Smoothed Rank Regression for the Accelerated Failure Time Competing Risks Model with Missing Cause of Failure

Zhiping Qiu, Alan T.K. Wan, Yong Zhou and Peter B. Gilbert

Huaqiao University, City University of Hong Kong,

Shanghai University of Finance and Economics, Chinese Academy of Sciences,

Fred Hutchinson Cancer Research Institute, University of Washington

Supplementary Material

In this note, we provide the the results of two additional simulation experiments and proofs of the main results in the paper.

S1 Additional Simulation Experiments

Experiment 3

This experiment is conducted in response to a query by a referee about the difference in computational time between the proposed smooth approach and the discontinuous rank approach. This experiment is based on the same setup as in Experiment 1, except that we confine our attention to Z_i following a $U[0,1]$ distribution, ϵ_i following a $N(0,0.5^2)$ distribution, and the missing data mechanism of $r(\boldsymbol{W}_i) =$

$\exp(\tilde{T}_i - Z_i)/\{1+\exp(\tilde{T}_i - Z_i)\}$. Thus, on average, 42% of failures are due to the cause of interest, 28% of failures are due to the other cause, and the data missing percentage is approximately 70%. We only report results based on the IPW missing data handling method. Results based on other missing data handling methods are similar and we omit them for brevity. For the non-smooth approach, estimates of the unknown parameters are obtained as solutions to the estimating equation:

$$n^{-3/2}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{R_i}{\widehat{\pi}(\boldsymbol{Q}_i)}I(J_i=2)\delta_i(\boldsymbol{Z}_i-\boldsymbol{Z}_j)I\{e_j^{\boldsymbol{\beta}}\geq e_i^{\boldsymbol{\beta}}\}=0,$$

where $e_i^{\boldsymbol{\beta}}=\log(\tilde{T}_i)-\boldsymbol{Z}_i^T\boldsymbol{\beta}, i=1,2,\cdots,n$. Note that the l.h.s. of above equation is the gradient of the following convex function

$$\boldsymbol{L}(\boldsymbol{\beta})=n^{-3/2}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{R_i}{\widehat{\pi}(\boldsymbol{Q}_i)}I(J_i=2)\delta_i(e_i^{\boldsymbol{\beta}}-e_j^{\boldsymbol{\beta}})^{-},$$

where $a^{-}=|a|I\{a<0\}$. We use the package "fminsearch" in Matlab to minimise $\boldsymbol{L}(\boldsymbol{\beta})$ with respect to $\boldsymbol{\beta}$ and obtain the estimates of β. As discussed previously, for estimating the asymptotic covariance of the estimator, we have to resort to resampling (Jin, Lin, Wei and Ying, 2003). Specifically, we first construct the perturbed objective function

$$\boldsymbol{L}^{*}(\boldsymbol{\beta})=n^{-3/2}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{R_i}{\widehat{\pi}(\boldsymbol{Q}_i)}I(J_i=2)\delta_i(e_i^{\boldsymbol{\beta}}-e_j^{\boldsymbol{\beta}})^{-}\xi_i,$$

where ξ_i follows the exponential distribution with mean 1. Given the data $(\boldsymbol{Q}_i^T, R_i, J_i)^T$, we repeat the resampling process 50 times, and use the standard deviation (SD) of the 50 re-sampled estimates to compute the standard errors of the estimate (SE). Table 4 below reports the results for $n = 400$ observations based on 1000 replications. It can be seen that in addition to delivering more accurate estimates, the smoothed approach has a significant advantage over the non-smoothed approach in terms of computational time.

Table 4: Simulation results of Experiment 3

METHOD	BIAS	SE	SD	CP	computational time
Smoothed	0.001	0.058	0.058	94.7%	314.051s
Non-smoothed	0.020	0.067	0.068	93.9%	21166.368s

Experiment 4

This experiment provides insights on the impact of bandwidth choices on the results. This experiment is conducted in response to a question by a referee. Related studies by Ma and Huang (2007), Song et al. (2007), Lin and Peng (2013) and Qiu, Qin and Zhou (2016) have shown that bandwidth choices do not impact the results significantly. Here, we conduct a simple simulation experiment to examine the sensitivity of results to bandwidth choices. We consider the same setup as in Experiment 2, except that we restrict our attention to $\epsilon_i \sim N(0, 0.25)$ and Scenario 1 of the missing data mechanism. We set the smoothing parameter σ_n to $0.1 \times n^{-0:26}, 0.3 \times n^{-0:26}, 0.5 \times n^{-0:26}, 0.7 \times n^{-0:26}$ and $0.9 \times n^{-0:26}$. The results presented in Table 5 show that for a given estimation method, the results across the different bandwidths are very similar. The assignment of bandwidth σ_n is thus straightforward and does not involve any search.

Table 5: Simulation results of Experiment 4

		$\beta_{01}=1$				$\beta_{02}=1$			
σ_n		BIAS	SE	SD	CP	BIAS	SE	SD	CP
$0.1*n^{-0.26}$	FULL	-0.001	0.111	0.112	94.7%	-0.002	0.064	0.064	94.2%
	CC	0.041	0.136	0.133	94.0%	-0.003	0.076	0.075	93.8%
	IPW	-0.026	0.133	0.134	94.5%	0.001	0.074	0.075	95.1%
	EEI	-0.003	0.117	0.128	96.2%	-0.002	0.070	0.072	95.4%
	AIPW	-0.006	0.132	0.127	94.5%	-0.004	0.075	0.074	95.4%
$0.3*n^{-0.26}$	FULL	-0.001	0.112	0.112	94.6%	-0.003	0.066	0.064	94.0%
	CC	0.040	0.137	0.133	93.3%	-0.004	0.079	0.075	94.1%
	IPW	-0.025	0.133	0.133	94.0%	-0.001	0.077	0.075	94.4%
	EEI	-0.002	0.121	0.131	96.0%	-0.003	0.072	0.074	95.3%
	AIPW	-0.006	0.138	0.129	93.4%	-0.007	0.080	0.073	93.8%
$0.5*n^{-0.26}$	FULL	-0.004	0.116	0.111	94.0%	-0.001	0.066	0.064	94.2%
	CC	0.039	0.141	0.133	92.0%	-0.004	0.080	0.075	93.7%
	IPW	-0.032	0.139	0.129	92.0%	-0.001	0.076	0.073	94.8%
	EEI	-0.008	0.125	0.126	95.7%	-0.004	0.070	0.072	95.7%
	AIPW	-0.009	0.137	0.126	93.2%	-0.005	0.073	0.073	95.1%
$0.7*n^{-0.26}$	FULL	0.002	0.111	0.113	95.6%	0.002	0.064	0.064	95.4%
	CC	0.050	0.130	0.134	94.2%	0.002	0.076	0.075	94.1%
	IPW	-0.019	0.132	0.137	95.4%	0.006	0.075	0.076	94.8%
	EEI	0.004	0.121	0.134	97.1%	0.002	0.071	0.077	96.3%
	AIPW	0.002	0.131	0.129	94.0%	0.001	0.074	0.076	95.4%
$0.9*n^{-0.26}$	FULL	0.006	0.120	0.114	94.4%	0.001	0.063	0.065	95.9%
	CC	0.054	0.141	0.136	93.0%	0.000	0.077	0.077	95.4%
	IPW	-0.018	0.144	0.135	93.6%	0.003	0.074	0.075	95.8%
	EEI	0.005	0.130	0.131	96.0%	-0.001	0.067	0.075	96.9%
	AIPW	0.003	0.146	0.129	93.3%	-0.003	0.072	0.073	95.6%

S2 Appendix: Proof of theorems

In our proof of theorems, for convenience purposes we assume that all elements of $\boldsymbol{W}_i$ are continuous. This assumption does not lead to any loss of generality.

