mathrm T}\mathbf X)-f_{\boldsymbol\beta^{\mathrm T}\mathbf X}(\boldsymbol\beta^{\mathrm T}\mathbf X)E\{I(Z\geqslant s)\mid\boldsymbol\beta^{\mathrm T}\mathbf X\}\right]K_h(\boldsymbol\beta^{\mathrm T}\mathbf X_i-\boldsymbol\beta^{\mathrm T}\mathbf X)$$
$$\times dM_i\{s,\boldsymbol\beta^{\mathrm T}\mathbf X,I(W\leqslant s),WI(W\leqslant s)\}\mid\Delta_i,\boldsymbol\beta^{\mathrm T}\mathbf X_i,Z_i,I(W\leqslant s),WI(W\leqslant s)\Bigg),$$

$$Q_{n22}=\sqrt{\frac{h}{n}}E\Bigg(\sum_{j=1}^{n}\int_0^t\frac{I\left\{\phi_n(s,\boldsymbol\beta^{\mathrm T}\mathbf X)>0\right\}}{f^2_{\boldsymbol\beta^{\mathrm T}\mathbf X}(\boldsymbol\beta^{\mathrm T}\mathbf X)E^2\{I(Z_j\geqslant s)\mid\boldsymbol\beta^{\mathrm T}\mathbf X\}}$$
$$\times\left[Y_j(s)K_h(\boldsymbol\beta^{\mathrm T}\mathbf X_j-\boldsymbol\beta^{\mathrm T}\mathbf X)-f_{\boldsymbol\beta^{\mathrm T}\mathbf X}(\boldsymbol\beta^{\mathrm T}\mathbf X)E\{I(Z\geqslant s)\mid\boldsymbol\beta^{\mathrm T}\mathbf X\}\right]K_h(\boldsymbol\beta^{\mathrm T}\mathbf X_i-\boldsymbol\beta^{\mathrm T}\mathbf X)$$
$$\times dM_i\{s,\boldsymbol\beta^{\mathrm T}\mathbf X,I(W\leqslant s),WI(W\leqslant s)\}\mid\Delta_j,\boldsymbol\beta^{\mathrm T}\mathbf X_j,Z_j,I(W\leqslant s),WI(W\leqslant s)\Bigg),$$

$$Q_{n23}=\sqrt{nh}E\Bigg(\int_0^t\frac{I\left\{\phi_n(s,\boldsymbol\beta^{\mathrm T}\mathbf X)>0\right\}}{f^2_{\boldsymbol\beta^{\mathrm T}\mathbf X}(\boldsymbol\beta^{\mathrm T}\mathbf X)E^2\{I(Z_j\geqslant s)\mid\boldsymbol\beta^{\mathrm T}\mathbf X\}}$$
$$\times\left[Y_j(s)K_h(\boldsymbol\beta^{\mathrm T}\mathbf X_j-\boldsymbol\beta^{\mathrm T}\mathbf X)-f_{\boldsymbol\beta^{\mathrm T}\mathbf X}(\boldsymbol\beta^{\mathrm T}\mathbf X)E\{I(Z\geqslant s)\mid\boldsymbol\beta^{\mathrm T}\mathbf X\}\right]$$

$$\times E\left\{K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i-\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})dM_i\{s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},I(W\leqslant s),WI(W\leqslant s)\}\right\}\Bigg).$$

$I\left\{\phi_n(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})>0\right\} = I\left[f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})E\{I(Z_j\geqslant s)\mid\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\}+O_p\{h^2+(nh)^{-1}\}>0\right] = 1$ almost surely. Thus, almost surely,

$$\begin{aligned}
Q_{n21} &= \sqrt{\frac{h}{n}}\sum_{i=1}^{n}\Bigg(\int_0^t\frac{1}{f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})E\{I(Z_j\geqslant s)\mid\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\}}\\
&\quad\times E\left[Y_j(s)K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_j-\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})-f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})E\{I(Z\geqslant s)\mid\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\}\right]\\
&\quad\times K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i-\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})dM_i\{s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},I(W\leqslant s),WI(W\leqslant s)\}\Bigg)\\
&= \sqrt{\frac{h}{n}}\sum_{i=1}^{n}\int_0^t\frac{O(h^2)K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i-\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})}{f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})E\{I(Z_j\geqslant s)\mid\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\}}dM_i\{s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},I(W\leqslant s),WI(W\leqslant s)\}\\
&\to 0.
\end{aligned}$$

uniformly as $h\to 0$. Similarly, almost surely,

$$\begin{aligned}
Q_{n22} &= \sqrt{\frac{h}{n}}\sum_{j=1}^{n}\int_0^t\frac{1}{f^2_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})E^2\{I(Z_j\geqslant s)\mid\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\}}\\
&\quad\times\left[Y_j(s)K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_j-\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})-f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})E\{I(Z\geqslant s)\mid\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\}\right]\\
&\quad\times E[K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i-\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})dM_i\{s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},I(W\leqslant s),WI(W\leqslant s)\}]\\
&= 0.
\end{aligned}$$

Obviously, $Q_{n23}=E(Q_{n22})=0$, hence $Q_{n2}\to 0$ in probability uniformly as $n\to\infty$.

For (S24)

$$\int_0^t\sqrt{nh}\frac{I\{\phi_n(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})\leqslant 0\}}{\sum_{j=1}^n Y_j(s)K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_j-\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})}\sum_{i=1}^{n}K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i-\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})dN_i(s)\to 0$$

in probability uniformly. We have obtained

$$\begin{aligned}
&\sqrt{nh}Q_n(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})\\
&= \sqrt{\frac{h}{n}}\sum_{i=1}^{n}\int_0^t\frac{I\left\{\phi_n(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})>0\right\}}{f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})E\{I(Z\geqslant s)\mid\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\}}K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i-\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})\\
&\quad\times dM_i\{s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},I(W\leqslant s),WI(W\leqslant s)\}
\end{aligned}$$

(S26)

$$+o_p(1).$$

Applying the martingale central limit theorem to (S26), we have

$$\begin{aligned}
&\frac{h}{n}\sum_{i=1}^{n}\int_0^t\frac{\lambda(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})I\{\phi_n(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})>0\}}{[f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})E\{I(Z\geqslant s)\mid\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\}]^2}K_h^2(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i-\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})Y_i(s)ds\\
&= \int_0^t\frac{\lambda(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})I\{\phi_n(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})>0\}}{[f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})E\{I(Z\geqslant s)\mid\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\}]^2}\frac{1}{n}\sum_{i=1}^{n}hK_h^2(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i-\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})Y_i(s)ds
\end{aligned}$$

$$= \int_0^t \frac{\lambda(s,\boldsymbol{\beta}^{\mathrm T}\mathbf{X}) I\{\phi_n(s,\boldsymbol{\beta}^{\mathrm T}\mathbf{X})>0\}}{[f_{\boldsymbol{\beta}^{\mathrm T}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm T}\mathbf{X}) E\{I(Z\geqslant s)\mid \boldsymbol{\beta}^{\mathrm T}\mathbf{X}\}]^2}$$

$$\times \left[f_{\boldsymbol{\beta}^{\mathrm T}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm T}\mathbf{X}) E\{I(Z_i\geqslant s)\mid \boldsymbol{\beta}^{\mathrm T}\mathbf{X}\} \int K^2(u)du + O_p(n^{-1/2}h^{-1/2}+h^2)\right] ds$$

$$\xrightarrow{p} \int K^2(u)du \int_0^t \frac{\lambda(s,\boldsymbol{\beta}^{\mathrm T}\mathbf{X})}{f_{\boldsymbol{\beta}^{\mathrm T}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm T}\mathbf{X}) E\{I(Z_j\geqslant s)\mid \boldsymbol{\beta}^{\mathrm T}\mathbf{X}\}} ds$$

(S27)$= \sigma^2(t,\boldsymbol{\beta}^{\mathrm T}\mathbf{X})$.

Next we inspect the following integration for any $\epsilon > 0$.

$$\sum_{i=1}^n \int_0^t \frac{h}{n} \frac{I\{\phi_n(s,\boldsymbol{\beta}^{\mathrm T}\mathbf{X})>0\}}{[f_{\boldsymbol{\beta}^{\mathrm T}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm T}\mathbf{X}) E\{I(Z\geqslant s)\mid \boldsymbol{\beta}^{\mathrm T}\mathbf{X}\}]^2} K_h^2(\boldsymbol{\beta}^{\mathrm T}\mathbf{X}_i - \boldsymbol{\beta}^{\mathrm T}\mathbf{X}) Y_i(s)$$

$$\times I\left[\sqrt{\frac{h}{n}} \left| \frac{I\{\phi_n(s,\boldsymbol{\beta}^{\mathrm T}\mathbf{X})>0\}}{f_{\boldsymbol{\beta}^{\mathrm T}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm T}\mathbf{X}) E\{I(Z\geqslant s)\mid \boldsymbol{\beta}^{\mathrm T}\mathbf{X}\}} K_h(\boldsymbol{\beta}^{\mathrm T}\mathbf{X}_i - \boldsymbol{\beta}^{\mathrm T}\mathbf{X})\right| > \epsilon \right] \lambda(s,\boldsymbol{\beta}^{\mathrm T}\mathbf{X}) ds$$

$$= \int_0^t \frac{I\{\phi_n(s,\boldsymbol{\beta}^{\mathrm T}\mathbf{X})>0\}}{[f_{\boldsymbol{\beta}^{\mathrm T}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm T}\mathbf{X}) E\{I(Z\geqslant s)\mid \boldsymbol{\beta}^{\mathrm T}\mathbf{X}\}]^2} \frac{1}{n}\sum_{i=1}^n \Bigg(hK_h^2(\boldsymbol{\beta}^{\mathrm T}\mathbf{X}_i - \boldsymbol{\beta}^{\mathrm T}\mathbf{X}) Y_i(s)$$

$$\times I\left[\sqrt{\frac{h}{n}} \left| \frac{I\{\phi_n(s,\boldsymbol{\beta}^{\mathrm T}\mathbf{X})>0\}}{f_{\boldsymbol{\beta}^{\mathrm T}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm T}\mathbf{X}) E\{I(Z\geqslant s)\mid \boldsymbol{\beta}^{\mathrm T}\mathbf{X}\}} K_h(\boldsymbol{\beta}^{\mathrm T}\mathbf{X}_i - \boldsymbol{\beta}^{\mathrm T}\mathbf{X})\right| > \epsilon \right]\Bigg) \lambda(s,\boldsymbol{\beta}^{\mathrm T}\mathbf{X}) ds$$

$$\leqslant \int_0^t \lambda(s,\boldsymbol{\beta}^{\mathrm T}\mathbf{X}) \frac{I\{\phi_n(s,\boldsymbol{\beta}^{\mathrm T}\mathbf{X})>0\}}{[f_{\boldsymbol{\beta}^{\mathrm T}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm T}\mathbf{X}) E\{I(Z\geqslant s)\mid \boldsymbol{\beta}^{\mathrm T}\mathbf{X}\}]^2} \left\{ \frac{1}{n}\sum_{i=1}^n hK_h^2(\boldsymbol{\beta}^{\mathrm T}\mathbf{X}_i - \boldsymbol{\beta}^{\mathrm T}\mathbf{X}) Y_i(s)\right\}$$

$$\times \sup_{1\leqslant i\leqslant n} I\left[\left| f_{\boldsymbol{\beta}^{\mathrm T}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm T}\mathbf{X}) E\{I(Z\geqslant s)\mid \boldsymbol{\beta}^{\mathrm T}\mathbf{X}\}\right| < \frac{\xi_i}{\epsilon\sqrt{nh}}\right] ds,$$

where $\xi_i = |I\{\phi_n(s,\boldsymbol{\beta}^{\mathrm T}\mathbf{X})>0\} K\{(\boldsymbol{\beta}^{\mathrm T}\mathbf{X}_i - \boldsymbol{\beta}^{\mathrm T}\mathbf{X})/h\}|$, which is bounded following Condition C1. In the above display,

$$\sup_{1\leqslant i\leqslant n} I\left[\left| f_{\boldsymbol{\beta}^{\mathrm T}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm T}\mathbf{X}) E\{I(Z_j\geqslant s)\mid \boldsymbol{\beta}^{\mathrm T}\mathbf{X}\}\right| < \frac{\xi_i}{\epsilon\sqrt{nh}}\right] = 0$$

as long as n is sufficiently large, because the right hand side converges to 0 by Condition C2 while the left hand side will be bounded away from 0 by Conditions C3 and C5. On the other hand,

$$\frac{1}{n}\sum_{i=1}^n hK_h^2(\boldsymbol{\beta}^{\mathrm T}\mathbf{X}_i - \boldsymbol{\beta}^{\mathrm T}\mathbf{X}) Y_i(s)$$

$$= f_{\boldsymbol{\beta}^{\mathrm T}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm T}\mathbf{X}) E\{I(Z_i\geqslant s)\mid \boldsymbol{\beta}^{\mathrm T}\mathbf{X}\} \int K^2(u)du + O_p(n^{-1/2}h^{-1/2}+h^2)$$

$$\to 0$$

in probability uniformly. Hence

$$\lim_{n\to\infty} \sum_{i=1}^n \int_0^t \frac{h}{n} \frac{I\{\phi_n(s,\boldsymbol{\beta}^{\mathrm T}\mathbf{X})>0\}}{[f_{\boldsymbol{\beta}^{\mathrm T}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm T}\mathbf{X}) E\{I(Z\geqslant s)\mid \boldsymbol{\beta}^{\mathrm T}\mathbf{X}\}]^2} K_h^2(\boldsymbol{\beta}^{\mathrm T}\mathbf{X}_i - \boldsymbol{\beta}^{\mathrm T}\mathbf{X}) Y_i(s)$$

$$\times I\left[\sqrt{\frac{h}{n}} \left| \frac{I\{\phi_n(s,\boldsymbol{\beta}^{\mathrm T}\mathbf{X})>0\}}{f_{\boldsymbol{\beta}^{\mathrm T}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm T}\mathbf{X}) E\{I(Z\geqslant s)\mid \boldsymbol{\beta}^{\mathrm T}\mathbf{X}\}} K_h(\boldsymbol{\beta}^{\mathrm T}\mathbf{X}_i - \boldsymbol{\beta}^{\mathrm T}\mathbf{X})\right| > \epsilon \right] \lambda(s,\boldsymbol{\beta}^{\mathrm T}\mathbf{X}) ds = 0$$

(S28)

with probability 1 uniformly for any $\epsilon > 0$.

In summary

$$\sqrt{nh}\left\{\widehat{\Lambda}(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})-\Lambda(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})\right\}\rightarrow N\{0,\sigma^2(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})\}$$

uniformly. □

4.3. *Proof of Theorem 1.* Because the result regarding (10) is the most difficult to establish, we provide only the proof concerning (10), the result concerning (9) is based on a similar proof.

For each n, let $\widehat{\boldsymbol{\beta}}_n$ satisfy

$$\frac{1}{n}\sum_{i=1}^{n}\Delta_i\frac{\widehat{\boldsymbol{\lambda}}_2\{Z_i,\widehat{\boldsymbol{\beta}}_n^{\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\}}{\widehat{\lambda}\{Z_i,\widehat{\boldsymbol{\beta}}_n^{\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\}}$$
$$\otimes\left[\mathbf{X}_{li}-\frac{\widehat{E}\left\{\mathbf{X}_{li}Y_i(Z_i)\mid\widehat{\boldsymbol{\beta}}_n^{\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\right\}}{\widehat{E}\left\{Y_i(Z_i)\mid\widehat{\boldsymbol{\beta}}_n^{\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\right\}}\right]=\mathbf{0}.$$

Under condition C6, there exists a subsequence of $\widehat{\boldsymbol{\beta}}_n, n=1,2,\ldots$, that converges. For notational simplicity, we still write $\widehat{\boldsymbol{\beta}}_n, n=1,2,\ldots$, as the subsequence that converges and let the limit be $\boldsymbol{\beta}^*$.