Proof of Theorem 1. We divide the proof into two parts.

Part A1. We can write

$$\begin{aligned}\frac{1}{\sqrt{n}}\nabla\boldsymbol{U}_1(\boldsymbol{\beta}_0) &= \frac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n\frac{R_i}{\pi(\boldsymbol{Q}_i)}\delta_i I(J_i=2)(\boldsymbol{Z}_i-\boldsymbol{Z}_j)^{\otimes 2}s\Big(\frac{r_j^{\boldsymbol{\beta}_0}-r_i^{\boldsymbol{\beta}_0}}{\sigma_n}\Big)\frac{1}{\sigma_n}\\ &\quad+\frac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n\Big(\frac{1}{\widehat{\pi}(\boldsymbol{Q}_i)}-\frac{1}{\pi(\boldsymbol{Q}_i)}\Big)R_i\delta_i I(J_i=2)(\boldsymbol{Z}_i-\boldsymbol{Z}_j)^{\otimes 2}s\Big(\frac{r_j^{\boldsymbol{\beta}_0}-r_i^{\boldsymbol{\beta}_0}}{\sigma_n}\Big)\frac{1}{\sigma_n},\end{aligned} \tag{S2.1}$$

where $s(\cdot)$ is the standard normal density function. By some tedious calculations and recognising the fact that $\sup_{\boldsymbol{q}} \big|\widehat{\pi}(\boldsymbol{q}) - \pi(\boldsymbol{q})\big| = O_p\left(h^r + (nh^d)^{-\frac{1}{2}}\right)$, we obtain

$$\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\big(\frac{1}{\widehat{\pi}(\boldsymbol{Q}_i)} - \frac{1}{\pi(\boldsymbol{Q}_i)}\big)R_i\delta_i I(J_i=2)(\boldsymbol{Z}_i-\boldsymbol{Z}_j)^{\otimes 2} s\Big(\frac{r_j^{\boldsymbol{\beta}_0} - r_i^{\boldsymbol{\beta}_0}}{\sigma_n}\Big)\frac{1}{\sigma_n} = o_p(h^r + (nh^d)^{-\frac{1}{2}}).$$

For the first item on the r.h.s. of (S2.1), note that

$$\begin{aligned}
&\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{R_i}{\pi(\boldsymbol{Q}_i)}\delta_i I(J_i=1)(\boldsymbol{Z}_i-\boldsymbol{Z}_j)^{\otimes 2} s\Big(\frac{r_j^{\boldsymbol{\beta}_0} - r_i^{\boldsymbol{\beta}_0}}{\sigma_n}\Big)\frac{1}{\sigma_n}\\
= \;& \frac{n(n-1)}{n^2}\frac{1}{C_n^2}\sum_{i<j}\frac{1}{2}\Big\{\frac{R_i}{\pi(\boldsymbol{Q}_i)}\delta_i I(J_i=2)(\boldsymbol{Z}_i-\boldsymbol{Z}_j)^{\otimes 2} s\Big(\frac{r_j^{\boldsymbol{\beta}_0} - r_i^{\boldsymbol{\beta}_0}}{\sigma_n}\Big)\frac{1}{\sigma_n}\\
&\quad + \frac{R_j}{\pi(\boldsymbol{Q}_j)}\delta_j I(J_j=2)(\boldsymbol{Z}_j-\boldsymbol{Z}_i)^{\otimes 2} s\Big(\frac{r_i^{\boldsymbol{\beta}_0} - r_j^{\boldsymbol{\beta}_0}}{\sigma_n}\Big)\frac{1}{\sigma_n}\Big\}.
\end{aligned}$$

By using the strong law of large numbers for U-statistics, we can obtain

$$\begin{aligned}
\frac{1}{\sqrt{n}}\nabla \boldsymbol{U}_1(\boldsymbol{\beta}_0) \xrightarrow{a.s.} E\Big\{(\boldsymbol{Z}_1-\boldsymbol{Z}_2)^{\otimes 2}\Big[\int_{-\infty}^{\infty}\overline{H}(u)\overline{F}_{01}(u)f_{02}(u)\xi(u)du\Big]\\
+ \lim_{n\to\infty}\Big(\frac{m_n}{\pi}\Big)^{\frac{1}{2}}\int_{-\infty}^{\infty}\tau(w)e^{-m_n w^2}dw\Big]\Big\},
\end{aligned}$$

where $\tau(w) = \int_{-\infty}^{\infty}\overline{H}(u)\overline{F}_{01}(u)f_{02}(u)\{\xi(u+w)-\xi(u)\}du$, $\xi(s) = f_{01}(s)\overline{F}_{02}(s)\overline{H}(s) + f_{02}(s)\overline{F}_{01}(s)\overline{H}(s) + \overline{F}_{01}(s)\overline{F}_{02}(s)h(s)$, $m_n = n/(2\sigma_n^2)$, $\overline{F}_{01}(\cdot)$ is the survival function of $\log(T_{11}) - \boldsymbol{Z}^T\boldsymbol{\beta}_0$, $\overline{F}_{02}(\cdot)$ is the survival function of $\log(T_{12}) - \boldsymbol{Z}^T\boldsymbol{\beta}_0$, and $\overline{H}(\cdot)$ is the survival function of $\log(C_1) - \boldsymbol{Z}^T\boldsymbol{\beta}_0$. Under conditions (C1)-(C9), the function $\tau(\cdot)$ is integrable, continuous and bounded on $\mathcal{R}$ with $\tau(0)=0$. Thus, the second term on the r.h.s. of (S2.1) vanishes (Kanwal, 1998, p.11). Therefore, we have

$$\frac{1}{\sqrt{n}}\nabla \boldsymbol{U}_1(\boldsymbol{\beta}_0) \xrightarrow{a.s.} \boldsymbol{A} = E\Big\{(\boldsymbol{Z}_1-\boldsymbol{Z}_2)^{\otimes 2}\Big[\int_{-\infty}^{\infty}\overline{H}(u)\overline{F}_{01}(u)f_{02}(u)\xi(u)du\Big]\Big\}.$$