From the uniform convergence in (S9), (S10), (S11), (S12) given in Lemma 1,

$$\frac{1}{n}\sum_{i=1}^{n}\Delta_i\frac{\widehat{\boldsymbol{\lambda}}_2\{Z_i,\widehat{\boldsymbol{\beta}}_n^{\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\}}{\widehat{\lambda}\{Z_i,\widehat{\boldsymbol{\beta}}_n^{\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\}}$$
$$\otimes\left[\mathbf{X}_{li}-\frac{\widehat{E}\left\{\mathbf{X}_{li}Y_i(Z_i)\mid\widehat{\boldsymbol{\beta}}_n^{\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\right\}}{\widehat{E}\left\{Y_i(Z_i)\mid\widehat{\boldsymbol{\beta}}_n^{\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\right\}}\right]$$
$$=\frac{1}{n}\sum_{i=1}^{n}\Delta_i\frac{\boldsymbol{\lambda}_2\{Z_i,\widehat{\boldsymbol{\beta}}_n^{\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\}+O_p\{(nbh^3)^{-1/2}+h^2+b^2\}}{\lambda\{Z_i,\widehat{\boldsymbol{\beta}}_n^{\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\}+O_p\{(nbh)^{-1/2}+h^2+b^2\}}$$
$$\otimes\left[\mathbf{X}_{li}-\frac{E\left\{\mathbf{X}_{li}Y_i(Z_i)\mid\widehat{\boldsymbol{\beta}}_n^{\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\right\}+O_p\{(nh)^{-1/2}+h^2\}}{E\left\{Y_i(Z_i)\mid\widehat{\boldsymbol{\beta}}_n^{\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\right\}+O_p\{(nh)^{-1/2}+h^2\}}\right]$$
$$=\frac{1}{n}\sum_{i=1}^{n}\Delta_i\left[\frac{\boldsymbol{\lambda}_2\{Z_i,\widehat{\boldsymbol{\beta}}_n^{\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\}}{\lambda\{Z_i,\widehat{\boldsymbol{\beta}}_n^{\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\}}+O_p\{(nbh^3)^{-1/2}+h^2+b^2\}\right]$$
$$\otimes\left[\mathbf{X}_{li}-\frac{E\left\{\mathbf{X}_{li}Y_i(Z_i)\mid\widehat{\boldsymbol{\beta}}_n^{\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\right\}}{E\left\{Y_i(Z_i)\mid\widehat{\boldsymbol{\beta}}_n^{\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\right\}}+O_p\{(nh)^{-1/2}+h^2\}\right]$$
$$=\frac{1}{n}\sum_{i=1}^{n}\Delta_i\frac{\boldsymbol{\lambda}_2\{Z_i,\widehat{\boldsymbol{\beta}}_n^{\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\}}{\lambda\{Z_i,\widehat{\boldsymbol{\beta}}_n^{\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\}}$$

$$\otimes\left[\mathbf{X}_{li}-\frac{E\left\{\mathbf{X}_{li}Y_i(Z_i)\mid\widehat{\boldsymbol{\beta}}_n^{\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\right\}}{E\left\{Y_i(Z_i)\mid\widehat{\boldsymbol{\beta}}_n^{\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\right\}}\right]+o_p(1).$$

Thus, for sufficiently large n,

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^n\Delta_i\frac{\boldsymbol{\lambda}_2\{Z_i,\widehat{\boldsymbol{\beta}}_n^{\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\}}{\lambda\{Z_i,\widehat{\boldsymbol{\beta}}_n^{\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\}}\\
&\quad\otimes\left[\mathbf{X}_{li}-\frac{E\left\{\mathbf{X}_{li}Y_i(Z_i)\mid\widehat{\boldsymbol{\beta}}_n^{\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\right\}}{E\left\{Y_i(Z_i)\mid\widehat{\boldsymbol{\beta}}_n^{\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\right\}}\right]\\
&=\frac{1}{n}\sum_{i=1}^n\Delta_i\frac{\boldsymbol{\lambda}_2\{Z_i,\boldsymbol{\beta}^{*\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\}}{\lambda\{Z_i,\boldsymbol{\beta}^{*\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\}}\\
&\quad\otimes\left[\mathbf{X}_{li}-\frac{E\left\{\mathbf{X}_{li}Y_i(Z_i)\mid\boldsymbol{\beta}^{*\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\right\}}{E\left\{Y_i(Z_i)\mid\boldsymbol{\beta}^{*\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\right\}}\right]+O_p(\widehat{\boldsymbol{\beta}}_n-\boldsymbol{\beta}^*)\\
&=\frac{1}{n}\sum_{i=1}^n\Delta_i\frac{\boldsymbol{\lambda}_2\{Z_i,\boldsymbol{\beta}^{*\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\}}{\lambda\{Z_i,\boldsymbol{\beta}^{*\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\}}\\
&\quad\otimes\left[\mathbf{X}_{li}-\frac{E\left\{\mathbf{X}_{li}Y_i(Z_i)\mid\boldsymbol{\beta}^{*\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\right\}}{E\left\{Y_i(Z_i)\mid\boldsymbol{\beta}^{*\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\right\}}\right]+o_p(1),
\end{aligned}$$

under Conditions C1-C2, where the last equality holds because $\widehat{\boldsymbol{\beta}}_n$ converges to $\boldsymbol{\beta}^*$. In addition,

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^n\Delta_i\frac{\boldsymbol{\lambda}_2\{Z_i,\boldsymbol{\beta}^{*\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\}}{\lambda\{Z_i,\boldsymbol{\beta}^{*\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\}}\\
&\quad\otimes\left[\mathbf{X}_{li}-\frac{E\left\{\mathbf{X}_{li}Y_i(Z_i)\mid\boldsymbol{\beta}^{*\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\right\}}{E\left\{Y_i(Z_i)\mid\boldsymbol{\beta}^{*\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\right\}}\right]\\
&=E\left(\Delta\frac{\boldsymbol{\lambda}_2\{Z,\boldsymbol{\beta}^{*\mathrm{T}}\mathbf{X},I(W\leqslant Z),WI(W\leqslant Z)\}}{\lambda\{Z,\boldsymbol{\beta}^{*\mathrm{T}}\mathbf{X},I(W\leqslant Z),WI(W\leqslant Z)\}}\right.\\
&\quad\left.\otimes\left[\mathbf{X}_l-\frac{E\left\{\mathbf{X}_lY(Z)\mid\boldsymbol{\beta}^{*\mathrm{T}}\mathbf{X},I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\right\}}{E\left\{Y(Z)\mid\boldsymbol{\beta}^{*\mathrm{T}}\mathbf{X},I(W\leqslant Z),WI(W\leqslant Z)\right\}}\right]\right)\\
&\quad+o_p(1)
\end{aligned}$$

under Conditions C1-C2. Thus, for sufficient large n

$$\mathbf{0}=\frac{1}{n}\sum_{i=1}^n\Delta_i\frac{\widehat{\boldsymbol{\lambda}}_2\{Z_i,\widehat{\boldsymbol{\beta}}_n^{\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\}}{\widehat{\lambda}\{Z_i,\widehat{\boldsymbol{\beta}}_n^{\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\}}$$

$$\otimes\left[\mathbf{X}_{li}-\frac{\widehat{E}\left\{\mathbf{X}_{li}Y_i(Z_i)\mid\widehat{\boldsymbol\beta}_n^{\mathrm T}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\right\}}{\widehat{E}\left\{Y_i(Z_i)\mid\widehat{\boldsymbol\beta}_n^{\mathrm T}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\right\}}\right]$$

$$=\frac{1}{n}\sum_{i=1}^n\Delta_i\frac{\boldsymbol\lambda_2\{Z_i,\widehat{\boldsymbol\beta}_n^{\mathrm T}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\}}{\lambda\{Z_i,\widehat{\boldsymbol\beta}_n^{\mathrm T}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\}}$$

$$\otimes\left[\mathbf{X}_{li}-\frac{E\left\{\mathbf{X}_{li}Y_i(Z_i)\mid\widehat{\boldsymbol\beta}_n^{\mathrm T}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\right\}}{E\left\{Y_i(Z_i)\mid\widehat{\boldsymbol\beta}_n^{\mathrm T}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\right\}}\right]+o_p(1)$$

$$=\frac{1}{n}\sum_{i=1}^n\Delta_i\frac{\boldsymbol\lambda_2\{Z_i,\boldsymbol\beta^{*\mathrm T}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\}}{\lambda\{Z_i,\boldsymbol\beta^{*\mathrm T}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\}}$$

$$\otimes\left[\mathbf{X}_{li}-\frac{E\left\{\mathbf{X}_{li}Y_i(Z_i)\mid\boldsymbol\beta^{*\mathrm T}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\right\}}{E\left\{Y_i(Z_i)\mid\boldsymbol\beta^{*\mathrm T}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\right\}}\right]+o_p(1)$$

$$=E\left(\Delta\frac{\boldsymbol\lambda_2\{Z,\boldsymbol\beta^{*\mathrm T}\mathbf{X},I(W\leqslant Z),WI(W\leqslant Z)\}}{\lambda\{Z,\boldsymbol\beta^{*\mathrm T}\mathbf{X},I(W\leqslant Z),WI(W\leqslant Z)\}}\right.$$

$$\left.\otimes\left[\mathbf{X}_l-\frac{E\left\{\mathbf{X}_lY(Z)\mid\boldsymbol\beta^{*\mathrm T}\mathbf{X},I(W\leqslant Z),WI(W\leqslant Z)\right\}}{E\left\{Y(Z)\mid\boldsymbol\beta^{*\mathrm T}\mathbf{X},I(W\leqslant Z),WI(W\leqslant Z)\right\}}\right]\right)+o_p(1)$$

under Conditions C1-C2 and C6. Note that

$$E\left(\Delta\frac{\boldsymbol\lambda_2\{Z,\boldsymbol\beta^{*\mathrm T}\mathbf{X},I(W\leqslant Z),WI(W\leqslant Z)\}}{\lambda\{Z,\boldsymbol\beta^{*\mathrm T}\mathbf{X},I(W\leqslant Z),WI(W\leqslant Z)\}}\right.$$

$$\left.\otimes\left[\mathbf{X}_l-\frac{E\left\{\mathbf{X}_lY(Z)\mid\boldsymbol\beta^{*\mathrm T}\mathbf{X},I(W\leqslant Z),WI(W\leqslant Z)\right\}}{E\left\{Y(Z)\mid\boldsymbol\beta^{*\mathrm T}\mathbf{X},I(W\leqslant Z),WI(W\leqslant Z)\right\}}\right]\right)$$

is a nonrandom quantity that does not depend on n, hence it is zero. Thus the uniqueness requirement in Condition C7 ensures that $\boldsymbol\beta^*=\boldsymbol\beta_0$.

We show by contradiction that the subsequence that converges includes all but a finite number of n's. Suppose that, for contradiction, this were not true, then we could obtain an infinite sequence of $\widehat{\boldsymbol\beta}_n$'s that did not converge to $\boldsymbol\beta_0$. Given that this infinite sequence lies in a compact set $\mathcal{B}$, we can extract another subsequence of $\widehat{\boldsymbol\beta}_n$'s that converges to, say, $\boldsymbol\beta^{**}$. However, by using the prior derivations, we conclude $\boldsymbol\beta^{**}=\boldsymbol\beta_0$, a contradiction to that $\widehat{\boldsymbol\beta}_n$'s did not converge to $\boldsymbol\beta^0$. Thus we conclude $\widehat{\boldsymbol\beta}-\boldsymbol\beta_0\to\mathbf{0}$ in probability when $n\to\infty$ under Conditions C1-C6. □

4.4. *Proof of Theorem 2.* The terms $\mathbf{A}$ and $\mathbf{B}$ in Theorem 2 are

$$\mathbf{A}=E\left\{\frac{\partial}{\partial\mathrm{vecl}(\boldsymbol\beta)^{\mathrm T}}\mathrm{vecl}\left(\Delta\mathbf{g}\{Z,\boldsymbol\beta^{\mathrm T}\mathbf{X},I(W\leqslant Z),WI(W\leqslant Z)\}\right.\right.$$

$$\left.\left.\otimes\left[\mathbf{a}(\mathbf{X}_l)-\frac{E\left\{\mathbf{a}(\mathbf{X}_l)Y(Z)\mid\boldsymbol\beta^{\mathrm T}\mathbf{X},I(W\leqslant Z),WI(W\leqslant Z)\right\}}{E\left\{Y(Z)\mid\boldsymbol\beta^{\mathrm T}\mathbf{X},I(W\leqslant Z),WI(W\leqslant Z)\right\}}\right]\right)\right\},$$

$$\mathbf{B} = E\Big\{\text{vecl}\Big(\Delta\mathbf{g}\{Z, \boldsymbol\beta^{\mathrm T}\mathbf{X}, I(W \leqslant Z), WI(W \leqslant Z)\} \otimes \left[\mathbf{a}(\mathbf{X}_l) - \frac{E\{\mathbf{a}(\mathbf{X}_l)Y(Z) \mid \boldsymbol\beta^{\mathrm T}\mathbf{X}, I(W \leqslant Z), WI(W \leqslant Z)\}}{E\{Y(Z) \mid \boldsymbol\beta^{\mathrm T}\mathbf{X}, I(W \leqslant Z), WI(W \leqslant Z)\}}\right]\Big)^{\otimes 2}\Big\}.$$

Here $\text{vecl}(\mathbf{A})$ represents the vectorization of the lower $(p-d)\times d$ block of a generic matrix $\mathbf{A}$ and $\mathbf{A}^{\otimes 2} = \mathbf{A}\mathbf{A}^{\mathrm T}$ for any matrix or vector $\mathbf{A}$. Note that in (9), $\mathbf{a}(\mathbf{X}_l) = \mathbf{X}_l$ and in (10), $\mathbf{a}(\mathbf{X}_l) = \mathbf{X}_l$, $g\{Z, \boldsymbol\beta^{\mathrm T}\mathbf{X}, I(W \leqslant Z), WI(W \leqslant Z)\} = \widehat{\mathbf{m}}_{12}\{Z, \boldsymbol\beta^{\mathrm T}\mathbf{X}, I(W \leqslant Z), WI(W \leqslant Z)\}/\{\widehat{m}_1\{Z, \boldsymbol\beta^{\mathrm T}\mathbf{X}, I(W \leqslant Z), WI(W \leqslant Z)\} + 1\} - \widehat{\mathbf{m}}_2\{Z, \boldsymbol\beta^{\mathrm T}\mathbf{X}, I(W \leqslant Z), WI(W \leqslant Z)\}/\widehat{m}\{Z, \boldsymbol\beta^{\mathrm T}\mathbf{X}, I(W \leqslant Z), WI(W \leqslant Z)\}$.

We only present the proof concerning (10); the result concerning (9) follows with a similar and simpler proof.

We first expand (10) as

$$\text{(S29)}\quad \mathbf{0} = n^{-1/2}\sum_{i=1}^n \Delta_i \frac{\widehat{\boldsymbol\lambda}_2\{Z_i, \widehat{\boldsymbol\beta}^{\mathrm T}\mathbf{X}_i, I(W_i \leqslant Z_i), W_iI(W_i \leqslant Z_i)\}}{\widehat\lambda\{Z_i, \widehat{\boldsymbol\beta}^{\mathrm T}\mathbf{X}_i, I(W_i \leqslant Z_i), W_iI(W_i \leqslant Z_i)\}} \otimes \left[\mathbf{X}_{li} - \frac{\widehat E\left\{\mathbf{X}_{li}Y_i(Z_i) \mid \widehat{\boldsymbol\beta}^{\mathrm T}\mathbf{X}_i, I(W_i \leqslant Z_i), W_iI(W_i \leqslant Z_i)\right\}}{\widehat E\left\{Y_i(Z_i) \mid \widehat{\boldsymbol\beta}^{\mathrm T}\mathbf{X}_i, I(W_i \leqslant Z_i), W_iI(W_i \leqslant Z_i)\right\}}\right]$$

$$\text{(S30)}\quad = n^{-1/2}\sum_{i=1}^n \Delta_i \frac{\widehat{\boldsymbol\lambda}_2\{Z_i, \boldsymbol\beta_0^{\mathrm T}\mathbf{X}_i, I(W_i \leqslant Z_i), W_iI(W_i \leqslant Z_i)\}}{\widehat\lambda\{Z_i, \boldsymbol\beta_0^{\mathrm T}\mathbf{X}_i, I(W_i \leqslant Z_i), W_iI(W_i \leqslant Z_i)\}} \otimes \left[\mathbf{X}_{li} - \frac{\widehat E\{\mathbf{X}_{li}Y_i(Z_i) \mid \boldsymbol\beta_0^{\mathrm T}\mathbf{X}_i, I(W_i \leqslant Z_i), W_iI(W_i \leqslant Z_i)\}}{\widehat E\{Y_i(Z_i) \mid \boldsymbol\beta_0^{\mathrm T}\mathbf{X}_i, I(W_i \leqslant Z_i), W_iI(W_i \leqslant Z_i)\}}\right]$$

$$\text{(S31)}\quad + \frac{1}{n}\sum_{i=1}^n \left\{\frac{\partial}{\partial(\mathbf{X}_i^{\mathrm T}\boldsymbol\beta)}\left(\Delta_i \frac{\widehat{\boldsymbol\lambda}_2\{Z_i, \boldsymbol\beta^{\mathrm T}\mathbf{X}_i, I(W_i \leqslant Z_i), W_iI(W_i \leqslant Z_i)\}}{\widehat\lambda\{Z_i, \boldsymbol\beta^{\mathrm T}\mathbf{X}_i, I(W_i \leqslant Z_i), W_iI(W_i \leqslant Z_i)\}}\right.\right.$$

$$\text{(S32)}\quad \left.\left.\otimes \left[\mathbf{X}_{li} - \frac{\widehat E\{\mathbf{X}_{li}Y_i(Z_i) \mid \boldsymbol\beta^{\mathrm T}\mathbf{X}_i, I(W_i \leqslant Z_i), W_iI(W_i \leqslant Z_i)\}}{\widehat E\{Y_i(Z_i) \mid \boldsymbol\beta^{\mathrm T}\mathbf{X}_i, I(W_i \leqslant Z_i), W_iI(W_i \leqslant Z_i)\}}\right]\right) \otimes \mathbf{X}_{li}^{\mathrm T}\right\}\Bigg|_{\boldsymbol\beta = \widetilde{\boldsymbol\beta}} \times \sqrt{n}(\widehat{\boldsymbol\beta} - \boldsymbol\beta_0),$$

where $\widetilde{\boldsymbol\beta}$ is on the line connecting $\boldsymbol\beta_0$ and $\widehat{\boldsymbol\beta}$.

We first consider (S32). Because of Theorem 1 and Lemma 1,

$$\frac{1}{n}\sum_{i=1}^n \left\{\frac{\partial}{\partial(\mathbf{X}_i^{\mathrm T}\boldsymbol\beta)}\left(\Delta_i \frac{\widehat{\boldsymbol\lambda}_2\{Z_i, \boldsymbol\beta^{\mathrm T}\mathbf{X}_i, I(W_i \leqslant Z_i), W_iI(W_i \leqslant Z_i)\}}{\widehat\lambda\{Z_i, \boldsymbol\beta^{\mathrm T}\mathbf{X}_i, I(W_i \leqslant Z_i), W_iI(W_i \leqslant Z_i)\}} \otimes \left[\mathbf{X}_{li} - \frac{\widehat E\{\mathbf{X}_{li}Y_i(Z_i) \mid \boldsymbol\beta^{\mathrm T}\mathbf{X}_i, I(W_i \leqslant Z_i), W_iI(W_i \leqslant Z_i)\}}{\widehat E\{Y_i(Z_i) \mid \boldsymbol\beta^{\mathrm T}\mathbf{X}_i, I(W_i \leqslant Z_i), W_iI(W_i \leqslant Z_i)\}}\right]\right) \otimes \mathbf{X}_{li}^{\mathrm T}\right\}\Bigg|_{\boldsymbol\beta = \widetilde{\boldsymbol\beta}}$$

$$= \frac{1}{n}\sum_{i=1}^n \left\{\frac{\partial}{\partial(\mathbf{X}_i^{\mathrm T}\boldsymbol\beta_0)}\left(\Delta_i \frac{\widehat{\boldsymbol\lambda}_2\{Z_i, \boldsymbol\beta_0^{\mathrm T}\mathbf{X}_i, I(W_i \leqslant Z_i), W_iI(W_i \leqslant Z_i)\}}{\widehat\lambda\{Z_i, \boldsymbol\beta_0^{\mathrm T}\mathbf{X}_i, I(W_i \leqslant Z_i), W_iI(W_i \leqslant Z_i)\}} \otimes \left[\mathbf{X}_{li} - \frac{\widehat E\{\mathbf{X}_{li}Y_i(Z_i) \mid \boldsymbol\beta_0^{\mathrm T}\mathbf{X}_i, I(W_i \leqslant Z_i), W_iI(W_i \leqslant Z_i)\}}{\widehat E\{Y_i(Z_i) \mid \boldsymbol\beta_0^{\mathrm T}\mathbf{X}_i, I(W_i \leqslant Z_i), W_iI(W_i \leqslant Z_i)\}}\right]\right) \otimes \mathbf{X}_{li}^{\mathrm T}\right\} + o_p(1)$$

$$= -\frac{1}{n}\sum_{i=1}^{n}\left(\Delta_i \frac{\widehat{\boldsymbol{\lambda}}_2^{\otimes 2}\{Z_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i \leqslant Z_i), W_i I(W_i \leqslant Z_i)\}}{\widehat{\lambda}^2\{Z_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i \leqslant Z_i), W_i I(W_i \leqslant Z_i)\}}\right.$$

$$\left.\otimes\left[\mathbf{X}_{li} - \frac{\widehat{E}\left\{\mathbf{X}_{li}Y_i(Z_i) \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i \leqslant Z_i), W_i I(W_i \leqslant Z_i)\right\}}{\widehat{E}\left\{Y_i(Z_i) \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i \leqslant Z_i), W_i I(W_i \leqslant Z_i)\right\}}\right] \otimes \mathbf{X}_{li}^{\mathrm{T}}\right)$$

(S33)

$$+\frac{1}{n}\sum_{i=1}^{n}\frac{\Delta_i}{\widehat{\lambda}\{Z_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i \leqslant Z_i), W_i I(W_i \leqslant Z_i)\}}$$

$$\frac{\partial}{\partial(\mathbf{X}_i^{\mathrm{T}}\boldsymbol{\beta}_0)}\left(\widehat{\boldsymbol{\lambda}}_2\{Z_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i \leqslant Z_i), W_i I(W_i \leqslant Z_i)\}\right.$$

$$\left.\otimes\left[\mathbf{X}_{li} - \frac{\widehat{E}\left\{\mathbf{X}_{li}Y_i(Z_i) \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i \leqslant Z_i), W_i I(W_i \leqslant Z_i)\right\}}{\widehat{E}\left\{Y_i(Z_i) \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i \leqslant Z_i), W_i I(W_i \leqslant Z_i)\right\}}\right]\right) \otimes \mathbf{X}_{li}^{\mathrm{T}}$$

(S34)

$$+o_p(1).$$

Because of Lemma 1, (S33) converges uniformly in probability to

$$-E\left(\int \frac{\boldsymbol{\lambda}_2^{\otimes 2}\{s, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\}}{\lambda^2\{s, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\}}\right.$$

$$\left.\otimes\left[\mathbf{X}_l - \frac{E\left\{\mathbf{X}_l Y(s) \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\right\}}{E\left\{Y(s) \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\right\}}\right] \otimes \mathbf{X}_l^{\mathrm{T}} dN(s)\right)$$

$$= -E\left(\int \frac{\boldsymbol{\lambda}_2^{\otimes 2}\{s, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\}}{\lambda^2\{s, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\}}\right.$$

$$\otimes\left[\mathbf{X}_l - \frac{E\left\{\mathbf{X}_l Y(s) \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\right\}}{E\left\{Y(s) \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\right\}}\right]$$

$$\left.\otimes \mathbf{X}_l^{\mathrm{T}} Y(s)\lambda\{s, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\}ds\right)$$

$$= -E\left(\int \frac{\boldsymbol{\lambda}_2^{\otimes 2}\{s, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\}}{\lambda\{s, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\}}\right.$$

$$\otimes\left[\mathbf{X}_l - \frac{E\left\{\mathbf{X}_l Y(s) \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\right\}}{E\left\{Y(s) \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\right\}}\right]$$

$$\left.\otimes\left[\mathbf{X}_l - \frac{E\left\{\mathbf{X}_l Y(s) \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\right\}}{E\left\{Y(s) \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\right\}}\right]^{\mathrm{T}} Y(s)ds\right)$$

$$-E\left(\int \frac{\boldsymbol{\lambda}_2^{\otimes 2}\{s, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\}}{\lambda\{s, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\}}\right.$$

$$\otimes\left[\mathbf{X}_l - \frac{E\left\{\mathbf{X}_l Y(s) \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\right\}}{E\left\{Y(s) \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}, I(W \leqslant s), WI(W \leqslant s)\right\}}\right]$$

$$
\begin{aligned}
&\otimes \frac{E\left\{\mathbf{X}_l Y(s) \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}, I(W\leqslant s), WI(W\leqslant s)\right\}}{E\left\{Y(s) \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}, I(W\leqslant s), WI(W\leqslant s)\right\}}^{\mathrm{T}} Y(s)ds\Bigg) \\
&= -E\{\mathbf{S}_{\mathrm{eff}}(\Delta, Z, \mathbf{X})^{\otimes 2}\},
\end{aligned}
$$

where the last equality is due to that the second term above is zero by first taking expectation conditional on $\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}$.

Similarly, from Lemma 1, the term in (S34) converges uniformly in probability to the limit of

$$
\begin{aligned}
&E\Bigg\{\frac{\Delta_i}{\lambda\{Z_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i\leqslant Z_i), W_iI(W_i\leqslant Z_i)\}} \\
&\times \frac{\partial}{\partial(\mathbf{X}_i^{\mathrm{T}}\boldsymbol{\beta}_0)}\Big(\widehat{\boldsymbol{\lambda}}_2\{Z_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i\leqslant Z_i), W_iI(W_i\leqslant Z_i)\} \\
&\otimes\left[\mathbf{X}_{li} - \frac{E\left\{\mathbf{X}_{li}Y_i(Z_i) \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i\leqslant Z_i), W_iI(W_i\leqslant Z_i)\right\}}{E\left\{Y_i(Z_i) \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i\leqslant Z_i), W_iI(W_i\leqslant Z_i)\right\}}\right]\Bigg)\otimes\mathbf{X}_{li}^{\mathrm{T}}\Bigg\}.
\end{aligned}
$$

Now let $\widehat{\boldsymbol{\lambda}}_{2,-i}(Z, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X})$ be the leave-one-out version of $\widehat{\boldsymbol{\lambda}}_2(Z, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X})$, i.e. it is constructed the same as $\widehat{\boldsymbol{\lambda}}_2(Z, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X})$ except that the ith observation is not used. Obviously,

$$
\begin{aligned}
&\frac{\Delta_i}{\lambda\{Z_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i\leqslant Z_i), W_iI(W_i\leqslant Z_i)\}} \\
&\quad\times \frac{\partial}{\partial(\mathbf{X}_i^{\mathrm{T}}\boldsymbol{\beta}_0)}\Big(\widehat{\boldsymbol{\lambda}}_2\{Z_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i\leqslant Z_i), W_iI(W_i\leqslant Z_i)\} \\
&\quad\quad\otimes\left[\mathbf{X}_{li} - \frac{E\left\{\mathbf{X}_{li}Y_i(Z_i) \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i\leqslant Z_i), W_iI(W_i\leqslant Z_i)\right\}}{E\left\{Y_i(Z_i) \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i\leqslant Z_i), W_iI(W_i\leqslant Z_i)\right\}}\right]\Bigg)\otimes\mathbf{X}_{li}^{\mathrm{T}} \\
&-\frac{\Delta_i}{\lambda\{Z_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i\leqslant Z_i), W_iI(W_i\leqslant Z_i)\}} \\
&\quad\times \frac{\partial}{\partial(\mathbf{X}_i^{\mathrm{T}}\boldsymbol{\beta}_0)}\Big(\widehat{\boldsymbol{\lambda}}_{2,-i}\{Z_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i\leqslant Z_i), W_iI(W_i\leqslant Z_i)\} \\
&\quad\quad\otimes\left[\mathbf{X}_{li} - \frac{E\left\{\mathbf{X}_{li}Y_i(Z_i) \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i\leqslant Z_i), W_iI(W_i\leqslant Z_i)\right\}}{E\left\{Y_i(Z_i) \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i\leqslant Z_i), W_iI(W_i\leqslant Z_i)\right\}}\right]\Bigg)\otimes\mathbf{X}_{li}^{\mathrm{T}} \\
&= o_p(1).
\end{aligned}
$$

Let E_i mean taking expectation with respect to the ith observation conditional on all other observations, then

$$
\begin{aligned}
&E_i\Bigg\{\frac{\Delta_i}{\lambda\{Z_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i\leqslant Z_i), W_iI(W_i\leqslant Z_i)\}}\frac{\partial}{\partial(\mathbf{X}_i^{\mathrm{T}}\boldsymbol{\beta}_0)} \\
&\times\Big(\widehat{\boldsymbol{\lambda}}_{2,-i}\{Z_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i\leqslant Z_i), W_iI(W_i\leqslant Z_i)\} \\
&\otimes\left[\mathbf{X}_{li} - \frac{E\left\{\mathbf{X}_{li}Y_i(Z_i) \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i\leqslant Z_i), W_iI(W_i\leqslant Z_i)\right\}}{E\left\{Y_i(Z_i) \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i\leqslant Z_i), W_iI(W_i\leqslant Z_i)\right\}}\right]\Bigg)\otimes\mathbf{X}_{li}^{\mathrm{T}}\Bigg\} \\
&= E_i\Bigg\{\frac{\partial}{\partial\boldsymbol{\beta}_0}\int\widehat{\boldsymbol{\lambda}}_{2,-i}\{s, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i\leqslant s), W_iI(W_i\leqslant s)\}
\end{aligned}
$$

$$
\otimes\left[\mathbf{X}_{li}-\frac{E\left\{\mathbf{X}_{li}Y_i(s)\mid\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant s),W_iI(W_i\leqslant s)\right\}}{E\left\{Y_i(s)\mid\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant s),W_iI(W_i\leqslant s)\right\}}\right]E\{Y_i(s)\mid\mathbf{X}_i\}ds\Bigg\}
$$

$$
=\frac{\partial}{\partial\boldsymbol{\beta}_0}E_i\Bigg\{\int\widehat{\boldsymbol{\lambda}}_{2,-i}\{s,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant s),W_iI(W_i\leqslant s)\}
$$

$$
\otimes\left[\mathbf{X}_{li}-\frac{E\left\{\mathbf{X}_{li}Y_i(s)\mid\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant s),W_iI(W_i\leqslant s)\right\}}{E\left\{Y_i(s)\mid\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant s),W_iI(W_i\leqslant s)\right\}}\right]Y_i(s)ds\Bigg\}
$$

$$
=\mathbf{0}.
$$

The last equality holds because the integrand has expectation zero conditional on $\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i$ and all other observations, and the third to last equality follows because the expectation is with respect to $\mathbf{X}_i$ and does not involve $\boldsymbol{\beta}_0$. Therefore, the term in (S34) converges in probability uniformly to

$$
E\Bigg\{\frac{\Delta_i}{\lambda\{Z_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\}}
\times\frac{\partial}{\partial(\mathbf{X}_i^{\mathrm{T}}\boldsymbol{\beta}_0)}\Big(\widehat{\boldsymbol{\lambda}}_{2,-i}\{Z_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\}
\otimes\left[\mathbf{X}_{li}-\frac{E\left\{\mathbf{X}_{li}Y_i(Z_i)\mid\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\right\}}{E\left\{Y_i(Z_i)\mid\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\right\}}\right]\Big)\otimes\mathbf{X}_{li}^{\mathrm{T}}\Bigg\}=0
$$

Combining the results concerning (S33) and (S34), thus the expression in (S32) is $-E\{\mathbf{S}_{\mathrm{eff}}(\Delta,Z,\mathbf{X})^{\otimes 2}\}+o_p(1)$.

Next we decompose (S31) into

$$
\begin{aligned}
&n^{-1/2}\sum_{i=1}^{n}\Delta_i\frac{\widehat{\boldsymbol{\lambda}}_2\{Z_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\}}{\widehat{\lambda}\{Z_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\}}\\
&\quad\otimes\left[\mathbf{X}_{li}-\frac{\widehat{E}\left\{\mathbf{X}_{li}Y_i(Z_i)\mid\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\right\}}{\widehat{E}\left\{Y_i(Z_i)\mid\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\right\}}\right]\\
&=\mathbf{T}_1+\mathbf{T}_2+\mathbf{T}_3+\mathbf{T}_4,
\end{aligned}
\tag{S35}
$$

where

$$
\begin{aligned}
\mathbf{T}_1&=n^{-1/2}\sum_{i=1}^{n}\Delta_i\frac{\boldsymbol{\lambda}_2\{Z_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\}}{\lambda\{Z_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\}}\\
&\quad\otimes\left[\mathbf{X}_{li}-\frac{E\left\{\mathbf{X}_{li}Y_i(Z_i)\mid\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\right\}}{E\left\{Y_i(Z_i)\mid\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\right\}}\right],\\
\mathbf{T}_2&=n^{-1/2}\sum_{i=1}^{n}\Delta_i\left[\frac{\widehat{\boldsymbol{\lambda}}_2\{Z_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\}}{\widehat{\lambda}\{Z_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\}}\right.\\
&\quad\left.-\frac{\boldsymbol{\lambda}_2\{Z_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\}}{\lambda\{Z_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\}}\right]\\
&\quad\otimes\left[\mathbf{X}_{li}-\frac{E\left\{\mathbf{X}_{li}Y_i(Z_i)\mid\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\right\}}{E\left\{Y_i(Z_i)\mid\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i,I(W_i\leqslant Z_i),W_iI(W_i\leqslant Z_i)\right\}}\right],
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{T}_3 = n^{-1/2}\sum_{i=1}^n \Delta_i \frac{\boldsymbol{\lambda}_2\{Z_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i\leqslant Z_i), W_iI(W_i\leqslant Z_i)\}}{\lambda\{Z_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i\leqslant Z_i), W_iI(W_i\leqslant Z_i)\}} \\
\otimes\left[\frac{E\left\{\mathbf{X}_{li}Y_i(Z_i)\mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i\leqslant Z_i), W_iI(W_i\leqslant Z_i)\right\}}{E\left\{Y_i(Z_i)\mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i\leqslant Z_i), W_iI(W_i\leqslant Z_i)\right\}}\right. \\
\left.-\frac{\widehat{E}\left\{\mathbf{X}_{li}Y_i(Z_i)\mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i\leqslant Z_i), W_iI(W_i\leqslant Z_i)\right\}}{\widehat{E}\left\{Y_i(Z_i)\mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i\leqslant Z_i), W_iI(W_i\leqslant Z_i)\right\}}\right],
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{T}_4 = n^{-1/2}\sum_{i=1}^n \Delta_i \left[\frac{\widehat{\boldsymbol{\lambda}}_2\{Z_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i\leqslant Z_i), W_iI(W_i\leqslant Z_i)\}}{\widehat{\lambda}\{Z_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i\leqslant Z_i), W_iI(W_i\leqslant Z_i)\}}\right. \\
\left.-\frac{\boldsymbol{\lambda}_2\{Z_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i\leqslant Z_i), W_iI(W_i\leqslant Z_i)\}}{\lambda\{Z_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i\leqslant Z_i), W_iI(W_i\leqslant Z_i)\}}\right] \\
\otimes\left[\frac{E\left\{\mathbf{X}_{li}Y_i(Z_i)\mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i\leqslant Z_i), W_iI(W_i\leqslant Z_i)\right\}}{E\left\{Y_i(Z_i)\mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i\leqslant Z_i), W_iI(W_i\leqslant Z_i)\right\}}\right. \\
\left.-\frac{\widehat{E}\left\{\mathbf{X}_{li}Y_i(Z_i)\mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i\leqslant Z_i), W_iI(W_i\leqslant Z_i)\right\}}{\widehat{E}\left\{Y_i(Z_i)\mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i\leqslant Z_i), W_iI(W_i\leqslant Z_i)\right\}}\right].
\end{aligned}
$$