Part A2. By the similar proof as Heller (2007), we note that

$$\begin{aligned}
\boldsymbol{U}_1(\boldsymbol{\beta}_0) &= \frac{1}{n^{\frac{3}{2}}}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{R_i}{\pi(\boldsymbol{Q}_i)}\delta_i I(J_i=2)(\boldsymbol{Z}_i-\boldsymbol{Z}_j)I\{r_j^{\boldsymbol{\beta}_0}\ge r_i^{\boldsymbol{\beta}_0}\} \\
&\quad -\frac{1}{n^{\frac{3}{2}}}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\widehat{\pi}(\boldsymbol{Q}_i)-\pi(\boldsymbol{Q}_i)}{\pi^2(\boldsymbol{Q}_i)}R_i\delta_i I(J_i=2)(\boldsymbol{Z}_i-\boldsymbol{Z}_j)I\{r_j^{\boldsymbol{\beta}_0}\ge r_i^{\boldsymbol{\beta}_0}\} \\
&\quad +\frac{1}{n^{\frac{3}{2}}}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\left(\widehat{\pi}(\boldsymbol{Q}_i)-\pi(\boldsymbol{Q}_i)\right)^2}{\widehat{\pi}(\boldsymbol{Q}_i)\pi^2(\boldsymbol{Q}_i)}R_i\delta_i I(J_i=2)(\boldsymbol{Z}_i-\boldsymbol{Z}_j)I\{r_j^{\boldsymbol{\beta}_0}\ge r_i^{\boldsymbol{\beta}_0}\}+o_p(1) \\
&\equiv \boldsymbol{U}_{11}(\boldsymbol{\beta}_0)-\boldsymbol{U}_{12}(\boldsymbol{\beta}_0)+\boldsymbol{U}_{13}(\boldsymbol{\beta}_0)+o_p(1). \qquad \text{(S2.2)}
\end{aligned}$$

Now, by (C1) and (C8), and recognising that $\sup_{\boldsymbol{q}}\left|\widehat{\pi}(\boldsymbol{q})-\pi(\boldsymbol{q})\right| = O_p\left(h^r+(nh^d)^{-\frac{1}{2}}\right)$, we have

$$\| \boldsymbol{U}_{13}(\boldsymbol{\beta}_0) \| \le C\sqrt{n}\sup_{\boldsymbol{q}}\left|\widehat{\pi}(\boldsymbol{q})-\pi(\boldsymbol{q})\right|^2 = O_p\left(\sqrt{n}h^{2r}+\frac{1}{\sqrt{n}h^d}\right), \qquad \text{(S2.3)}$$

where C is an arbitrary constant. Thus, by condition (C4), $\boldsymbol{U}_{13}(\boldsymbol{\beta}_0)=o_p(1)$.

From the definition of $\widehat{\pi}(\boldsymbol{Q}_i)$, it follows that

$$
\begin{aligned}
&\boldsymbol{U}_{12}(\boldsymbol{\beta}_0)\\
= &\ \frac{1}{n^{\frac{3}{2}}}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\big(\widehat{r}(\boldsymbol{W}_i)-r(\boldsymbol{W}_i)\big)\widehat{G}_n(\boldsymbol{W}_i)}{\pi^2(\boldsymbol{Q}_i)g(\boldsymbol{W}_i)}R_i\delta_i I(J_i=2)(\boldsymbol{Z}_i-\boldsymbol{Z}_j)I\big\{r_j^{\boldsymbol{\beta}_0}\geq r_i^{\boldsymbol{\beta}_0}\big\}\\
&\ -\frac{1}{n^{\frac{3}{2}}}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\big(\widehat{r}(\boldsymbol{W}_i)-r(\boldsymbol{W}_i)\big)\big(\widehat{G}_n(\boldsymbol{W}_i)-g(\boldsymbol{W}_i)\big)}{\pi^2(\boldsymbol{Q}_i)g(\boldsymbol{W}_i)}R_i\delta_i I(J_i=2)\\
&\qquad\qquad \times(\boldsymbol{Z}_i-\boldsymbol{Z}_j)I\big\{r_j^{\boldsymbol{\beta}_0}\geq r_i^{\boldsymbol{\beta}_0}\big\}\\
= &\ \frac{1}{n^{\frac{5}{2}}}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{l=1}^{n}\big(R_l-r(\boldsymbol{W}_i)\big)\pi^{-2}(\boldsymbol{Q}_i)g^{-1}(\boldsymbol{W}_i)\delta_l K_h(\boldsymbol{W}_i-\boldsymbol{W}_l)R_i\delta_i I(J_i=2)\\
&\qquad\qquad\qquad\qquad \times(\boldsymbol{Z}_i-\boldsymbol{Z}_j)I\big\{r_j^{\boldsymbol{\beta}_0}\geq r_i^{\boldsymbol{\beta}_0}\big\}\\
&\ -\frac{1}{n^{\frac{3}{2}}}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\big(\widehat{r}(\boldsymbol{W}_i)-r(\boldsymbol{W}_i)\big)\big(\widehat{G}_n(\boldsymbol{W}_i)-g(\boldsymbol{W}_i)\big)}{\pi^2(\boldsymbol{Q}_i)g(\boldsymbol{W}_i)}R_i\delta_i I(J_i=2)\\
&\ (\boldsymbol{Z}_i-\boldsymbol{Z}_j)I\big\{r_j^{\boldsymbol{\beta}_0}\geq r_i^{\boldsymbol{\beta}_0}\big\}\\
\equiv &\ \boldsymbol{U}_{121}(\boldsymbol{\beta}_0)+\boldsymbol{U}_{122}(\boldsymbol{\beta}_0).
\end{aligned}\tag{S2.4}
$$

Recognising that $\sup_{\boldsymbol{w}}\big|\widehat{r}(\boldsymbol{w})-r(\boldsymbol{w})\big| = O_p\left(h^r+(nh^d)^{-\frac{1}{2}}\right)$ and $\sup_{\boldsymbol{w}}\big|\widehat{G}_n(\boldsymbol{w})-g(\boldsymbol{w})\big| = O_p\left(h^r+(nh^d)^{-\frac{1}{2}}\right)$, we obtain

$$
\boldsymbol{U}_{122}(\boldsymbol{\beta}_0)=O_p\left(\sqrt{n}h^{2r}+\frac{1}{\sqrt{n}h^d}\right)=o_p(1). \tag{S2.5}
$$

Our next task is to prove

$$
\boldsymbol{U}_{121}(\boldsymbol{\beta}_0)=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Big(\frac{R_i}{r(\boldsymbol{W}_i)}-1\Big)\delta_i\rho(\boldsymbol{W}_i)\boldsymbol{\varphi}(\boldsymbol{W}_i)+o_p(1), \tag{S2.6}
$$

where $\varphi(\boldsymbol{w})=E\Big\{(\boldsymbol{Z}_1-\boldsymbol{Z}_2)S\Big(\frac{r_2^{\boldsymbol{\beta}_0}-r_1^{\boldsymbol{\beta}_0}}{\sigma_n}\Big)\Big|\boldsymbol{W}_1=\boldsymbol{w}\Big\}$.