First,

$$
\begin{aligned}
\mathbf{T}_2 &= n^{-1/2}\sum_{i=1}^n \int\left[\frac{\widehat{\boldsymbol{\lambda}}_2\{s,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i\leqslant s), W_iI(W_i\leqslant s)\}}{\widehat{\lambda}\{s,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i\leqslant s), W_iI(W_i\leqslant s)\}}\right. \\
&\quad\left.-\frac{\boldsymbol{\lambda}_2\{s,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i\leqslant s), W_iI(W_i\leqslant s)\}}{\lambda\{s,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i\leqslant s), W_iI(W_i\leqslant s)\}}\right] \\
&\quad\otimes\left[\mathbf{X}_{li}-\frac{E\left\{\mathbf{X}_{li}Y_i(s)\mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i\leqslant s), W_iI(W_i\leqslant s)\right\}}{E\left\{Y_i(s)\mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i\leqslant s), W_iI(W_i\leqslant s)\right\}}\right] dN_i(s) \\
&= o_p\left(n^{-1/2}\sum_{i=1}^n \int\left[\mathbf{X}_{li}-\frac{E\left\{\mathbf{X}_{li}Y_i(s)\mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i\leqslant s), W_iI(W_i\leqslant s)\right\}}{E\left\{Y_i(s)\mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i\leqslant s), W_iI(W_i\leqslant s)\right\}}\right]\right. \\
&\quad\left.\times Y_i(s)\lambda(s,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_{li})ds\right) \\
&= o_p(1),
\end{aligned}
$$

where the last equality holds because the quantity inside the parentheses is a mean zero normal random quantity of order $O_p(1)$. Further,

$$
\begin{aligned}
&\mathbf{T}_3 \\
&= n^{-1/2}\sum_{i=1}^n \Delta_i \frac{\boldsymbol{\lambda}_2\{Z_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i\leqslant Z_i), W_iI(W_i\leqslant Z_i)\}}{\lambda\{Z_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i\leqslant Z_i), W_iI(W_i\leqslant Z_i)\}} \\
&\quad\otimes\left(-\frac{\widehat{E}\left\{\mathbf{X}_{li}Y_i(Z_i)\mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i\leqslant Z_i), W_iI(W_i\leqslant Z_i)\right\}}{E\left\{Y_i(Z_i)\mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i\leqslant Z_i), W_iI(W_i\leqslant Z_i)\right\}}\right. \\
&\quad+\frac{\widehat{E}\left\{Y_i(Z_i)\mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i\leqslant Z_i), W_iI(W_i\leqslant Z_i)\right\}}{\left[E\left\{Y_i(Z_i)\mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i\leqslant Z_i), W_iI(W_i\leqslant Z_i)\right\}\right]}
\end{aligned}
$$

$$
\times \frac{E\left\{\mathbf{X}_{li}Y_i(Z_i) \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i \leqslant Z_i), W_i I(W_i \leqslant Z_i)\right\}}{\left[E\left\{Y_i(Z_i) \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i \leqslant Z_i), W_i I(W_i \leqslant Z_i)\right\}\right]}\Bigg) + o_p(1)
$$

$$
= n^{-1/2}\sum_{i=1}^n \Delta_i \frac{\boldsymbol{\lambda}_2\{Z_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i \leqslant Z_i), W_i I(W_i \leqslant Z_i)\}}{\lambda\{Z_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i \leqslant Z_i), W_i I(W_i \leqslant Z_i)\}}
$$

$$
\otimes \Bigg(-\frac{n^{-1}\sum_{j=1}^n K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_j - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i)\mathbf{X}_{lj} I(Z_j \geqslant Z_i)}{f_{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i) E\left\{Y_i(Z_i) \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i \leqslant Z_i), W_i I(W_i \leqslant Z_i)\right\}}
$$

$$
+ \frac{E\left\{\mathbf{X}_{li}Y_i(Z_i) \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i \leqslant Z_i), W_i I(W_i \leqslant Z_i)\right\}}{f_{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i) E\left\{Y_i(Z_i) \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i \leqslant Z_i), W_i I(W_i \leqslant Z_i)\right\}}
$$

$$
\times \left\{ n^{-1}\sum_{j=1}^n K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_j - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i) - f_{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i)\right\}
$$

$$
+ \frac{E\left\{\mathbf{X}_{li}Y_i(Z_i) \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i \leqslant Z_i), W_i I(W_i \leqslant Z_i)\right\}}{\left[E\left\{Y_i(Z_i) \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i \leqslant Z_i), W_i I(W_i \leqslant Z_i)\right\}\right]^2}
$$

$$
\otimes \Bigg[\frac{n^{-1}\sum_{j=1}^n K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_j - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i) I(Z_j \geqslant Z_i)}{f_{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i)}
$$

$$
- \frac{E\left\{Y_i(Z_i) \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i \leqslant Z_i), W_i I(W_i \leqslant Z_i)\right\}\}}{f_{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i)}
$$

$$
\times \frac{\{n^{-1}\sum_{j=1}^n K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_j - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i) - f_{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i)\}}{f_{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i)}\Bigg]\Bigg) + o_p(1)
$$

$$
= n^{-3/2}\sum_{i=1}^n\sum_{j=1}^n \Delta_i \frac{\boldsymbol{\lambda}_2\{Z_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i \leqslant Z_i), W_i I(W_i \leqslant Z_i)\}}{\lambda\{Z_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i \leqslant Z_i), W_i I(W_i \leqslant Z_i)\}}
$$

$$
\otimes \Bigg[-\frac{K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_j - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i)\mathbf{X}_{lj} I(Z_j \geqslant Z_i)}{f_{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i) E\left\{Y_i(Z_i) \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i \leqslant Z_i), W_i I(W_i \leqslant Z_i)\right\}}
$$

$$
+ \frac{E\left\{\mathbf{X}_{li}Y_i(Z_i) \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, W_i\right\} K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_j - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i) I(Z_j \geqslant Z_i)}{f_{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i)\left[E\left\{Y_i(Z_i) \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i \leqslant Z_i), W_i I(W_i \leqslant Z_i)\right\}\right]^2}\Bigg] + o_p(1)
$$

$$
= \mathbf{T}_{31} + \mathbf{T}_{32} + \mathbf{T}_{33} + o_p(1),
$$

where

$$
\mathbf{T}_{31} = n^{-1/2}\sum_{i=1}^n \Delta_i \frac{\boldsymbol{\lambda}_2\{Z_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i \leqslant Z_i), W_i I(W_i \leqslant Z_i)\}}{\lambda\{Z_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i \leqslant Z_i), W_i I(W_i \leqslant Z_i)\}}
$$

$$
\otimes E\Bigg[-\frac{K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_j - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i)\mathbf{X}_{lj} I(Z_j \geqslant Z_i)}{f_{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i) E\left\{Y_i(Z_i) \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i \leqslant Z_i), W_i I(W_i \leqslant Z_i)\right\}}
$$

$$
+ \frac{E\left\{\mathbf{X}_{li}Y_i(Z_i) \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, W_i\right\} K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_j - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i) I(Z_j \geqslant Z_i)}{f_{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i)\left[E\left\{Y_i(Z_i) \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i \leqslant Z_i), W_i I(W_i \leqslant Z_i)\right\}\right]^2}
$$

$$
\mid \Delta_i, Z_i, \mathbf{X}_i, I(W_i \leqslant Z_i), W_i I(W_i \leqslant Z_i)\Big]
$$

$$\mathbf{T}_{32} = n^{-1/2}\sum_{j=1}^{n} E\left(\Delta_i \frac{\boldsymbol{\lambda}_2\{Z_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i \leqslant Z_i), W_i I(W_i \leqslant Z_i)\}}{\lambda\{Z_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i \leqslant Z_i), W_i I(W_i \leqslant Z_i)\}}\right.$$
$$\otimes\left[-\frac{K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_j - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i)\mathbf{X}_{lj} I(Z_j \geqslant Z_i)}{f_{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i) E\left\{Y_i(Z_i) \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i \leqslant Z_i), W_i I(W_i \leqslant Z_i)\right\}}\right.$$
$$+\frac{E\left\{\mathbf{X}_{li} Y_i(Z_i) \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i \leqslant Z_i), W_i I(W_i \leqslant Z_i)\right\}}{f_{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i)[E\left\{Y_i(Z_i) \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i \leqslant Z_i), W_i I(W_i \leqslant Z_i)\right\}]^2}$$
$$\left.\left.\times K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_j - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i) I(Z_j \geqslant Z_i)\right] \mid \Delta_j, Z_j, \mathbf{X}_j, I(W_i \leqslant Z_i), W_i I(W_i \leqslant Z_i)\right)$$

$$\mathbf{T}_{33} = -n^{1/2} E\left(\Delta_i \frac{\boldsymbol{\lambda}_2\{Z_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i \leqslant Z_i), W_i I(W_i \leqslant Z_i)\}}{\lambda\{Z_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i \leqslant Z_i), W_i I(W_i \leqslant Z_i)\}}\right.$$
$$\otimes E\left[-\frac{K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_j - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i)\mathbf{X}_{lj} I(Z_j \geqslant Z_i)}{f_{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i) E\left\{Y_i(Z_i) \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i \leqslant Z_i), W_i I(W_i \leqslant Z_i)\right\}}\right.$$
$$\left.\left.+\frac{E\left\{\mathbf{X}_{li} Y_i(Z_i) \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i \leqslant Z_i), W_i I(W_i \leqslant Z_i)\right\} K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_j - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i) I(Z_j \geqslant Z_i)}{f_{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i)[E\left\{Y_i(Z_i) \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i \leqslant Z_i), W_i I(W_i \leqslant Z_i)\right\}]^2}\right]\right).$$

Here we used U-statistic property in the last equality. Now when $nh^4 \to 0$,

$$\mathbf{T}_{31} = n^{-1/2}\sum_{i=1}^{n} \Delta_i \frac{\boldsymbol{\lambda}_2\{Z_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i \leqslant Z_i), W_i I(W_i \leqslant Z_i)\}}{\lambda\{Z_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i \leqslant Z_i), W_i I(W_i \leqslant Z_i)\}}$$
$$\otimes\left[-\frac{E\{\mathbf{X}_{li} Y_i(Z_i) \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i \leqslant Z_i), W_i I(W_i \leqslant Z_i)\}}{E\left\{Y_i(Z_i) \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i \leqslant Z_i), W_i I(W_i \leqslant Z_i)\right\}}\right.$$
$$+\frac{E\left\{\mathbf{X}_{li} Y_i(Z_i) \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i \leqslant Z_i), W_i I(W_i \leqslant Z_i)\right\}}{[E\left\{Y_i(Z_i) \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i \leqslant Z_i), W_i I(W_i \leqslant Z_i)\right\}]^2}$$
$$\left.\times E\left\{Y_i(Z_i) \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i \leqslant Z_i), W_i I(W_i \leqslant Z_i)\right\}\right] + O(n^{1/2}h^2)$$
$$= o_p(1).$$

Thus, $\mathbf{T}_{33} = o_p(1)$ as well. Also,

$$\mathbf{T}_{32}$$
$$= n^{-1/2}\sum_{j=1}^{n} E\left(\Delta_i \frac{\boldsymbol{\lambda}_2\{Z_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i \leqslant Z_i), W_i I(W_i \leqslant Z_i)\}}{\lambda\{Z_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i \leqslant Z_i), W_i I(W_i \leqslant Z_i)\}}\right.$$
$$\otimes\left[-\frac{K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_j - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i)\mathbf{X}_{lj} I(Z_j \geqslant Z_i)}{f_{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i) E\left\{I(Z \geqslant Z_i) \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X} = \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, Z_i, I(W_i \leqslant Z_i), W_i I(W_i \leqslant Z_i)\right\}}\right.$$
$$\left.+\frac{E\left\{\mathbf{X}_l I(Z \geqslant Z_i) \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X} = \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, Z_i, I(W_i \leqslant Z_i), W_i I(W_i \leqslant Z_i)\right\} K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_j - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i) I(Z_j \geqslant Z_i)}{f_{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i)[E\left\{I(Z \geqslant Z_i) \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X} = \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, Z_i, I(W_i \leqslant Z_i), W_i I(W_i \leqslant Z_i)\right\}]^2}\right]$$
$$\left.\mid \Delta_j, Z_j, \mathbf{X}_j, I(W_i \leqslant Z_i), W_i I(W_i \leqslant Z_i)\right)$$
$$= n^{-1/2}\sum_{j=1}^{n} E\left\{E\left(\Delta_i \frac{\boldsymbol{\lambda}_2\{Z_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i \leqslant Z_i), W_i I(W_i \leqslant Z_i)\}}{\lambda\{Z_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i, I(W_i \leqslant Z_i), W_i I(W_i \leqslant Z_i)\}}\right.\right.$$

$$\otimes\left[-\frac{\mathbf{x}_{lj}I(z_j \geqslant Z_i)}{f_{\boldsymbol\beta_0^{\rm T}\mathbf{X}}(\boldsymbol\beta_0^{\rm T}\mathbf{X}_i)E\left\{I(Z \geqslant Z_i) \mid \boldsymbol\beta_0^{\rm T}\mathbf{X} = \boldsymbol\beta_0^{\rm T}\mathbf{X}_i, Z_i, I(W_i \leqslant Z_i), W_iI(W_i \leqslant Z_i)\right\}}\right.$$
$$\left.+\frac{E\left\{\mathbf{X}_lI(Z \geqslant Z_i) \mid \boldsymbol\beta_0^{\rm T}\mathbf{X} = \boldsymbol\beta_0^{\rm T}\mathbf{X}_i, Z_i, I(W_i \leqslant Z_i), W_iI(W_i \leqslant Z_i)\right\}I(z_j \geqslant Z_i)}{f_{\boldsymbol\beta_0^{\rm T}\mathbf{X}}(\boldsymbol\beta_0^{\rm T}\mathbf{X}_i)[E\left\{I(Z \geqslant Z_i) \mid \boldsymbol\beta_0^{\rm T}\mathbf{X} = \boldsymbol\beta_0^{\rm T}\mathbf{X}_i, Z_i, I(W_i \leqslant Z_i), W_iI(W_i \leqslant Z_i)\right\}]^2}\right]$$
$$\mid \boldsymbol\beta_0^{\rm T}\mathbf{X}_i, I(W_i \leqslant Z_i), W_iI(W_i \leqslant Z_i)\big)\, K_h(\boldsymbol\beta_0^{\rm T}\mathbf{x}_j - \boldsymbol\beta_0^{\rm T}\mathbf{X}_i)\Big\}$$
$$= n^{-1/2}\sum_{j=1}^n E\left(\int_0^{z_j}\frac{\boldsymbol\lambda_2\{s, \boldsymbol\beta_0^{\rm T}\mathbf{x}_j, I(w_i \leqslant s), w_iI(w_i \leqslant s)\}}{E\left\{S_c(s,\mathbf{X}) \mid \boldsymbol\beta_0^{\rm T}\mathbf{X} = \boldsymbol\beta_0^{\rm T}\mathbf{x}_j\right\}}\right.$$
$$\left.\otimes\left[\frac{E\left\{\mathbf{X}_lS_c(s,\mathbf{X}) \mid \boldsymbol\beta_0^{\rm T}\mathbf{X} = \boldsymbol\beta_0^{\rm T}\mathbf{x}_j\right\}}{E\left\{S_c(s,\mathbf{X}) \mid \boldsymbol\beta_0^{\rm T}\mathbf{X} = \boldsymbol\beta_0^{\rm T}\mathbf{x}_j\right\}} - \mathbf{x}_{lj}\right]S_c(s,\mathbf{X}_i)ds \mid \boldsymbol\beta_0^{\rm T}\mathbf{X}_i = \boldsymbol\beta_0^{\rm T}\mathbf{x}_j\right) + O_p(n^{1/2}h^2)$$
$$= n^{-1/2}\sum_{j=1}^n\int Y_j(s)\lambda\{s, \boldsymbol\beta_0^{\rm T}\mathbf{x}_j, I(w_i \leqslant s), w_iI(w_i \leqslant s)\}\frac{\boldsymbol\lambda_2\{s, \boldsymbol\beta_0^{\rm T}\mathbf{x}_j, I(w_i \leqslant s), w_iI(w_i \leqslant s)\}}{\lambda\{s, \boldsymbol\beta_0^{\rm T}\mathbf{x}_j, I(w_i \leqslant s), w_iI(w_i \leqslant s)\}}$$
$$\otimes\left[\frac{E\left\{\mathbf{X}_{lj}Y_j(s) \mid \boldsymbol\beta_0^{\rm T}\mathbf{x}_j, I(w_i \leqslant s), w_iI(w_i \leqslant s)\right\}}{E\left\{Y_j(s) \mid \boldsymbol\beta_0^{\rm T}\mathbf{x}_j, I(w_i \leqslant s), w_iI(w_i \leqslant s)\right\}} - \mathbf{x}_{lj}\right]ds + O_p(n^{1/2}h^2).$$

When $nh^4 \to 0$, plugging the results of $\mathbf{T}_1$ and $\mathbf{T}_{32}$ to (S35), the expression in (S31) is

$$n^{-1/2}\sum_{i=1}^n\Delta_i\frac{\widehat{\boldsymbol\lambda}_2\{Z_i, \boldsymbol\beta_0^{\rm T}\mathbf{X}_i, I(W_i \leqslant Z_i), W_iI(W_i \leqslant Z_i)\}}{\widehat\lambda\{Z_i, \boldsymbol\beta_0^{\rm T}\mathbf{X}_i, I(W_i \leqslant Z_i), W_iI(W_i \leqslant Z_i)\}}$$
$$\otimes\left[\mathbf{X}_{li} - \frac{\widehat E\left\{\mathbf{X}_{li}Y_i(Z_i) \mid \boldsymbol\beta_0^{\rm T}\mathbf{X}_i, I(W_i \leqslant Z_i), W_iI(W_i \leqslant Z_i)\right\}}{\widehat E\left\{Y_i(Z_i) \mid \boldsymbol\beta_0^{\rm T}\mathbf{X}_i, I(W_i \leqslant Z_i), W_iI(W_i \leqslant Z_i)\right\}}\right]$$
$$= n^{-1/2}\sum_{i=1}^n\int\frac{\boldsymbol\lambda_2\{t, \boldsymbol\beta_0^{\rm T}\mathbf{X}_i, I(W_i \leqslant t), W_iI(W_i \leqslant t)\}}{\lambda\{t, \boldsymbol\beta_0^{\rm T}\mathbf{X}_i, I(W_i \leqslant t), W_iI(W_i \leqslant t)\}}$$
$$\otimes\left[\mathbf{X}_{li} - \frac{E\left\{\mathbf{X}_{li}Y_i(t) \mid \boldsymbol\beta_0^{\rm T}\mathbf{X}_i, I(W_i \leqslant Z_i), W_iI(W_i \leqslant Z_i)\right\}}{E\left\{Y_i(t) \mid \boldsymbol\beta_0^{\rm T}\mathbf{X}_i, I(W_i \leqslant Z_i), W_iI(W_i \leqslant Z_i)\right\}}\right]$$
$$\times dM_i\{t, \boldsymbol\beta_0^{\rm T}\mathbf{X}_i, I(W_i \leqslant t), W_iI(W_i \leqslant t)\} + o_p(1)$$
$$= n^{-1/2}\sum_{i=1}^n\mathbf{S}_{\rm eff}\{\Delta_i, Z_i, \mathbf{X}_i, I(W_i \leqslant t), W_iI(W_i \leqslant Z_i)\} + o_p(1).$$