Note that

$$\boldsymbol{U}_{121}(\boldsymbol{\beta}_0) \quad = \quad \boldsymbol{U}_{121}^{[1]}(\boldsymbol{\beta}_0) + \boldsymbol{U}_{121}^{[2]}(\boldsymbol{\beta}_0), \tag{S2.7}$$

where

$$\begin{aligned} \boldsymbol{U}_{121}^{[1]}(\boldsymbol{\beta}_0) \quad = \quad & \frac{1}{n^{\frac{5}{2}}} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{l=1}^{n} \big(R_l - r(\boldsymbol{W}_l)\big) \pi^{-2}(\boldsymbol{Q}_i) g^{-1}(\boldsymbol{W}_i) \delta_l K_h(\boldsymbol{W}_i - \boldsymbol{W}_l) \\ & \times R_i \delta_i I(J_i = 2)(\boldsymbol{Z}_i - \boldsymbol{Z}_j) I\big\{ r_j^{\boldsymbol{\beta}_0} \geq r_i^{\boldsymbol{\beta}_0} \big\} \end{aligned}$$

and

$$\begin{aligned} \boldsymbol{U}_{121}^{[2]}(\boldsymbol{\beta}_0) \quad = \quad & \frac{1}{n^{\frac{5}{2}}} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{l=1}^{n} \big(r(\boldsymbol{W}_l) - r(\boldsymbol{W}_i)\big) \pi^{-2}(\boldsymbol{Q}_i) g^{-1}(\boldsymbol{W}_i) \delta_l K_h(\boldsymbol{W}_i - \boldsymbol{W}_l) \\ & \times R_i \delta_i I(J_i = 2)(\boldsymbol{Z}_i - \boldsymbol{Z}_j) I\big\{ r_j^{\boldsymbol{\beta}_0} \geq r_i^{\boldsymbol{\beta}_0} \big\}. \end{aligned}$$

To analyse $\boldsymbol{U}_{121}^{[1]}(\boldsymbol{\beta}_0)$, similar to Zhou, Wan and Wang (2008), let us define

$$\begin{aligned} \boldsymbol{h}(\boldsymbol{S}_i, \boldsymbol{S}_j, \boldsymbol{S}_l) \quad = \quad & \big(R_l - r(\boldsymbol{W}_l)\big) \pi^{-2}(\boldsymbol{Q}_i) g^{-1}(\boldsymbol{W}_i) \delta_l K_h(\boldsymbol{W}_i - \boldsymbol{W}_l) \\ & \times R_i \delta_i I(J_i = 2)(\boldsymbol{Z}_i - \boldsymbol{Z}_j) I\big\{ r_j^{\boldsymbol{\beta}_0} \geq r_i^{\boldsymbol{\beta}_0} \big\} \end{aligned}$$

and

$$\begin{aligned} \boldsymbol{H}(\boldsymbol{S}_i, \boldsymbol{S}_j, \boldsymbol{S}_l) \quad = \quad & \boldsymbol{h}(\boldsymbol{S}_i, \boldsymbol{S}_j, \boldsymbol{S}_l) + \boldsymbol{h}(\boldsymbol{S}_i, \boldsymbol{S}_l, \boldsymbol{S}_j) + \boldsymbol{h}(\boldsymbol{S}_j, \boldsymbol{S}_i, \boldsymbol{S}_l) \\ & + \boldsymbol{h}(\boldsymbol{S}_l, \boldsymbol{S}_i, \boldsymbol{S}_j) + \boldsymbol{h}(\boldsymbol{S}_j, \boldsymbol{S}_l, \boldsymbol{S}_i) + \boldsymbol{h}(\boldsymbol{S}_l, \boldsymbol{S}_j, \boldsymbol{S}_i), \end{aligned}$$

where $\boldsymbol{S}_i = (\boldsymbol{Q}_i^T, R_i, J_i)^T$, $i, j, l = 1, 2, \ldots, n$. Thus,

$$\begin{aligned}\boldsymbol{U}_{121}^{[1]}(\boldsymbol{\beta}_0) &= n^{-\frac{5}{2}} \sum_{i\neq j, l=i} \boldsymbol{h}(\boldsymbol{S}_i, \boldsymbol{S}_j, \boldsymbol{S}_l) + n^{-\frac{5}{2}} \sum_{i\neq j, l=j} \boldsymbol{h}(\boldsymbol{S}_i, \boldsymbol{S}_j, \boldsymbol{S}_l) + n^{-\frac{5}{2}} \sum_{i\neq j, l\neq i, l\neq l} \boldsymbol{h}(\boldsymbol{S}_i, \boldsymbol{S}_j, \boldsymbol{S}_l) \\ &= n^{-\frac{5}{2}} \sum_{i\neq j, l=i} \boldsymbol{h}(\boldsymbol{S}_i, \boldsymbol{S}_j, \boldsymbol{S}_l) + n^{-\frac{5}{2}} \sum_{i\neq j, l=j} \boldsymbol{h}(\boldsymbol{S}_i, \boldsymbol{S}_j, \boldsymbol{S}_l) + n^{-\frac{5}{2}} \sum_{i<j<l} \boldsymbol{H}(\boldsymbol{S}_i, \boldsymbol{S}_j, \boldsymbol{S}_l).\end{aligned} \tag{S2.8}$$

Let us consider each of the three terms on the r.h.s. of (S2.8). By the theory of U-statistics (van der Vaart, 2000), it can be shown easily that

$$n^{-\frac{5}{2}} \sum_{i\neq j, l=i} \boldsymbol{h}(\boldsymbol{S}_i, \boldsymbol{S}_j, \boldsymbol{S}_l) = O_p(n^{-1}) \ \text{ and } \ n^{-\frac{5}{2}} \sum_{i\neq j, l=j} \boldsymbol{h}(\boldsymbol{S}_i, \boldsymbol{S}_j, \boldsymbol{S}_l) = O_p(n^{-1}). \tag{S2.9}$$

The third term on the r.h.s. of (S2.8) is a U-statistic with symmetric kernel function $\boldsymbol{H}(\cdot,\cdot,\cdot)$. Note that $E\big\{\boldsymbol{H}(\boldsymbol{S}_i, \boldsymbol{S}_j, \boldsymbol{S}_l)\big\} = 0$ and $E\big\{[R_l - r(\boldsymbol{W}_l)]\big|\boldsymbol{W}_l, \delta_l = 1\big\} = 0$. Then, by some manipulations, we can show that $E\big\{\boldsymbol{h}(\boldsymbol{S}_i, \boldsymbol{S}_j, \boldsymbol{S}_l)\big|\boldsymbol{S}_i\big\} = E\big\{\boldsymbol{h}(\boldsymbol{S}_i, \boldsymbol{S}_l, \boldsymbol{S}_j)\big|\boldsymbol{S}_i\big\} = E\big\{\boldsymbol{h}(\boldsymbol{S}_j, \boldsymbol{S}_i, \boldsymbol{S}_l)\big|\boldsymbol{S}_i\big\} = E\big\{\boldsymbol{h}(\boldsymbol{S}_l, \boldsymbol{S}_i, \boldsymbol{S}_j)\big|\boldsymbol{S}_i\big\} = 0$. Also, by standard non-parametric procedures, we can write

$$\begin{aligned} & E\big\{\boldsymbol{h}(\boldsymbol{S}_j, \boldsymbol{S}_l, \boldsymbol{S}_i)\big|\boldsymbol{S}_i\big\} \\ = \ & (R_i - r(\boldsymbol{W}_i))\delta_i E\Big\{K_h(\boldsymbol{W}_j - \boldsymbol{W}_i)\pi^{-2}(\boldsymbol{Q}_j)g^{-1}(\boldsymbol{W}_j) \\ & \qquad \times R_j \delta_j I(J_i = 2)(\boldsymbol{Z}_i - \boldsymbol{Z}_j) I\{r_l^{\boldsymbol{\beta}_0} \geq r_j^{\boldsymbol{\beta}_0}\}\Big|\boldsymbol{S}_i\Big\} \\ = \ & \Big(\frac{R_i}{r(\boldsymbol{W}_i)} - 1\Big)\delta_i \rho(\boldsymbol{W}_i)\varphi(\boldsymbol{W}_i) + O_p(h^r). \end{aligned}$$