Finally,

$$\mathbf{T}_4 = n^{-1/2}\sum_{i=1}^n\Delta_i\left[\frac{\widehat{\boldsymbol\lambda}_2\{Z_i, \boldsymbol\beta_0^{\rm T}\mathbf{X}_i, I(W_i \leqslant Z_i), W_iI(W_i \leqslant Z_i)\}}{\widehat\lambda\{Z_i, \boldsymbol\beta_0^{\rm T}\mathbf{X}_i, I(W_i \leqslant Z_i), W_iI(W_i \leqslant Z_i)\}}\right.$$
$$\left.-\frac{\boldsymbol\lambda_2\{Z_i, \boldsymbol\beta_0^{\rm T}\mathbf{X}_i, I(W_i \leqslant Z_i), W_iI(W_i \leqslant Z_i)\}}{\lambda\{Z_i, \boldsymbol\beta_0^{\rm T}\mathbf{X}_i, I(W_i \leqslant Z_i), W_iI(W_i \leqslant Z_i)\}}\right]$$
$$\times\left[\frac{E\left\{\mathbf{X}_{li}Y_i(Z_i) \mid \boldsymbol\beta_0^{\rm T}\mathbf{X}_i, I(W_i \leqslant Z_i), W_iI(W_i \leqslant Z_i)\right\}}{E\left\{Y_i(Z_i) \mid \boldsymbol\beta_0^{\rm T}\mathbf{X}_i, I(W_i \leqslant Z_i), W_iI(W_i \leqslant Z_i)\right\}}\right.$$
$$\left.-\frac{\widehat E\left\{\mathbf{X}_{li}Y_i(Z_i) \mid \boldsymbol\beta_0^{\rm T}\mathbf{X}_i, I(W_i \leqslant Z_i), W_iI(W_i \leqslant Z_i)\right\}}{\widehat E\left\{Y_i(Z_i) \mid \boldsymbol\beta_0^{\rm T}\mathbf{X}_i, I(W_i \leqslant Z_i), W_iI(W_i \leqslant Z_i)\right\}}\right]$$

$$
\begin{aligned}
&= o_p\left(n^{-1/2}\sum_{i=1}^n \Delta_i\left[\frac{E\left\{\mathbf{X}_{li}Y_i(Z_i)\mid \boldsymbol\beta_0^{\mathrm T}\mathbf{X}_i, I(W_i\leqslant Z_i), W_iI(W_i\leqslant Z_i)\right\}}{E\left\{Y_i(Z_i)\mid \boldsymbol\beta_0^{\mathrm T}\mathbf{X}_i, I(W_i\leqslant Z_i), W_iI(W_i\leqslant Z_i)\right\}}\right.\right.\\
&\qquad\left.\left.-\frac{\widehat E\left\{\mathbf{X}_{li}Y_i(Z_i)\mid \boldsymbol\beta_0^{\mathrm T}\mathbf{X}_i, I(W_i\leqslant Z_i), W_iI(W_i\leqslant Z_i)\right\}}{\widehat E\left\{Y_i(Z_i)\mid \boldsymbol\beta_0^{\mathrm T}\mathbf{X}_i, I(W_i\leqslant Z_i), W_iI(W_i\leqslant Z_i)\right\}}\right]\right)\\
&= o_p\left(n^{-1/2}\sum_{i=1}^n\int Y_i(s)\lambda\{s,\boldsymbol\beta_0^{\mathrm T}\mathbf{X}_i, I(W_i\leqslant s), W_iI(W_i\leqslant s)\}\right.\\
&\qquad\left.\times\left[\frac{E\left\{\mathbf{X}_{li}Y_i(s)\mid \boldsymbol\beta_0^{\mathrm T}\mathbf{X}_i, I(W_i\leqslant Z_i), W_iI(W_i\leqslant Z_i)\right\}}{E\left\{Y_i(s)\mid \boldsymbol\beta_0^{\mathrm T}\mathbf{X}_i, I(W_i\leqslant Z_i), W_iI(W_i\leqslant Z_i)\right\}}-\mathbf{x}_{li}\right]ds\right)+o_p(n^{1/2}h^2)\\
&= o_p(1),
\end{aligned}
$$

where the last equality holds because the integrand has mean zero conditional on $\boldsymbol\beta_0^{\mathrm T}\mathbf{X}$, and the second to last equality follows through the same derivation as that of $\mathbf{T}_3$. Utilizing these results in (S31), in conjunction with (S32), establishes the theorem. □

4.5. *Proof of Theorem 3.* The asymptotic variances of $\widehat m_N(t,\boldsymbol\beta^{\mathrm T}\mathbf{X})$ and $\widehat m_T(t,\boldsymbol\beta^{\mathrm T}\mathbf{X},W)$ are

$$
\begin{aligned}
\sigma_N^2(t,\boldsymbol\beta^{\mathrm T}\mathbf{X}) &= e^{2\Lambda_N(t,\boldsymbol\beta^{\mathrm T}\mathbf{X})}\frac{\int K^2(u)du}{f_{\boldsymbol\beta^{\mathrm T}\mathbf{X}}(\boldsymbol\beta^{\mathrm T}\mathbf{X})}\int_0^\tau\frac{\lambda_N(r,\boldsymbol\beta^{\mathrm T}\mathbf{X})}{E\{I(Z\geqslant r)\mid\boldsymbol\beta^{\mathrm T}\mathbf{X}\}}\\
&\quad\times\left\{I(r<t)\int_t^\tau e^{-\Lambda_N(s,\boldsymbol\beta^{\mathrm T}\mathbf{X})}ds+\int_{\max(r,t)}^\tau e^{-\Lambda_N(s,\boldsymbol\beta^{\mathrm T}\mathbf{X})}ds\right\}^2dr,\\
\sigma_T^2(t,\boldsymbol\beta^{\mathrm T}\mathbf{X},W) &= e^{2\Lambda_T(t,\boldsymbol\beta^{\mathrm T}\mathbf{X},W)}\frac{\int K^2(u)du}{f_{\boldsymbol\beta^{\mathrm T}\mathbf{X},W}(\boldsymbol\beta^{\mathrm T}\mathbf{X},W)}\int_0^\tau\frac{\lambda_T(r,\boldsymbol\beta^{\mathrm T}\mathbf{X},W)}{E\{I(Z\geqslant r)\mid\boldsymbol\beta^{\mathrm T}\mathbf{X},W\}}\\
&\quad\times\left\{I(r<t)\int_t^\tau e^{-\Lambda_T(s,\boldsymbol\beta^{\mathrm T}\mathbf{X},W)}ds+\int_{\max(r,t)}^\tau e^{-\Lambda_T(s,\boldsymbol\beta^{\mathrm T}\mathbf{X},W)}ds\right\}^2dr,
\end{aligned}
$$

respectively.

We begin by analyzing $\widehat m_N(t,\boldsymbol\beta^{\mathrm T}\mathbf{X})$. Our focus is on examining $\sqrt{nh}\,\{\widehat m_N(t,\boldsymbol\beta^{\mathrm T}\mathbf{X})-m_N(t,\boldsymbol\beta^{\mathrm T}\mathbf{X})\}$, because $\sqrt{nh}\left\{\widehat m_N(t,\widehat{\boldsymbol\beta}^{\mathrm T}\mathbf{X})-\widehat m_N(t,\boldsymbol\beta^{\mathrm T}\mathbf{X})\right\}=O_p(\sqrt{h})=o_p(1)$ based on Theorems 1 and 2.

We expand $\sqrt{nh}\,\{\widehat m_N(t,\boldsymbol\beta^{\mathrm T}\mathbf{X})-m_N(t,\boldsymbol\beta^{\mathrm T}\mathbf{X})\}$ as

$$
\begin{aligned}
&\sqrt{nh}\,\{\widehat m_N(t,\boldsymbol\beta^{\mathrm T}\mathbf{X})-m_N(t,\boldsymbol\beta^{\mathrm T}\mathbf{X})\}\\
&= \sqrt{nh}\left\{e^{\widehat\Lambda_N(t,\boldsymbol\beta^{\mathrm T}\mathbf{X})}-e^{\Lambda_N(t,\boldsymbol\beta^{\mathrm T}\mathbf{X})}\right\}\int_t^\tau e^{-\Lambda_N(s,\boldsymbol\beta^{\mathrm T}\mathbf{X})}ds \qquad \text{(S36)}\\
&\quad+\sqrt{nh}e^{\Lambda_N(t,\boldsymbol\beta^{\mathrm T}\mathbf{X})}\int_t^\tau\left\{e^{-\widehat\Lambda_N(s,\boldsymbol\beta^{\mathrm T}\mathbf{X})}-e^{-\Lambda_N(s,\boldsymbol\beta^{\mathrm T}\mathbf{X})}\right\}ds. \qquad \text{(S37)}\\
&\quad+\sqrt{nh}\left\{e^{\widehat\Lambda_N(t,\boldsymbol\beta^{\mathrm T}\mathbf{X})}-e^{\Lambda_N(t,\boldsymbol\beta^{\mathrm T}\mathbf{X})}\right\}\int_t^\tau\left\{e^{-\widehat\Lambda_N(s,\boldsymbol\beta^{\mathrm T}\mathbf{X})}-e^{-\Lambda_N(s,\boldsymbol\beta^{\mathrm T}\mathbf{X})}\right\}ds. \qquad \text{(S38)}
\end{aligned}
$$

It is easy to see that the term in (S38) satisfies

$$
\sqrt{nh}\left\{e^{\widehat\Lambda_N(t,\boldsymbol\beta^{\mathrm T}\mathbf{X})}-e^{\Lambda_N(t,\boldsymbol\beta^{\mathrm T}\mathbf{X})}\right\}\int_t^\tau\left\{e^{-\widehat\Lambda_N(s,\boldsymbol\beta^{\mathrm T}\mathbf{X})}-e^{-\Lambda_N(s,\boldsymbol\beta^{\mathrm T}\mathbf{X})}\right\}ds
$$

$$= \sqrt{nh} O_p\{\widehat{\Lambda}_N(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) - \Lambda_N(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})\} \int_t^\tau e^{-\Lambda_N(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})} O_p\{\widehat{\Lambda}_N(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) - \Lambda_N(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})\} ds$$
$$= O_p(\sqrt{nh}) O_p\{h^4 + (nh)^{-1}\}$$
$$= o_p(1),$$

by Condition C2.

We inspect the terms in (S36) and (S37). For (S36), based on Lemma 1,

$$\sqrt{nh}\left\{e^{\widehat{\Lambda}_N(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})} - e^{\Lambda_N(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})}\right\} \int_t^\tau e^{-\Lambda_N(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})} ds$$
$$= \sqrt{nh} e^{\Lambda_N(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})} \left(\widehat{\Lambda}_N(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) - \Lambda_N(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) + O_p[\{\widehat{\Lambda}_N(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) - \Lambda_N(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})\}^2]\right)$$
$$\times \int_t^\tau e^{-\Lambda_N(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})} ds$$
$$= \sqrt{nh} e^{\Lambda_N(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})} \left\{\widehat{\Lambda}_N(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) - \Lambda_N(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})\right\} \int_t^\tau e^{-\Lambda_N(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})} ds + o_p(1),$$

where the last step uses Condition C2.

For (S37), using Condition C2 as well, we get

$$\sqrt{nh} e^{\Lambda_N(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})} \int_t^\tau \left\{e^{-\widehat{\Lambda}_N(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})} - e^{-\Lambda_N(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})}\right\} ds$$
$$= \sqrt{nh} e^{\Lambda_N(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})} \int_t^\tau \left[\widehat{\Lambda}_N(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) - \Lambda_N(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) + O_p\{(nh)^{-1} + h^4\}\right] e^{-\Lambda_N(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})} ds$$
$$= \sqrt{nh} e^{\Lambda_N(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})} \int_t^\tau \left\{\widehat{\Lambda}_N(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) - \Lambda_N(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})\right\} e^{-\Lambda_N(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})} ds + o_p(1).$$

Now combine the leading terms in (S36) and (S37) and use the expansion of $\widehat{\Lambda}_N(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) - \Lambda_N(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})$ in Lemma 2,

$$\sqrt{nh} e^{\Lambda_N(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})} \left\{\widehat{\Lambda}_N(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) - \Lambda_N(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})\right\} \int_t^\tau e^{-\Lambda_N(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})} ds$$
$$+ \sqrt{nh} e^{\Lambda_N(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})} \int_t^\tau \left\{\widehat{\Lambda}_N(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) - \Lambda_N(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})\right\} e^{-\Lambda_N(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})} ds$$
$$= e^{\Lambda_N(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})} \sum_{i=1}^n \int_0^\tau \sqrt{\frac{h}{n}} \frac{I\left\{\phi_n(r,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) > 0\right\} K_h(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})}{f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) E\{I(Z \geqslant r) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\}}$$
$$\text{(S39)} \times \left\{I(r < t) \int_t^\tau e^{-\Lambda_N(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})} ds + \int_{\max(r,t)}^\tau e^{-\Lambda_N(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})} ds\right\} dM_i(r,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) + o_p(1).$$

Note that $I\left\{\phi_n(r,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) > 0\right\} = 1$ almost surely and according to Lemma 1

$$\frac{1}{n}\sum_{i=1}^n h K_h^2(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) Y_i(r) = f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) E\{I(Z \geqslant r) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\} \int K^2(u) du + o_p(1).$$

The leading term in (S39) converges to $N\{0, \sigma_N^2(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})\}$ uniformly by the martingale central limit theorem, where

$$\sigma_N^2(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) = e^{2\Lambda_N(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})} \frac{\int K^2(u) du}{f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})} \int_0^\tau \frac{\lambda_N(r,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})}{E\{I(Z \geqslant r) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\}}$$

$$\times \left\{ I(r<t) \int_t^\tau e^{-\Lambda_N(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})} ds + \int_{\max(r,t)}^\tau e^{-\Lambda_N(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})} ds \right\}^2 dr.$$

Therefore $\sqrt{nh}\,\{\widehat{m}_N(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) - m_N(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})\} \to N\{0,\sigma_N^2(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})\}$ uniformly for all t and $\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}$.

Similarly, $\sqrt{nh}\,\{\widehat{m}_T(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},W) - m_T(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},W)\} \to N\{0,\sigma_T^2(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},W)\}$ uniformly for all t and $\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}$ where

$$\begin{aligned}
&\sigma_T^2(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},W)\\
&= e^{2\Lambda_T(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},W)} \frac{\int K^2(u)du}{f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},W}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},W)} \int_0^\tau \frac{\lambda_T(r,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},W)}{E\{I(Z\geqslant r)\mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},W\}}\\
&\times \left\{ I(r<t) \int_t^\tau e^{-\Lambda_T(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},W)} ds + \int_{\max(r,t)}^\tau e^{-\Lambda_T(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},W)} ds \right\}^2 dr.
\end{aligned}$$

Furthermore, $\widehat{m}\{t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},I(W\leqslant t),WI(W\leqslant t)\} = \widehat{m}_T(t-W,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})I(W\leqslant t) + \widehat{m}_N(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})\{1-I(W\leqslant t)\}$ and

$$\begin{aligned}
&\sqrt{nh}\left\{\widehat{m}\{t,\widehat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{X},I(W\leqslant t),WI(W\leqslant t)\} - m\{t,\widehat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{X},I(W\leqslant t),WI(W\leqslant t)\}\right\}\\
&\to N[0,\sigma_m^2\{t,\widehat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{X},I(W\leqslant t),WI(W\leqslant t)\}]
\end{aligned}$$

uniformly for all t and $\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}$ where $\sigma_m^2\{t,\widehat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{X},I(W\leqslant t),WI(W\leqslant t)\} = \sigma_T^2(t-W,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})I(W\leqslant t) + \sigma_N^2(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})\{1-I(W\leqslant t)\}$.

□

5. Relaxation of the Complete Follow-up Assumption. To weaken the complete follow-up assumption that the event time is supported on $[0,\tau]$, we relax the compact support assumption on the event time, while allow a sample size dependent end of the study time τ_n. Because the estimation and inference of $\boldsymbol{\beta},\Lambda\{t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x},I(W\leqslant t),WI(W\leqslant t)\}$ does not rely on the complete follow-up assumption, hence under the weakened assumptions, the same analysis as in the main text leads to the same results for $\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}$ as in Theorems 1 and 2.

We assume

$$\text{(S40)}\quad \int_{\tau_n}^\infty e^{-\Lambda\{s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x},I(w\leqslant s),wI(w\leqslant s)\}} ds = o(n^{\frac{1-2\nu}{4\nu}}) e^{-\Lambda\{\tau_n-\tau_0,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x},I(w\leqslant \tau_n-\tau_0),wI(w\leqslant\tau_n-\tau_0)\}},$$

where τ_0 is a small positive constant. (S40) is equivalent to

$$\lim_{n\to\infty} \frac{\tau_n' e^{\Lambda\{\tau_n-\tau_0,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x},I(w\leqslant\tau_n-\tau_0),wI(w\leqslant\tau_n-\tau_0)\}-\Lambda\{\tau_n,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x},I(w\leqslant\tau_n),wI(w\leqslant\tau_n)\}}}{\frac{2\nu-1}{4\nu}n^{\frac{1-6\nu}{4\nu}} + n^{\frac{1-2\nu}{4\nu}}\lambda\{\tau_n-\tau_0,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x},I(w\leqslant\tau_n-\tau_0),wI(w\leqslant\tau_n-\tau_0)\}\tau_n'} = 0,$$

where $\tau_n' \equiv d\tau_n/dn$. Clearly, a sufficient condition for (S40) to hold is that the cumulative hazard function $\Lambda(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}) \geqslant C(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x})t^{1+\alpha}$ for some constant $\alpha>0$, and $\tau_n = n^\omega$ for some constant $\omega>0$. We point out two facts about this condition. First, ω can be very small as long as it is fixed, for example, we can set $\omega=1/2$. This implies that the end of the study time τ_n, although required to increase with sample size n, can increase very slowly. Second, the tail condition on the distribution family of T given $\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}$ is very weak. For example, all sub-Gaussian distributions satisfy this requirement, this naturally includes all the Cox proportional hazard model with the baseline hazard t^c for $c>0$ and the Weibull family with increasing risk, which are often used as baseline to generate various survival models in practice.