Similarly,

$$E\{\boldsymbol{h}(\boldsymbol{S}_l, \boldsymbol{S}_j, \boldsymbol{S}_i)\big|\boldsymbol{S}_i\} = \Big(\frac{R_i}{r(\boldsymbol{W}_i)} - 1\Big)\delta_i \rho(\boldsymbol{W}_i)\varphi(\boldsymbol{W}_i) + O_p(h^r).$$

Therefore, the projection of the kernel function $\boldsymbol{H}(\boldsymbol{S}_i, \boldsymbol{S}_j, \boldsymbol{S}_l)$ is given by

$$E\{\boldsymbol{H}(\boldsymbol{S}_i, \boldsymbol{S}_j, \boldsymbol{S}_l)\big|\boldsymbol{S}_i\} = 2\Big(\frac{R_i}{r(\boldsymbol{W}_i)} - 1\Big)\delta_i \rho(\boldsymbol{W}_i)\varphi(\boldsymbol{W}_i) + O_p(h^r).$$

Thus, by the theory of U-statistics (van der Vaart, 2000, Chap.12),

$$n^{-\frac{5}{2}} \sum_{i<j<l} \boldsymbol{H}(\boldsymbol{S}_i, \boldsymbol{S}_j, \boldsymbol{S}_l) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \Big(\frac{R_i}{r(\boldsymbol{W}_i)} - 1\Big)\delta_i \rho(\boldsymbol{W}_i)\varphi(\boldsymbol{W}_i) + O_p(\sqrt{n}h^r). \tag{S2.10}$$

Combining (S2.8), (S2.9) and (S2.10), it follows that

$$\boldsymbol{U}_{121}^{[1]}(\boldsymbol{\beta}_0) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \Big(\frac{R_i}{r(\boldsymbol{W}_i)} - 1\Big)\delta_i \rho(\boldsymbol{W}_i)\varphi(\boldsymbol{W}_i) + O_p(\sqrt{n}h^r). \tag{S2.11}$$

On the other hand, by some complex calculations as in Zhou, Wan and Wang (2008), we obtain

$$\| \boldsymbol{U}_{121}^{[2]}(\boldsymbol{\beta}_0) \| \leq C\sqrt{n}h^r + o_p(1). \tag{S2.12}$$

Thus, by (S2.7), (S2.11), (S2.12) and condition (C4),

$$\boldsymbol{U}_{121}(\boldsymbol{\beta}_0) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \Big(\frac{R_i}{r(\boldsymbol{W}_i)} - 1\Big)\delta_i \rho(\boldsymbol{W}_i)\varphi(\boldsymbol{W}_i) + o_p(1), \tag{S2.13}$$

and this proves (S2.6). Further, by combining (S2.4)-(S2.6), we have

$$\boldsymbol{U}_{12}(\boldsymbol{\beta}_0) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \Big(\frac{R_i}{r(\boldsymbol{W}_i)} - 1 \Big) \delta_i \rho(\boldsymbol{W}_i) \varphi(\boldsymbol{W}_i) + o_p(1). \tag{S2.14}$$

Analogous to the above derivation and by the theory of U-statistics, we can also obtain

$$\begin{aligned} \boldsymbol{U}_{11}(\boldsymbol{\beta}_0) &= n^{-\frac{3}{2}} \sum_{i=1}^{n} \sum_{j=1}^{n} \delta_i I(J_i = 2)(\boldsymbol{Z}_i - \boldsymbol{Z}_j) I\big\{ r_j^{\boldsymbol{\beta}_0} \geq r_i^{\boldsymbol{\beta}_0} \big\} \\ &\quad + \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \Big(\frac{R_i}{r(\boldsymbol{W}_i)} - 1 \Big) \delta_i I(J_i = 2) \varphi(\boldsymbol{W}_i). \end{aligned} \tag{S2.15}$$

Therefore, by (S2.2), (S2.3), (S2.14) and (S2.15), it follows that

$$\begin{aligned} \boldsymbol{U}_{1}(\boldsymbol{\beta}_0) &= n^{-\frac{3}{2}} \sum_{i=1}^{n} \sum_{j=1}^{n} \delta_i I(J_i = 2)(\boldsymbol{Z}_i - \boldsymbol{Z}_j) I\big\{ r_j^{\boldsymbol{\beta}_0} \geq r_i^{\boldsymbol{\beta}_0} \big\} \\ &\quad + \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \Big(\frac{R_i}{r(\boldsymbol{W}_i)} - 1 \Big) \delta_i \big(I(J_i = 2) - \rho(\boldsymbol{W}_i) \big) \varphi(\boldsymbol{W}_i) + o_p(1). \end{aligned} \tag{S2.16}$$

Note that the first and second terms on the r.h.s. of (S2.16) are uncorrelated. Hence

$$\boldsymbol{U}_1(\boldsymbol{\beta}_0) \xrightarrow{d} N\big(0, \boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2\big),$$

by the Central Limit Theorem. The proof of Theorem 1 can be completed by the Taylor series expansion. We omit the details here for brevity.

Proof of Theorem 2. We divide the proof into two parts.

Part B1. Note that

$$\begin{aligned}
&\frac{1}{\sqrt{n}}\nabla \boldsymbol{U}_2(\boldsymbol{\beta}_0)\\
= &\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\Big\{\big[R_i I(J_i=2)+(1-R_i)\rho(\boldsymbol{W}_i)\big]\delta_i(\boldsymbol{Z}_i-\boldsymbol{Z}_j)^{\otimes 2}s\Big(\frac{r_j^{\boldsymbol{\beta}_0}-r_i^{\boldsymbol{\beta}_0}}{\sigma_n}\Big)\frac{1}{\sigma_n}\Big\}\\
&+\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}(1-R_i)\big(\widehat{\rho}(\boldsymbol{W}_i)-\rho(\boldsymbol{W}_i)\big)\delta_i(\boldsymbol{Z}_i-\boldsymbol{Z}_j)^{\otimes 2}s\Big(\frac{r_j^{\boldsymbol{\beta}_0}-r_i^{\boldsymbol{\beta}_0}}{\sigma_n}\Big)\frac{1}{\sigma_n}.
\end{aligned}$$