Let

$$m^*\{t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x},I(W\leqslant t),WI(W\leqslant t)\}$$
$$=e^{\Lambda\{t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x},I(W\leqslant t),WI(W\leqslant t)\}}\int_t^{\tau_n}e^{-\Lambda\{s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x},I(W\leqslant s),WI(W\leqslant s)\}}ds.$$

Then, under Condition C2,

$$|m\{t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x},I(W\leqslant t),WI(W\leqslant t)\}-m^*\{t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x},I(W\leqslant t),WI(W\leqslant t)\}|$$
$$=o(n^{\frac{1-2\nu}{4\nu}})=o\{(nh)^{-1/2}\}. \tag{S41}$$

Note that our estimator described in the main text, when viewed as an estimator for $m^*\{t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x},I(W\leqslant t),WI(W\leqslant t)\}$, satisfies the same properties as described in Theorem 3, i.e. $\sqrt{nh}[\widehat{m}\{t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x},I(W\leqslant t),WI(W\leqslant t)\}-m^*\{t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x},I(W\leqslant t),WI(W\leqslant t)\}]/\sigma_{m^*}\{t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x},I(W\leqslant t),WI(W\leqslant t)\}\to N(0,1)$ in distribution when $n\to\infty$, where

$$\sigma^2_{m^*}\{t,\widehat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{x},I(W\leqslant t),WI(W\leqslant t)\}$$
$$=I(w\leqslant t)e^{2\Lambda_T(t-w,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x},w)}\frac{\int K^2(u)du}{f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},W}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x},w)}\int_0^{\tau_n}\frac{\lambda_T(r-w,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x},w)}{E\{I(Z\geqslant r)\mid\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x},w\}}$$
$$\times\left\{I(r<t-w)\int_{t-w}^{\tau_n}e^{-\Lambda_T(s-w,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x},w)}ds+\int_{\max(r,t-w)}^{\tau_n}e^{-\Lambda_T(s-w,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x},w)}ds\right\}dr$$
$$+\{1-I(w\leqslant t)\}e^{2\Lambda_N(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x})}\frac{\int K^2(u)du}{f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x})}\int_0^{\tau_n}\frac{\lambda_N(r,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x})}{E\{I(Z\geqslant r)\mid\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}\}}$$
$$\times\left\{I(r<t)\int_t^{\tau_n}e^{-\Lambda_N(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x})}ds+\int_{\max(r,t)}^{\tau_n}e^{-\Lambda_N(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x})}ds\right\}dr.$$

Indeed, the proof of Theorem 3 will still follow by substituting τ with τ_n, in combination with (S41) and the fact that $\sigma_{m^*}(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x})$ converges to

$$\sigma^2_{m}\{t,\widehat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{x},I(W\leqslant t),WI(W\leqslant t)\}$$
$$=I(w\leqslant t)e^{2\Lambda_T(t-w,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x},w)}\frac{\int K^2(u)du}{f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},W}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x},w)}\int_0^{\infty}\frac{\lambda_T(r-w,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x},w)}{E\{I(Z\geqslant r)\mid\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x},w\}}$$
$$\times\left\{I(r<t-w)\int_{t-w}^{\infty}e^{-\Lambda_T(s-w,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x},w)}ds+\int_{\max(r,t-w)}^{\infty}e^{-\Lambda_T(s-w,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x},w)}ds\right\}dr$$
$$+\{1-I(w\leqslant t)\}e^{2\Lambda_N(t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x})}\frac{\int K^2(u)du}{f_{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x})}\int_0^{\infty}\frac{\lambda_N(r,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x})}{E\{I(Z\geqslant r)\mid\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}\}}$$
$$\times\left\{I(r<t)\int_t^{\infty}e^{-\Lambda_N(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x})}ds+\int_{\max(r,t)}^{\infty}e^{-\Lambda_N(s,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x})}ds\right\}dr.$$

This further leads to

$$\sqrt{nh}[\widehat{m}\{t,\widehat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{x},I(W\leqslant t),WI(W\leqslant t)\}-m\{t,\widehat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{x},I(W\leqslant t),WI(W\leqslant t)\}]$$
$$\to N[0,\sigma^2_m\{t,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x},I(W\leqslant t),WI(W\leqslant t)\}]$$

in distribution when $n\to\infty$.

6. Code. The MATLAB code of estimating the mean residual life function $\mathbf{m}$ and the parameter $\boldsymbol{\beta}$ is given in github.com/Ge-zichuan/Mean-residual-life.

REFERENCES

FLEMING, T. R. and HARRINGTON, D. P. (1991). *Counting Processes and Survival Analysis*. Wiley, New York.

HÄRDLE, W., SPOKOINY, V., SPERLICH, S. et al. (1997). Semiparametric Single Index Versus Fixed Link Function Modelling. *The Annals of Statistics* **25** 212–243.

MA, Y. and ZHU, L. (2013). Efficient estimation in sufficient dimension reduction. *Annals of Statistics* **41** 250–268.

SILVERMAN, B. W. (1978). Weak and Strong Uniform Consistency of the Kernel Estimate of a Density and its Derivatives. *The Annals of Statistics* 177–184.

SILVERMAN, B. W. (1986). *Density Estimation for Statistics and Data Analysis* **26**. CRC Press, Boca Raton, Florida.

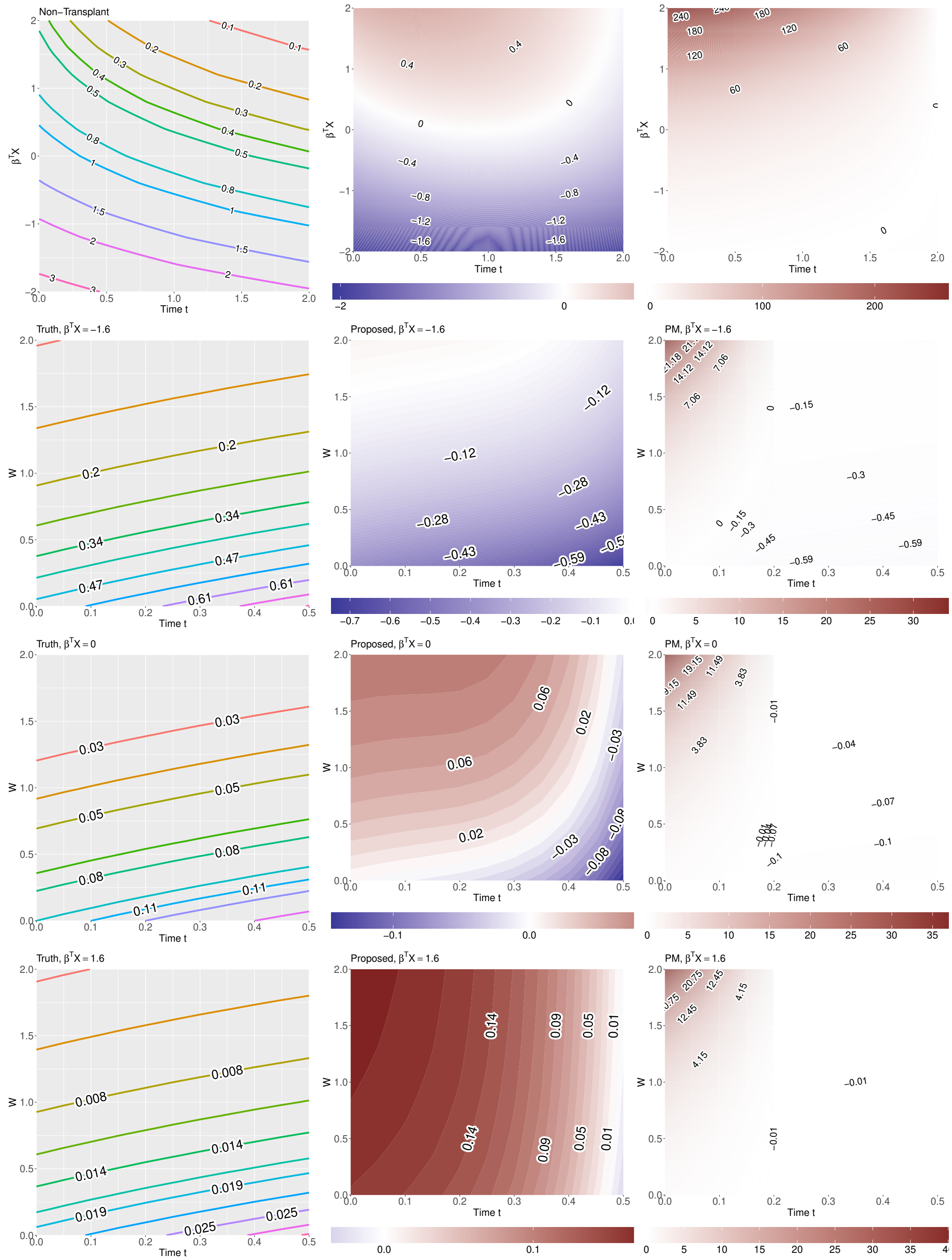

Fig S1: Performance of the semiparametric method on estimating the mean residual life function of **Study 1** without censoring. First row: $m_N(t, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})$ and its estimates. Second row: $m_T(t, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, W)$ and its estimates at $\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X} = -1.6$; Third row: $m_T(t, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, W)$ and its estimates at $\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X} = 0$; Fourth row: $m_T(t, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, W)$ and its estimates at $\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X} = 1.6$. First column: contour plot of the true $m_N(t, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})$ and $\widehat{m}_N(t, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})$. Second column: contour plot of the residuals of the proposed method; Third column: contour plot of the residuals of the PM method;

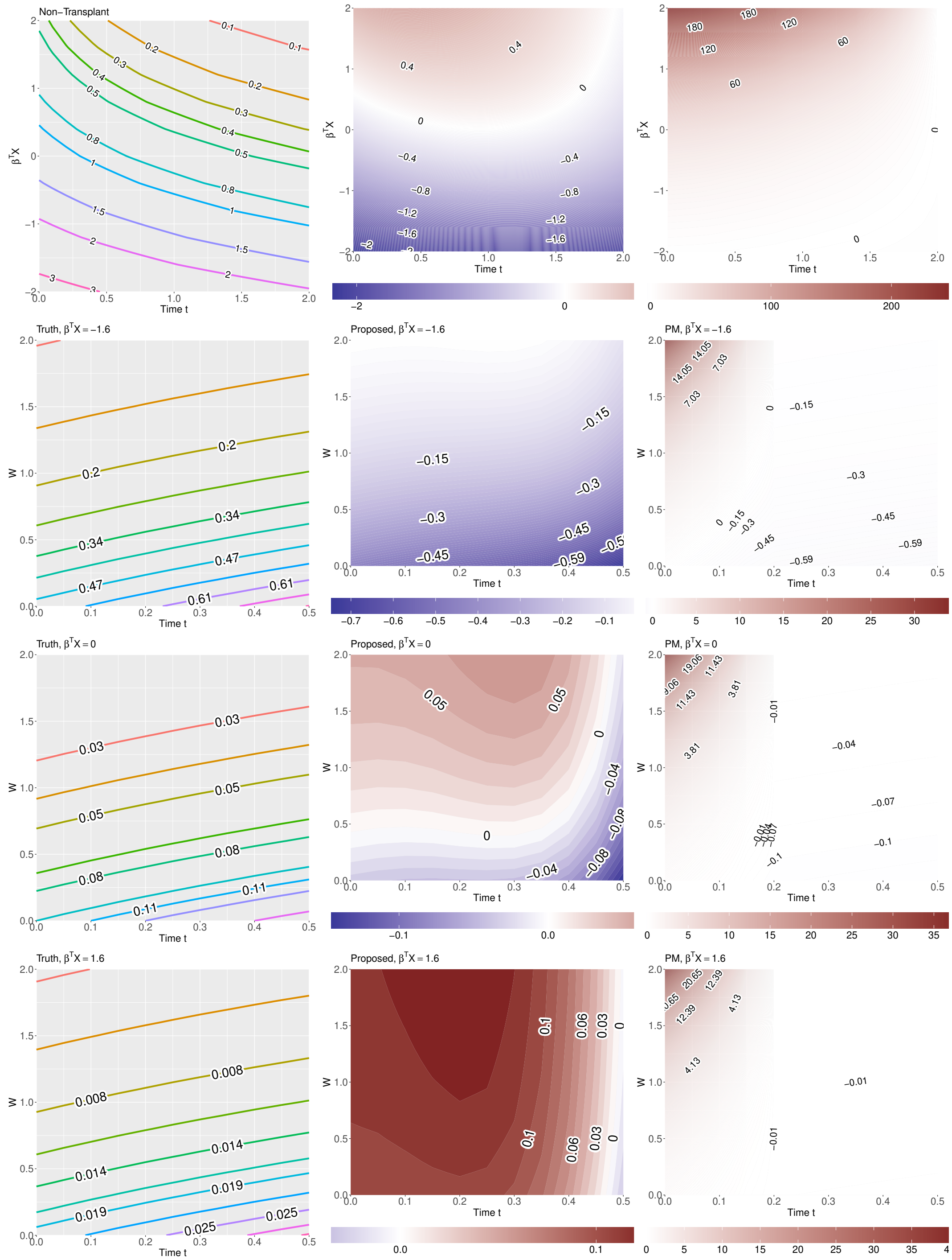

Fig S2: Performance of the semiparametric method on estimating the mean residual life function of **Study 1** under the censoring rate of 20%. First row: $m_N(t, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})$ and its estimates. Second row: $m_T(t, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, W)$ and its estimates at $\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X} = -1.6$; Third row: $m_T(t, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, W)$ and its estimates at $\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X} = 0$; Fourth row: $m_T(t, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, W)$ and its estimates at $\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X} = 1.6$. First column: contour plot of the true $m_N(t, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})$ and $\widehat{m}_N(t, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})$. Second column: contour plot of the residuals of the proposed method; Third column: contour plot of the residuals of the PM method;

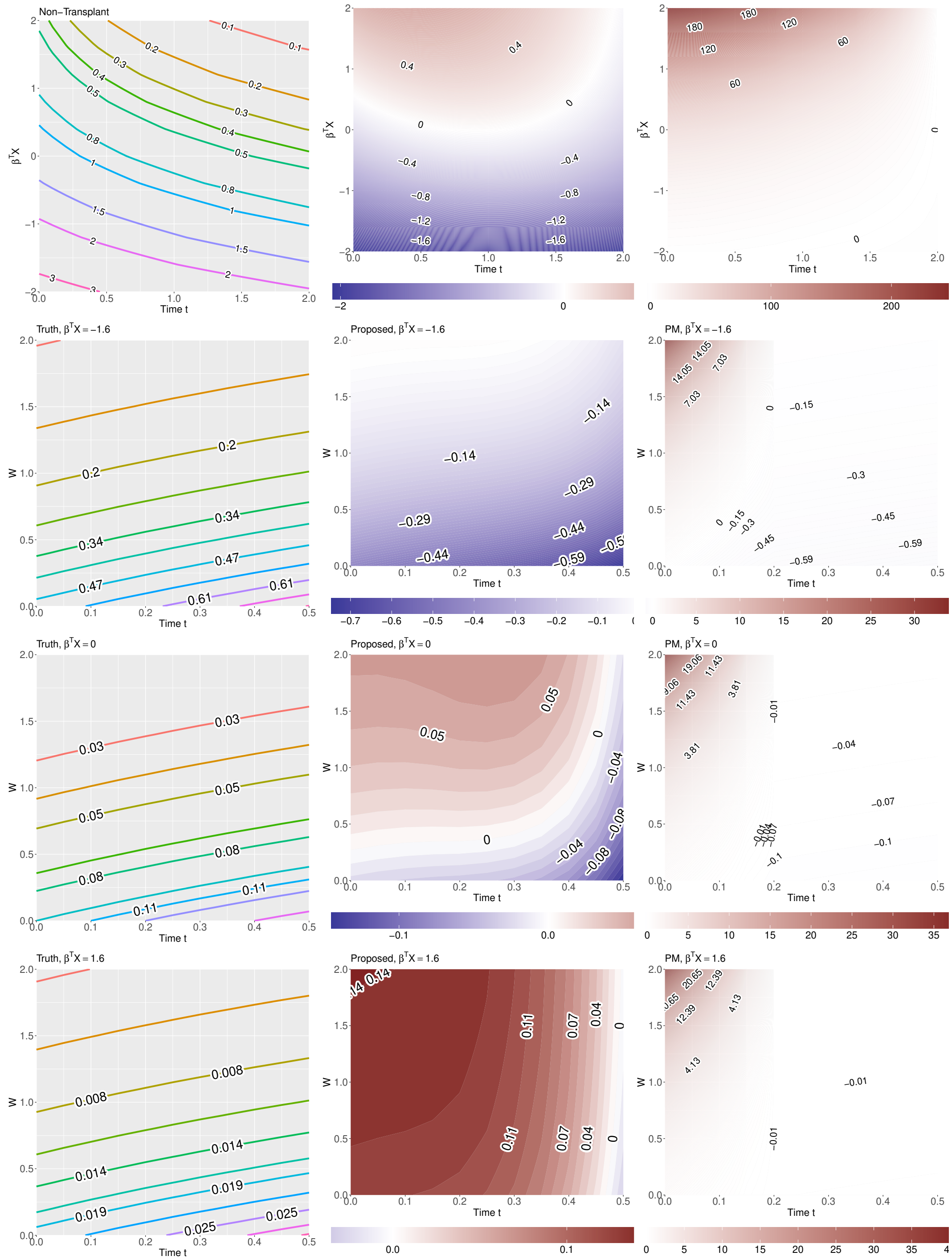

Fig S3: Performance of the semiparametric method on estimating the mean residual life function of **Study 1** under the censoring rate of 40%. First row: $m_N(t, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})$ and its estimates. Second row: $m_T(t, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, W)$ and its estimates at $\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X} = -1.6$; Third row: $m_T(t, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, W)$ and its estimates at $\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X} = 0$; Fourth row: $m_T(t, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, W)$ and its estimates at $\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X} = 1.6$. First column: contour plot of the true $m_N(t, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})$ and $\widehat{m}_N(t, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})$. Second column: contour plot of the residuals of the proposed method; Third column: contour plot of the residuals of the PM method;

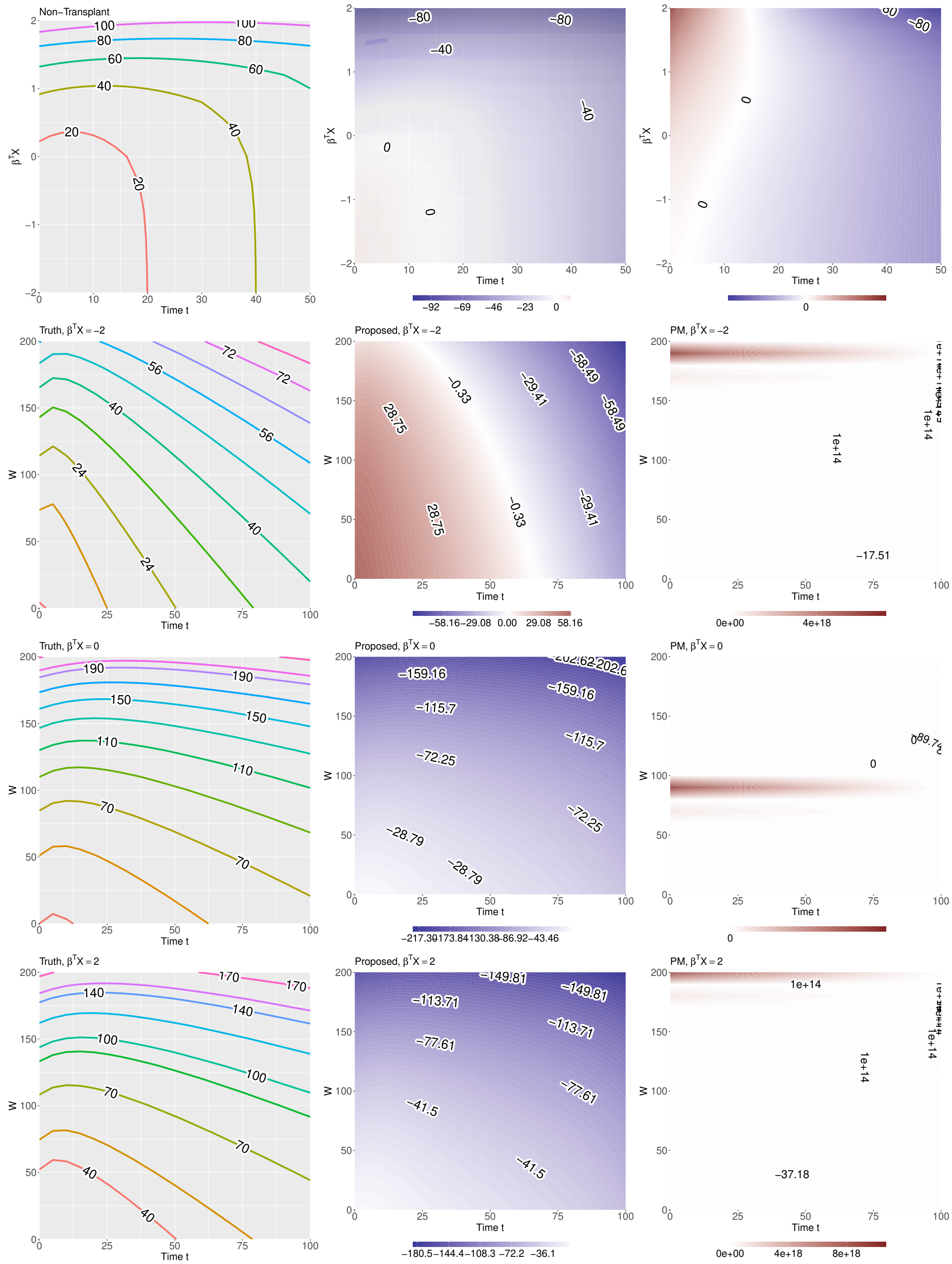

Fig S4: Performance of the semiparametric method on estimating the mean residual life function of **Study 2** with no censoring. First row: $m_N(t, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})$ and its estimates. Second row: $m_T(t, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, W)$ and its estimates at $\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X} = -2$; Third row: $m_T(t, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, W)$ and its estimates at $\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X} = 0$; Fourth row: $m_T(t, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, W)$ and its estimates at $\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X} = 2$. First column: contour plot of the true $m_N(t, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})$ and $\widehat{m}_N(t, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})$. Second column: contour plot of the residuals of the proposed method; Third column: contour plot of the residuals of the PM method;

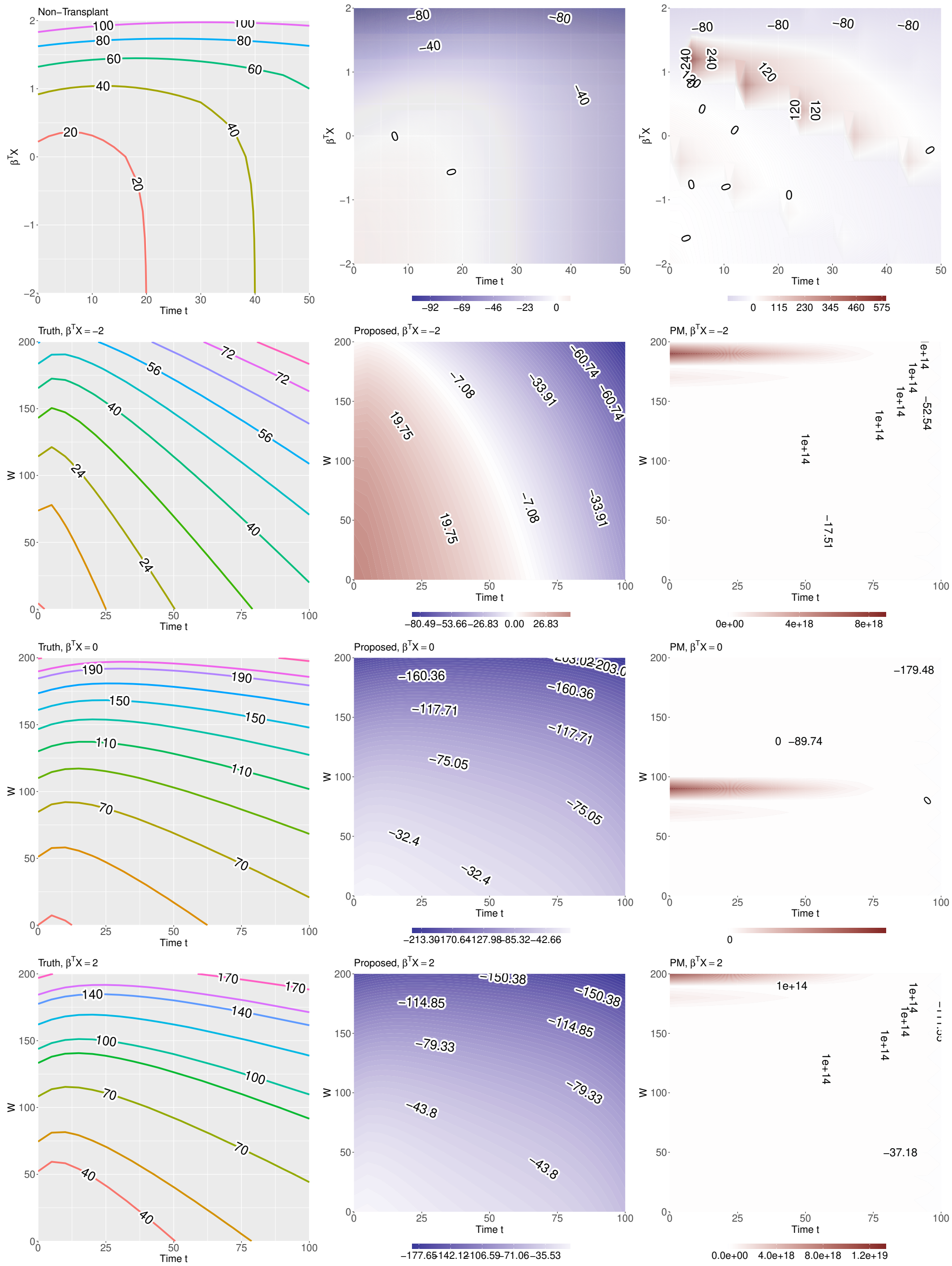

Fig S5: Performance of the semiparametric method on estimating the mean residual life function of **Study 2** under the censoring rate of 20%. First row: $m_N(t, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})$ and its estimates. Second row: $m_T(t, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, W)$ and its estimates at $\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X} = -2$; Third row: $m_T(t, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, W)$ and its estimates at $\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X} = 0$; Fourth row: $m_T(t, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, W)$ and its estimates at $\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X} = 2$. First column: contour plot of the true $m_N(t, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})$ and $\widehat{m}_N(t, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})$. Second column: contour plot of the residuals of the proposed method; Third column: contour plot of the residuals of the PM method;

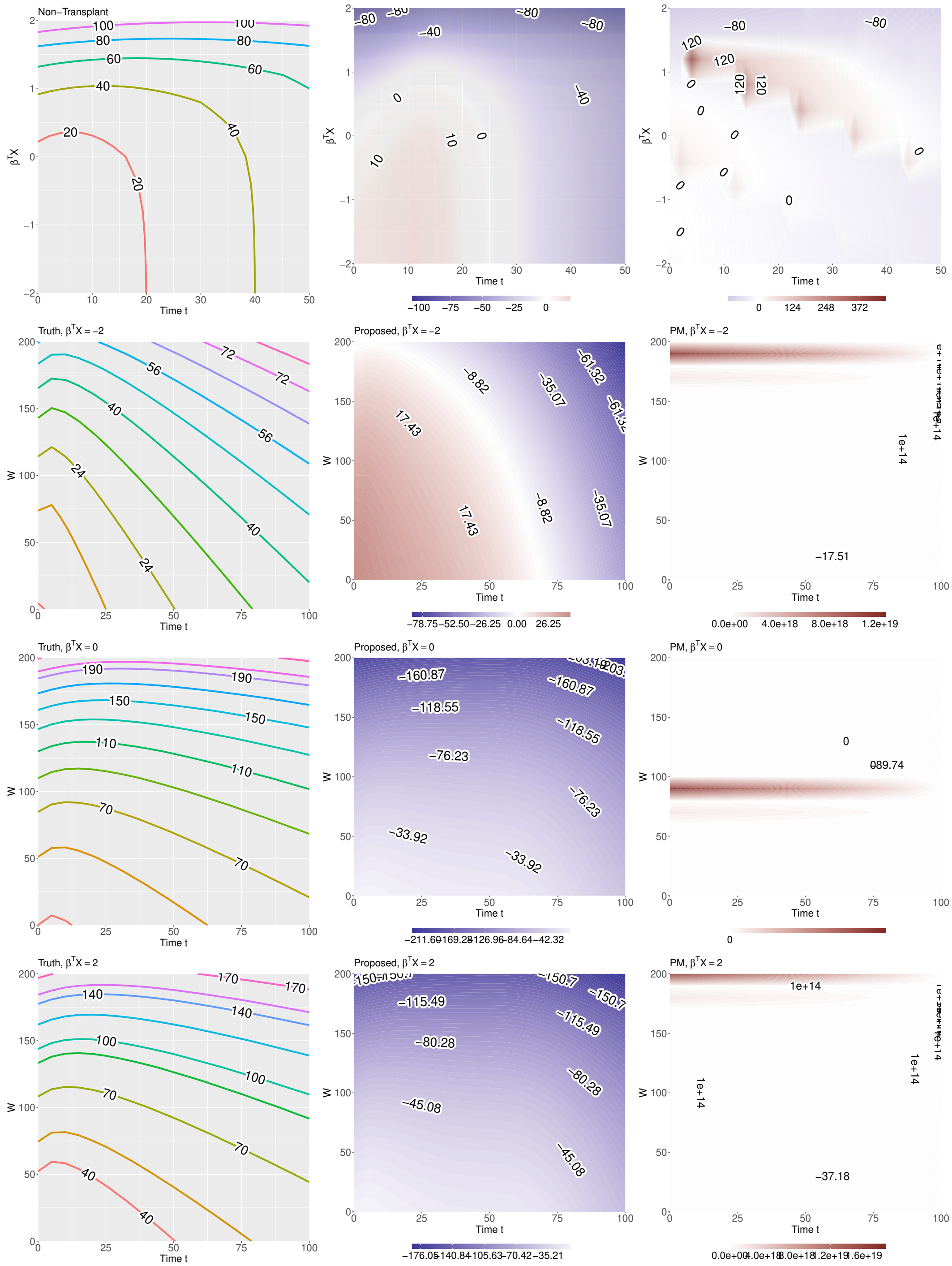

Fig S6: Performance of the semiparametric method on estimating the mean residual life function of **Study 2** under the censoring rate of 40%. First row: $m_N(t, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})$ and its estimates. Second row: $m_T(t, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, W)$ and its estimates at $\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X} = -2$; Third row: $m_T(t, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, W)$ and its estimates at $\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X} = 0$; Fourth row: $m_T(t, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, W)$ and its estimates at $\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X} = 2$. First column: contour plot of the true $m_N(t, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})$ and $\widehat{m}_N(t, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})$. Second column: contour plot of the residuals of the proposed method; Third column: contour plot of the residuals of the PM method;

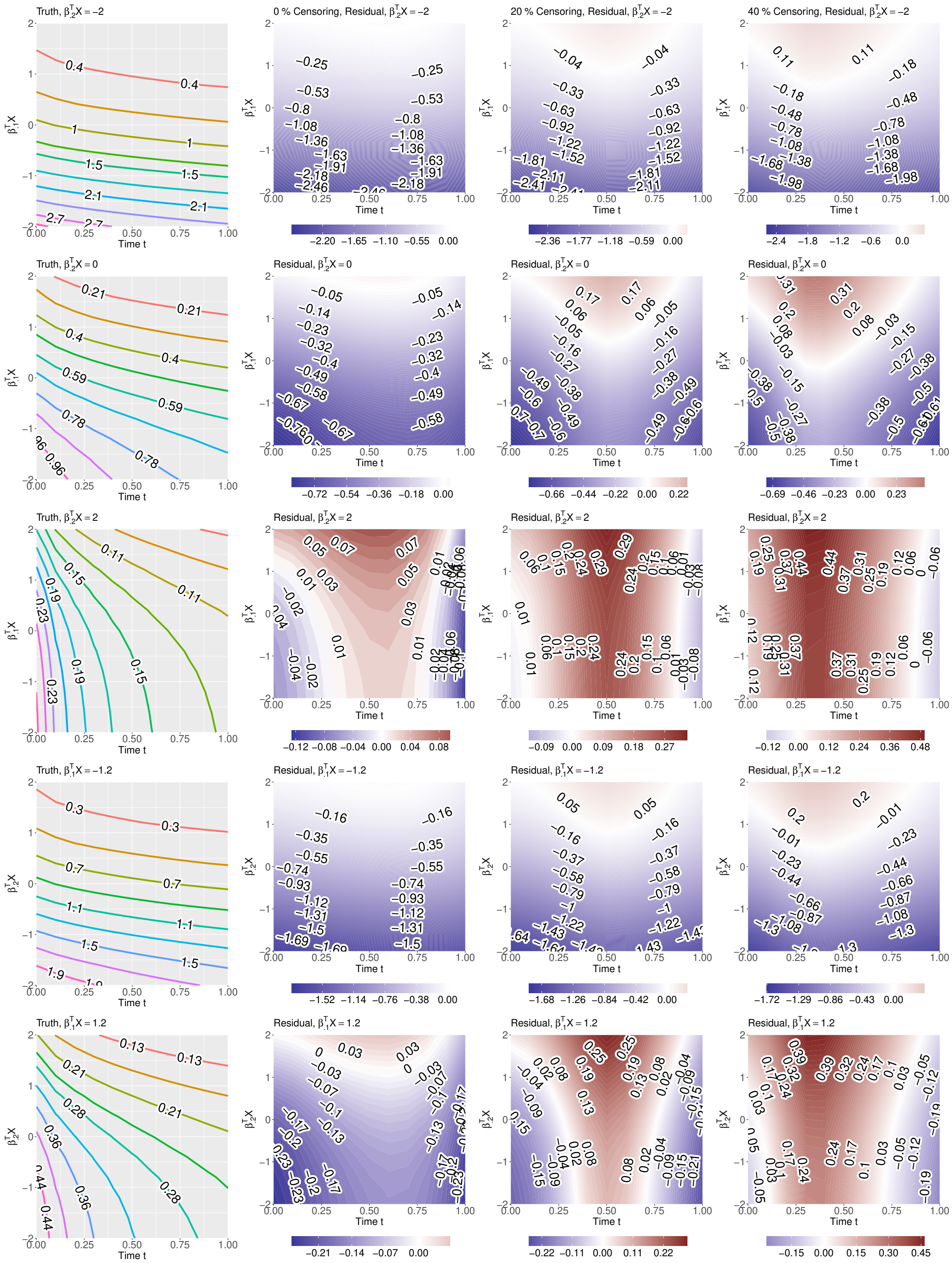

Fig S7: Performance of the proposed method on estimating the mean residual life function of the nontransplant group of **Study 3** at different $\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}$. Left column: contour plots of true $m_N(t, 0, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})$; Column 2 to 4: contour plots of the residual of $\widehat{m}_N(t, \boldsymbol{\beta}_2{}^{\mathrm{T}}\mathbf{X}, 0)$ under the censoring rates of 0%, 20%, and 40%.