By some tedious calculations and the fact that $\sup_{\boldsymbol{w}}\big|\widehat{\rho}(\boldsymbol{w})-\rho(\boldsymbol{w})\big|=O_p\left(h^r+(nh^d)^{-\frac{1}{2}}\right)$, it follows that

$$\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}(1-R_i)\big(\widehat{\rho}(\boldsymbol{W}_i)-\rho(\boldsymbol{W}_i)\big)\delta_i(\boldsymbol{Z}_i-\boldsymbol{Z}_j)^{\otimes 2}s\Big(\frac{r_j^{\boldsymbol{\beta}_0}-r_i^{\boldsymbol{\beta}_0}}{\sigma_n}\Big)\frac{1}{\sigma_n}=o_p(h^r+(nh^d)^{-\frac{1}{2}}).$$

Thus, recognising that $E\big[R_i I(J_i=2)+(1-R_i)\rho(\boldsymbol{W}_i)\big]=E[I(J_i=2)]$ and by derivations similar to those used in the proof of Theorem 1, we obtain

$$\frac{1}{\sqrt{n}}\nabla \boldsymbol{U}_2(\boldsymbol{\beta}_0)\xrightarrow{a.s.}\boldsymbol{A}=E\Big\{(\boldsymbol{Z}_1-\boldsymbol{Z}_2)^{\otimes 2}\Big[\int_{-\infty}^{\infty}\overline{H}(u)\overline{F}_{01}(u)f_{02}(u)\xi(u)du\Big]\Big\}.$$

Part B2. First, note that

$$\begin{aligned}
\boldsymbol{U}_2(\boldsymbol{\beta}_0) &= n^{-\frac{3}{2}}\sum_{i=1}^{n}\sum_{j=1}^{n}\big[R_i I(J_i=2)+(1-R_i)\rho(\boldsymbol{W}_i)\big]\delta_i(\boldsymbol{Z}_i-\boldsymbol{Z}_j)I\{r_j^{\boldsymbol{\beta}_0}\ge r_i^{\boldsymbol{\beta}_0}\}\\
&\quad+n^{-\frac{3}{2}}\sum_{i=1}^{n}\sum_{j=1}^{n}(1-R_i)\big(\widehat{\rho}(\boldsymbol{W}_i)-\rho(\boldsymbol{W}_i)\big)\delta_i(\boldsymbol{Z}_i-\boldsymbol{Z}_j)I\{r_j^{\boldsymbol{\beta}_0}\ge r_i^{\boldsymbol{\beta}_0}\}+o_p(1)\\
&\equiv \boldsymbol{U}_{21}(\boldsymbol{\beta}_0)+\boldsymbol{U}_{22}(\boldsymbol{\beta}_0)+o_p(1). \qquad \text{(S2.17)}
\end{aligned}$$

Now, we can write

$$\begin{aligned}
\boldsymbol{U}_{22}(\boldsymbol{\beta}_0) &= n^{-\frac{3}{2}}\sum_{i=1}^{n}\sum_{j=1}^{n}(1-R_i)\big(\widehat{\rho}(\boldsymbol{W}_i)-\rho(\boldsymbol{W}_i)\big)\widehat{M}_n(\boldsymbol{W}_i) \\
&\qquad\times m^{-1}(\boldsymbol{W}_i)\delta_i(\boldsymbol{Z}_i-\boldsymbol{Z}_j)I\big\{r_j^{\boldsymbol{\beta}_0}\geq r_i^{\boldsymbol{\beta}_0}\big\} \\
&\quad -n^{-\frac{3}{2}}\sum_{i=1}^{n}\sum_{j=1}^{n}(1-R_i)\big(\widehat{\rho}(\boldsymbol{W}_i)-\rho(\boldsymbol{W}_i)\big)\big(\widehat{M}_n(\boldsymbol{W}_i)-m(\boldsymbol{W}_i)\big) \\
&\qquad\times m^{-1}(\boldsymbol{W}_i)\delta_i(\boldsymbol{Z}_i-\boldsymbol{Z}_j)I\big\{r_j^{\boldsymbol{\beta}_0}\geq r_i^{\boldsymbol{\beta}_0}\big\} \\
&\equiv \boldsymbol{U}_{22}^{[1]}(\boldsymbol{\beta}_0)+\boldsymbol{U}_{22}^{[2]}(\boldsymbol{\beta}_0),
\end{aligned} \tag{S2.18}$$

where $m(\boldsymbol{w}) = \pi(\boldsymbol{w})g(\boldsymbol{w})$. We can easily show by steps similar to those for proving Theorem 1 that $\boldsymbol{U}_{22}^{[2]}(\boldsymbol{\beta}_0) = o_p(1)$. Also, by the definition of $\widehat{M}_n(\boldsymbol{W}_i)$, we have

$$\begin{aligned}
&\boldsymbol{U}_{22}^{[1]}(\boldsymbol{\beta}_0) \\
&= n^{-\frac{5}{2}}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{l=1}^{n}(1-R_i)\big(I(J_l=1)-\rho(\boldsymbol{W}_i)\big)R_l\delta_lK_h(\boldsymbol{W}_i-\boldsymbol{W}_l) \\
&\qquad\times m^{-1}(\boldsymbol{W}_i)\delta_i(\boldsymbol{Z}_i-\boldsymbol{Z}_j)I\big\{r_j^{\boldsymbol{\beta}_0}\geq r_i^{\boldsymbol{\beta}_0}\big\} \\
&= n^{-\frac{5}{2}}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{l=1}^{n}(1-R_i)\big(I(J_l=1)-\rho(\boldsymbol{W}_l)\big)R_l\delta_lK_h(\boldsymbol{W}_i-\boldsymbol{W}_l) \\
&\qquad\times m^{-1}(\boldsymbol{W}_i)\delta_i(\boldsymbol{Z}_i-\boldsymbol{Z}_j)I\big\{r_j^{\boldsymbol{\beta}_0}\geq r_i^{\boldsymbol{\beta}_0}\big\} \\
&\quad +n^{-\frac{5}{2}}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{l=1}^{n}(1-R_i)\big(\rho(\boldsymbol{W}_l)-\rho(\boldsymbol{W}_i)\big)R_l\delta_lK_h(\boldsymbol{W}_i-\boldsymbol{W}_l) \\
&\qquad\times m^{-1}(\boldsymbol{W}_i)\delta_i(\boldsymbol{Z}_i-\boldsymbol{Z}_j)I\big\{r_j^{\boldsymbol{\beta}_0}\geq r_i^{\boldsymbol{\beta}_0}\big\} \\
&\equiv \boldsymbol{I}_1+\boldsymbol{I}_2.
\end{aligned} \tag{S2.19}$$

By arguments similar to those used for proving Theorem 1, we have

$$\boldsymbol{I}_2 = o_p(1), \tag{S2.20}$$

and

$$\boldsymbol{I}_1 = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \big(I(J_i = 2) - \rho(\boldsymbol{W}_i)\big) R_i \delta_i \big(1 - r(\boldsymbol{W}_i)\big) r^{-1}(\boldsymbol{W}_i) \boldsymbol{\varphi}(\boldsymbol{W}_i) + O_p(\sqrt{n} h^r). \tag{S2.21}$$