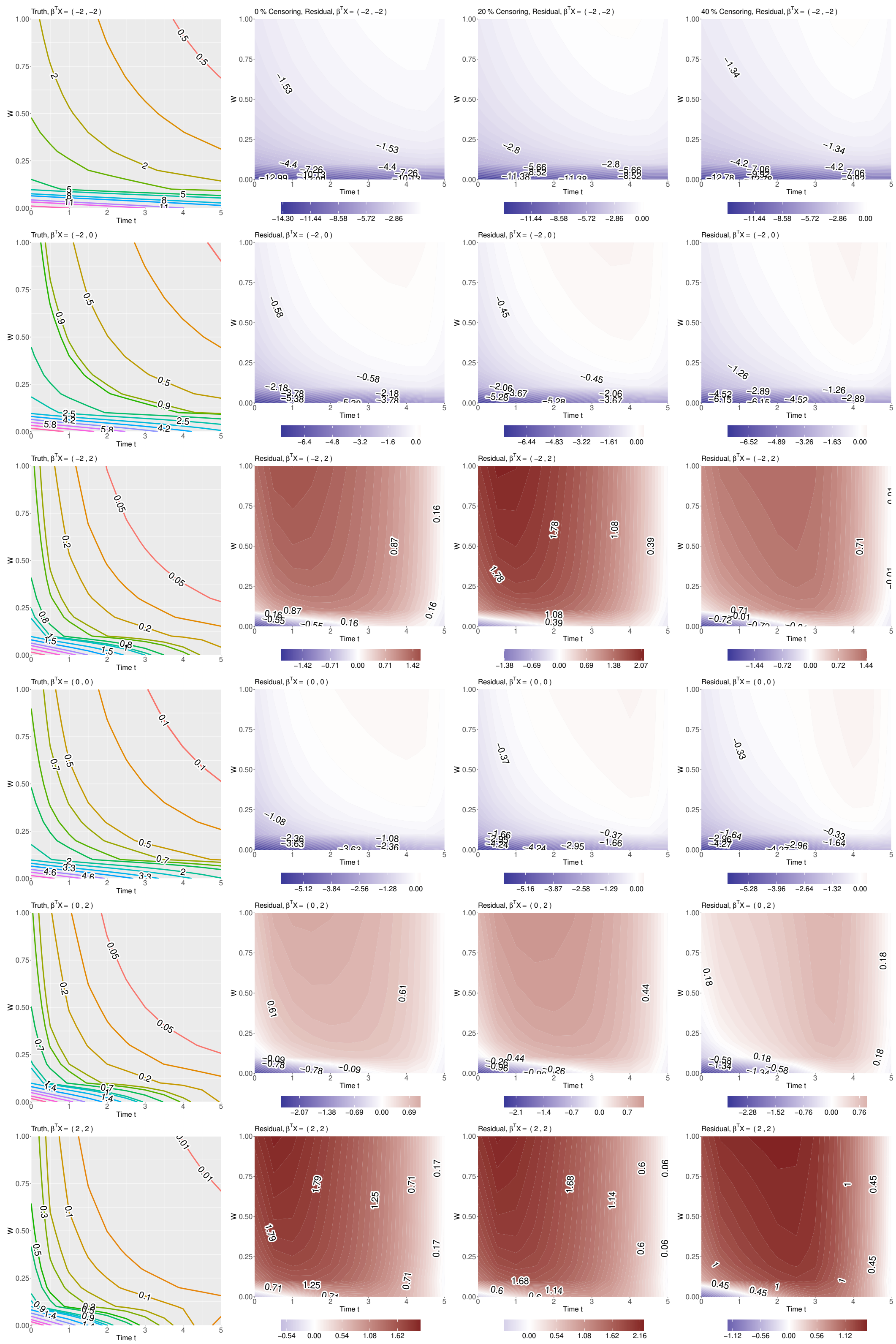

Fig S8: Performance of the proposed method on estimating the mean residual life function of the transplant group of **Study 3** at different $\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}$. Left column: contour plots of true $m_T(t, 0, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X})$; Column 2 to 4: contour plots of the residuals of $\widehat{m}_T(t, \boldsymbol{\beta}_2{}^{\mathrm{T}}\mathbf{X}, 0)$ under the censoring rates of 0%, 20%, and 40%.

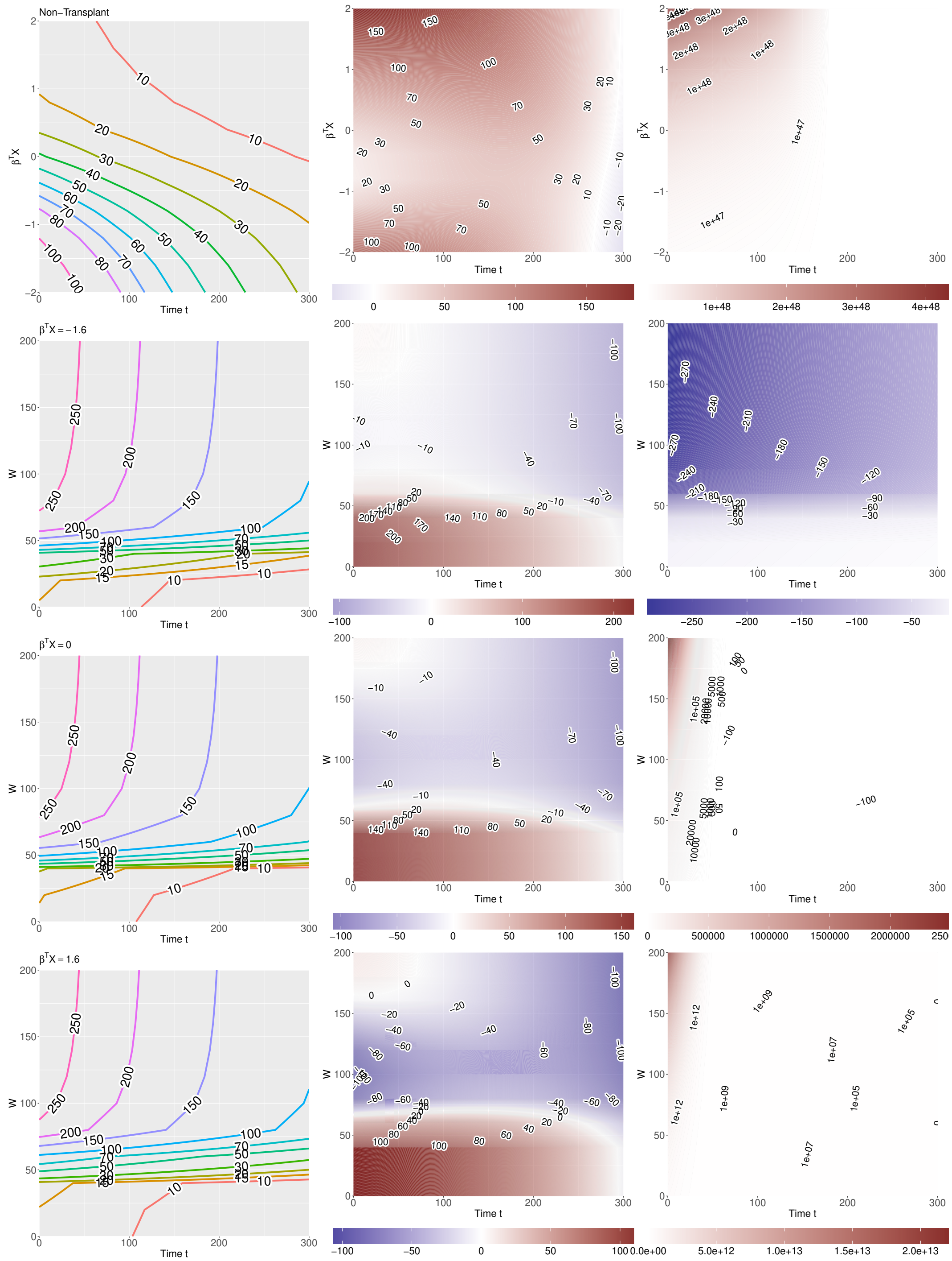

Fig S9: Performance of the mean residual life function estimation in **Study 4**. Row 1: Performance of $m_N(t, \boldsymbol{\beta}_1{}^{\mathrm{T}}\mathbf{X})$; Row 2 to 4: $\widehat{m}_T(t, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}, W)$ at $\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X} = -1.6, 0, 1.6$. Left column: Contour plot of true mean residual function. Middle column: Residual from the proposed method. Right column: Residual from the PM method.

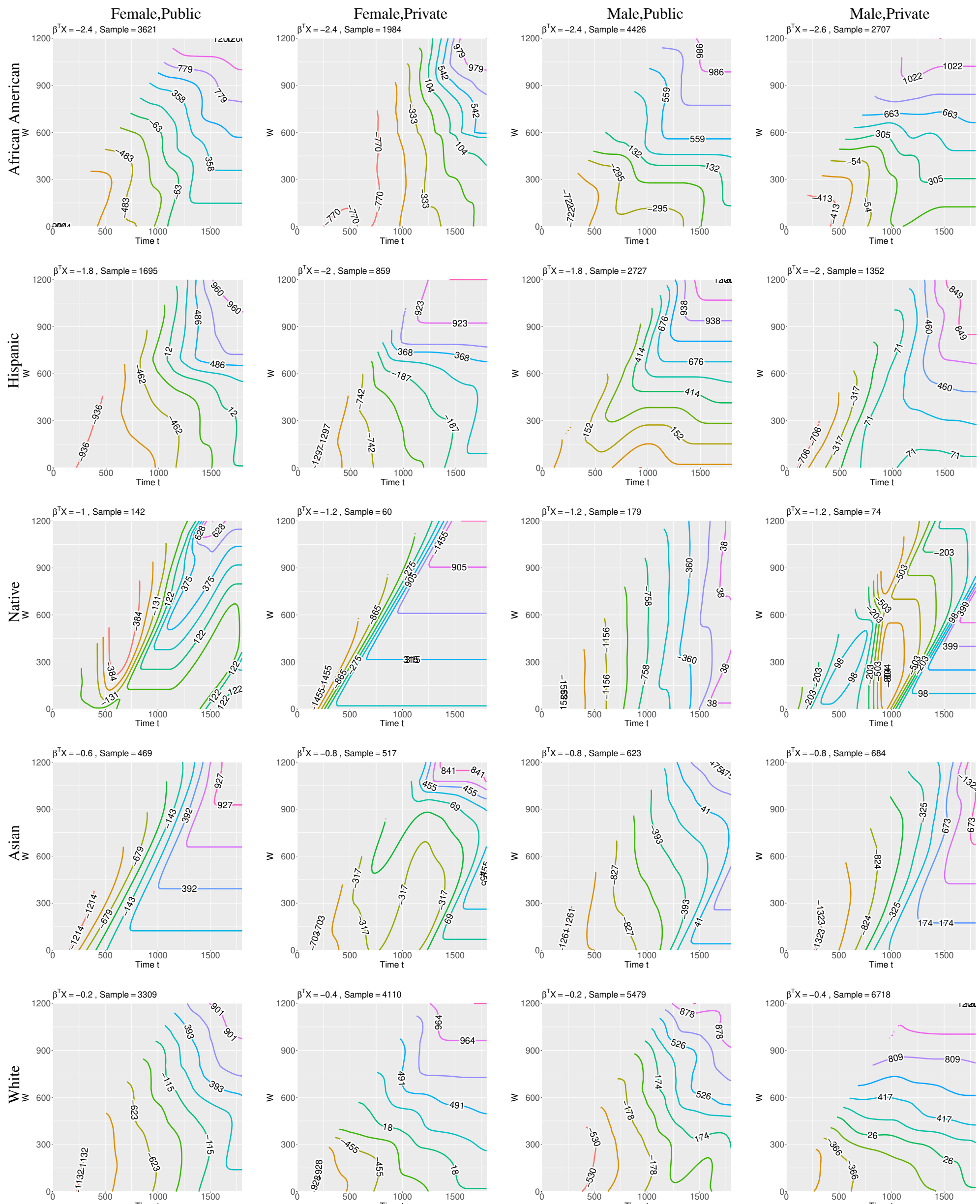

Fig S10: Mean residual life improvement from UNOS/OPTN data. Stratified by Race, Gender, and Insurance Status with minimum $\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}$ per strata.

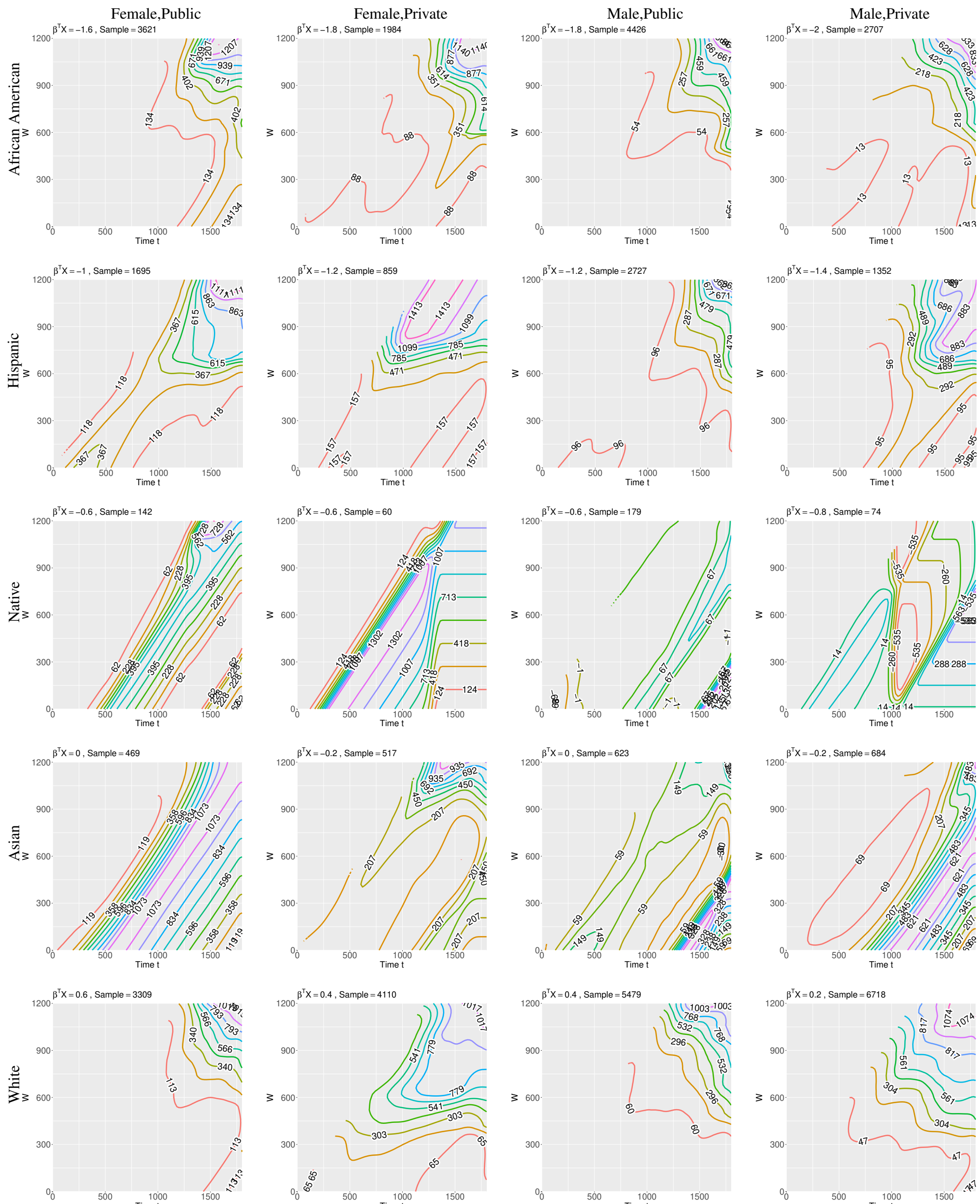

Fig S11: Mean residual life improvement from UNOS/OPTN data. Stratified by Race, Gender, and Insurance Status with median $\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}$ per strata.

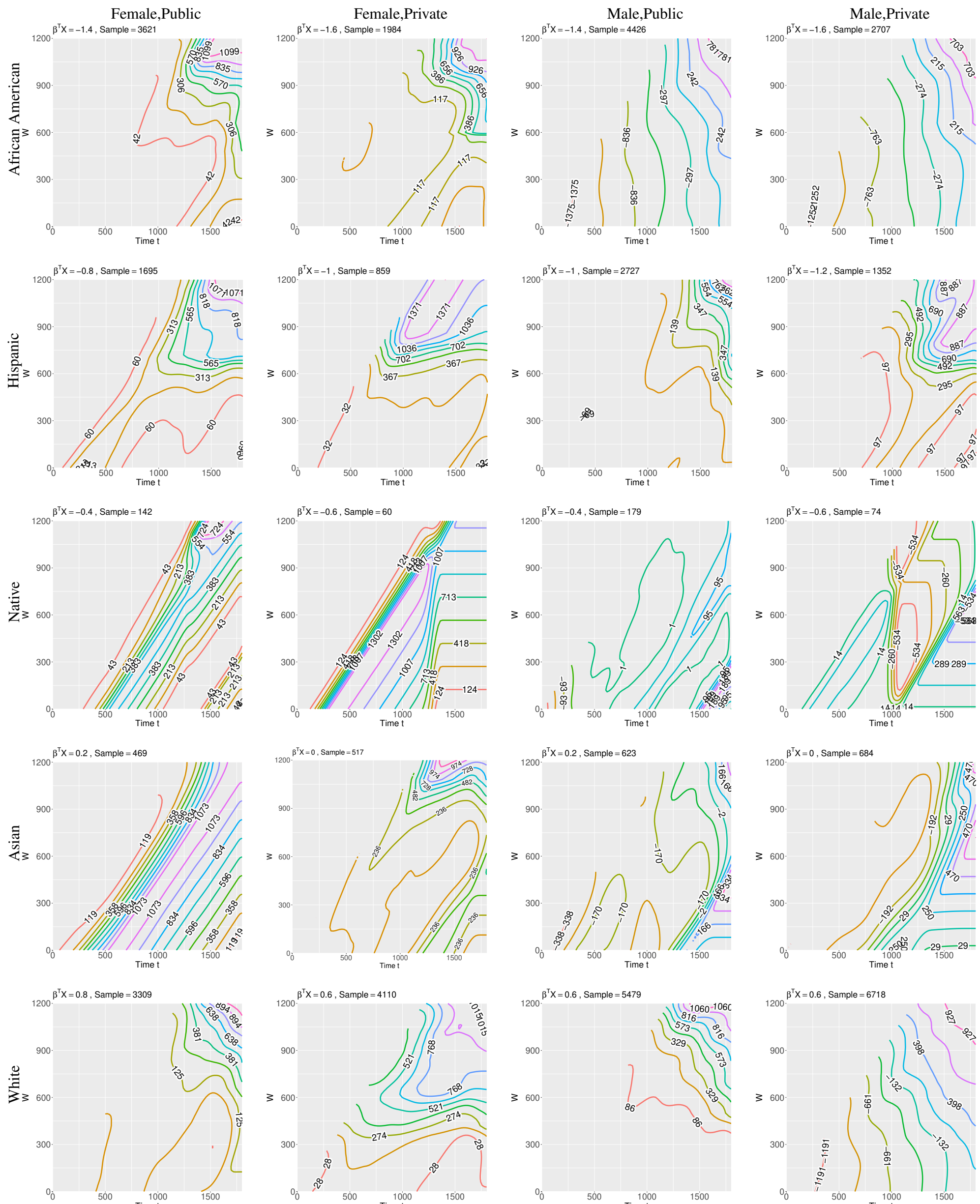

Fig S12: Mean residual life improvement from UNOS/OPTN data. Stratified by Race, Gender, and Insurance Status with maximum $\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}$ per strata.