Using (S2.17)-(S2.21) together, and by the theory of U-statistics, we obtain

$$\begin{aligned}
& \boldsymbol{U}_2(\boldsymbol{\beta}_0) \\
= \; & n^{-\frac{3}{2}} \sum_{i=1}^{n} \sum_{j=1}^{n} I(J_i = 2) \delta_i (\boldsymbol{Z}_i - \boldsymbol{Z}_j) I\big\{r_j^{\boldsymbol{\beta}_0} \geq r_i^{\boldsymbol{\beta}_0}\big\} \\
& - n^{-\frac{3}{2}} \sum_{i=1}^{n} \sum_{j=1}^{n} (1 - R_i)\big(I(J_i = 2) - \rho(\boldsymbol{W}_i)\big) \delta_i (\boldsymbol{Z}_i - \boldsymbol{Z}_j) I\big\{r_j^{\boldsymbol{\beta}_0} \geq r_i^{\boldsymbol{\beta}_0}\big\} \\
& + \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \big(I(J_i = 2) - \rho(\boldsymbol{W}_i)\big) R_i \delta_i \big(1 - r(\boldsymbol{W}_i)\big) r^{-1}(\boldsymbol{W}_i) \boldsymbol{\varphi}(\boldsymbol{W}_i) + o_p(1) \\
= \; & n^{-\frac{3}{2}} \sum_{i=1}^{n} \sum_{j=1}^{n} I(J_i = 2) \delta_i (\boldsymbol{Z}_i - \boldsymbol{Z}_j) I\big\{r_j^{\boldsymbol{\beta}_0} \geq r_i^{\boldsymbol{\beta}_0}\big\} \\
& + \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \big(I(J_i = 2) - \rho(\boldsymbol{W}_i)\big) \Big(\frac{R_i}{r(\boldsymbol{W}_i)} - 1\Big) \delta_i \boldsymbol{\varphi}(\boldsymbol{W}_i) + o_p(1).
\end{aligned}$$

The proof of Theorem 2 may be completed by using steps analogous to those used for proving Theorem 1.

Proof of Theorem 3. We divide the proof into two parts.

Part C1. Write

$$
\begin{aligned}
& \frac{1}{\sqrt{n}} \nabla \boldsymbol{U}_3(\boldsymbol{\beta}_0) \\
= \quad & \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \delta_i \Big[\frac{R_i}{\pi(\boldsymbol{Q}_i)} I(J_i = 2) + (1 - \frac{R_i}{\pi(\boldsymbol{Q}_i)}) \rho(\boldsymbol{W}_i) \Big] (\boldsymbol{Z}_i - \boldsymbol{Z}_j) s \left(\frac{r_j^{\boldsymbol{\beta}} - r_i^{\boldsymbol{\beta}}}{\sigma_n} \right) \frac{1}{\sigma_n} \\
& + \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \delta_i \Big[\frac{R_i}{\widehat{\pi}(\boldsymbol{Q}_i)} - \frac{R_i}{\pi(\boldsymbol{Q}_i)} \Big] (\boldsymbol{Z}_i - \boldsymbol{Z}_j) s \left(\frac{r_j^{\boldsymbol{\beta}} - r_i^{\boldsymbol{\beta}}}{\sigma_n} \right) \frac{1}{\sigma_n} \\
& + \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \delta_i \big(1 - \frac{R_i}{\pi(\boldsymbol{Q}_i)}\big) \big(\widehat{\rho}(\boldsymbol{W}_i) - \rho(\boldsymbol{W}_i) \big) (\boldsymbol{Z}_i - \boldsymbol{Z}_j) s \left(\frac{r_j^{\boldsymbol{\beta}} - r_i^{\boldsymbol{\beta}}}{\sigma_n} \right) \frac{1}{\sigma_n} \\
& + \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\widehat{\pi}(\boldsymbol{Q}_i) - \pi(\boldsymbol{Q}_i)}{\widehat{\pi}(\boldsymbol{Q}_i) \widehat{\pi}(\boldsymbol{Q}_i)} \big(\widehat{\rho}(\boldsymbol{W}_i) - \rho(\boldsymbol{W}_i) \big) (\boldsymbol{Z}_i - \boldsymbol{Z}_j) \delta_i R_i s \left(\frac{r_j^{\boldsymbol{\beta}} - r_i^{\boldsymbol{\beta}}}{\sigma_n} \right) \frac{1}{\sigma_n} \\
& + \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \big(\frac{R_i}{\pi(\boldsymbol{Q}_i)} - \frac{R_i}{\widehat{\pi}(\boldsymbol{Q}_i)} \big) \delta_i \rho(\boldsymbol{W}_i) (\boldsymbol{Z}_i - \boldsymbol{Z}_j) s \left(\frac{r_j^{\boldsymbol{\beta}} - r_i^{\boldsymbol{\beta}}}{\sigma_n} \right) \frac{1}{\sigma_n}. \qquad \text{(S2.22)}
\end{aligned}
$$

By steps analogous to those used for proving Theorems 1 and 2, we can show that the last four items of the r.h.s. of (S2.22) are $o_p(1)$. Furthermore, noting that $E\Big[\frac{R_i}{\pi(\boldsymbol{Q}_i)} I(J_i = 2) + (1 - \frac{R_i}{\pi(\boldsymbol{Q}_i)}) \rho(\boldsymbol{W}_i)\Big] = E\big[I(J_i = 2)\big]$, we have

$$
\frac{1}{\sqrt{n}} \nabla \boldsymbol{U}_3(\boldsymbol{\beta}_0) \xrightarrow{a.s.} \boldsymbol{A} = E\Big\{ (\boldsymbol{Z}_1 - \boldsymbol{Z}_2)^{\otimes 2} \Big[\int_{-\infty}^{\infty} \overline{H}(u) \overline{F}_{01}(u) f_{02}(u) \xi(u) du \Big] \Big\}.
$$

Part C2. Note that

$$
\begin{aligned}
& \boldsymbol{U}_3(\boldsymbol{\beta}_0) \\
= \;& n^{-3/2}\sum_{i=1}^n\sum_{j=1}^n \delta_i\big[\frac{R_i}{\pi(\boldsymbol{Q}_i)}I(J_i=2)+(1-\frac{R_i}{\pi(\boldsymbol{Q}_i)})\rho(\boldsymbol{W}_i)\big](\boldsymbol{Z}_i-\boldsymbol{Z}_j)I\big\{r_j^{\boldsymbol{\beta}_0}\geq r_i^{\boldsymbol{\beta}_0}\big\} \\
& +n^{-3/2}\sum_{i=1}^n\sum_{j=1}^n \delta_i\big[\frac{R_i}{\widehat{\pi}(\boldsymbol{Q}_i)}-\frac{R_i}{\pi(\boldsymbol{Q}_i)}\big]I(J_i=2)(\boldsymbol{Z}_i-\boldsymbol{Z}_j)I\big\{r_j^{\boldsymbol{\beta}_0}\geq r_i^{\boldsymbol{\beta}_0}\big\} \\
& +n^{-3/2}\sum_{i=1}^n\sum_{j=1}^n \delta_i\big[(1-\frac{R_i}{\widehat{\pi}(\boldsymbol{Q}_i)})\widehat{\rho}(\boldsymbol{W}_i)-\big(1-\frac{R_i}{\pi(\boldsymbol{Q}_i)}\big)\rho(\boldsymbol{W}_i)\big](\boldsymbol{Z}_i-\boldsymbol{Z}_j)I\big\{r_j^{\boldsymbol{\beta}_0}\geq r_i^{\boldsymbol{\beta}_0}\big\} \\
\equiv \;& \boldsymbol{U}_{31}(\boldsymbol{\beta}_0)+\boldsymbol{U}_{32}(\boldsymbol{\beta}_0)+\boldsymbol{U}_{33}(\boldsymbol{\beta}_0).
\end{aligned} \tag{S2.23}
$$

Similar to the proof of Theorem 1, we have

$$
\boldsymbol{U}_{32}(\boldsymbol{\beta}_0)=-\frac{1}{\sqrt{n}}\sum_{i=1}^n\big(\frac{R_i}{r(\boldsymbol{W}_i)}-1\big)\delta_i\rho(\boldsymbol{W}_i)\boldsymbol{\varphi}(\boldsymbol{W}_i)+o_p(1). \tag{S2.24}
$$

Also, note that

$$
\begin{aligned}
& \boldsymbol{U}_{33}(\boldsymbol{\beta}_0) \\
= \;& n^{-3/2}\sum_{i=1}^n\sum_{j=1}^n \delta_i\big(1-\frac{R_i}{\pi(\boldsymbol{Q}_i)}\big)\big(\widehat{\rho}(\boldsymbol{W}_i)-\rho(\boldsymbol{W}_i)\big)(\boldsymbol{Z}_i-\boldsymbol{Z}_j)I\big\{r_j^{\boldsymbol{\beta}_0}\geq r_i^{\boldsymbol{\beta}_0}\big\} \\
& +n^{-3/2}\sum_{i=1}^n\sum_{j=1}^n \frac{\widehat{\pi}(\boldsymbol{Q}_i)-\pi(\boldsymbol{Q}_i)}{\widehat{\pi}(\boldsymbol{Q}_i)\widehat{\pi}(\boldsymbol{Q}_i)}\big(\widehat{\rho}(\boldsymbol{W}_i)-\rho(\boldsymbol{W}_i)\big)(\boldsymbol{Z}_i-\boldsymbol{Z}_j)\delta_iR_iI\big\{r_j^{\boldsymbol{\beta}_0}\geq r_i^{\boldsymbol{\beta}_0}\big\} \\
& +n^{-3/2}\sum_{i=1}^n\sum_{j=1}^n \big(\frac{R_i}{\pi(\boldsymbol{Q}_i)}-\frac{R_i}{\widehat{\pi}(\boldsymbol{Q}_i)}\big)\delta_i\rho(\boldsymbol{W}_i)(\boldsymbol{Z}_i-\boldsymbol{Z}_j)I\big\{r_j^{\boldsymbol{\beta}_0}\geq r_i^{\boldsymbol{\beta}_0}\big\} \\
\equiv \;& \boldsymbol{U}_{33}^{[1]}(\boldsymbol{\beta}_0)+\boldsymbol{U}_{33}^{[2]}(\boldsymbol{\beta}_0)+\boldsymbol{U}_{33}^{[3]}(\boldsymbol{\beta}_0).
\end{aligned} \tag{S2.25}
$$

It is clear that $\boldsymbol{U}_{33}^{[2]}(\boldsymbol{\beta}_0) = o_p(1)$, and similar to the proof of Theorem 2, it follows that

$$\begin{aligned}
\boldsymbol{U}_{33}^{[1]}(\boldsymbol{\beta}_0) &= n^{-\frac{3}{2}} \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{l=1}^{n} \big(1 - \frac{R_i}{\pi(\boldsymbol{Q}_i)}\big)\big(I(J_i = 2) - \rho(\boldsymbol{W}_i)\big) R_l \delta_l \\
&\quad \times K_h(\boldsymbol{W}_i - \boldsymbol{W}_l) m^{-1}(\boldsymbol{W}_i)\delta_i(\boldsymbol{Z}_i - \boldsymbol{Z}_j) I\big\{r_j^{\boldsymbol{\beta}_0} \geq r_i^{\boldsymbol{\beta}_0}\big\} + o_p(1) \\
&= O_p(\sqrt{n}h^r) + o_p(1) = o_p(1). \qquad \text{(S2.26)}
\end{aligned}$$

Moreover, by arguments similar to those used for the proof of Theorem 1, we can write

$$\boldsymbol{U}_{33}^{[3]}(\boldsymbol{\beta}_0) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \big(\frac{R_i}{r(\boldsymbol{W}_i)} - 1\big)\delta_i \rho(\boldsymbol{W}_i)\boldsymbol{\varphi}(\boldsymbol{W}_i) + o_p(1). \qquad \text{(S2.27)}$$

Thus, using (S2.25)-(S2.27) together,

$$\boldsymbol{U}_{33}(\boldsymbol{\beta}_0) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \big(\frac{R_i}{r(\boldsymbol{W}_i)} - 1\big)\delta_i \rho(\boldsymbol{W}_i)\boldsymbol{\varphi}(\boldsymbol{W}_i) + o_p(1). \qquad \text{(S2.28)}$$

Finally, by combining (S2.23), (S2.24) and (S2.28) and the theory of U-statistics, we obtain

$$\begin{aligned}
\boldsymbol{U}_3(\boldsymbol{\beta}_0) &= n^{-\frac{3}{2}} \sum_{i=1}^{n}\sum_{j=1}^{n} I(J_i = 2)\delta_i(\boldsymbol{Z}_i - \boldsymbol{Z}_j) I\big\{r_j^{\boldsymbol{\beta}_0} \geq r_i^{\boldsymbol{\beta}_0}\big\} \\
&\quad + \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \big(I(J_i = 2) - \rho(\boldsymbol{W}_i)\big)\Big(\frac{R_i}{r(\boldsymbol{W}_i)} - 1\Big)\delta_i \boldsymbol{\varphi}(\boldsymbol{W}_i) + o_p(1).
\end{aligned}$$

The proof of Theorem 3 may be completed by using arguments analogous to those used for proving Theorem 1.

Additional Bibliography

Lin, H. and Peng, H. (2013). Smoothed rank correlation of the linear transformation regression model. *Comput. Statist. Data An.* **19**, 1370-1402.

Ma, S. and Huang, J. (2007). Combining multiple markers for classification using ROC. *Biometrics* **63**, 751-757.

Qiu, Z., Qin, J. and Zhou, Y. (2016). Composite Estimating Equation Method for the Accelerated Failure Time Model with Length-biased Sampling Data. *Scand. J. Statist.* **43**, 396-415.

Song, X., Ma, S., Huang, J. and Zhou, X. H. (2007). A semiparametric approach for the nonparametric transformation survival model with multiple covariates. *Biostatistics* **8**, 197-211.

van der Vaart, A. W. (2000). *Asymptotic Statistics.* Cambridge, UK: Cambridge University Press. *Ann. Inst. Statist. Math.* **68**, 855-876.

Zhou, Y., Wan, A. T. K. and Wang, X. (2008). Estimating equations inference with missing data. *J. Amer. Statist. Assoc.* **103**, 1187-1199.