BIOMETRICS 1, 1–15 DOI: 000
April 2017

Web-based Supplementary Materials for "Generalized accelerated recurrence time model for multivariate recurrent events data with missing event type"

Huijuan Ma[1], Limin Peng[1,*], Zhumin Zhang[2], and HuiChuan J. Lai[2]

[1]Department of Biostatistics and Bioinformatics, Emory University, Atlanta, GA, 30322, U.S.A.

[2]Department of Nutritional Sciences, University of Wisconsin–Madison, Madison, WI, 53706, U.S.A.

*email: lpeng@emory.edu

This paper has been submitted for consideration for publication in *Biometrics*

S1. Derivation of the Closed Forms of $\hat{\pi}(t_0, \mathbf{z}_0)$ and $\hat{p}_k(t_0, \mathbf{z}_0)$

First, we give the derivation of $\hat{\pi}(t_0, \mathbf{z}_0)$. As we stated in Section 2, under the MAR assumption, $\pi_k(t, \mathbf{z}) = E\{A_i(t)|dN_{ik}(t) = 1, \mathbf{Z} = \mathbf{z}\}$ is the same for each k. Suppose $\pi(t, \mathbf{z}) \approx \pi(t_0, \mathbf{z}_0) \doteq \pi$ when $t \in (t_0 - h, t_0 + h)$, $\mathbf{z}_1 \in (\mathbf{z}_{10} - \mathbf{h}, \mathbf{z}_{10} + \mathbf{h})$ and $\mathbf{z}_2 = \mathbf{z}_{20}$, where $\mathbf{h}$ is a d dimensional vector with each component as h. The likelihood function of π for subject i in $t \in (t_0 - h, t_0 + h)$, $\mathbf{z}_1 \in (\mathbf{z}_{10} - \mathbf{h}, \mathbf{z}_{10} + \mathbf{h})$ and $\mathbf{z}_2 = \mathbf{z}_{20}$ is $\pi^{A_i(t)}(1 - \pi)^{1-A_i(t)}$. The log-likelihood function of π for subject i in $t \in (t_0 - h, t_0 + h)$, $\mathbf{z}_1 \in (\mathbf{z}_{10} - \mathbf{h}, \mathbf{z}_{10} + \mathbf{h})$ and $\mathbf{z}_2 = \mathbf{z}_{20}$ is $A_i(t)\log(\pi) + \{1 - A_i(t)\}\log(1 - \pi)$. The full local log-likelihood for π is

$$\iota(\pi) = \sum_{i=1}^{n} \mathbf{K}_h(\mathbf{Z}_{1i} - \mathbf{z}_{10}) I(\mathbf{Z}_{2i} = \mathbf{z}_{20}) \int K_h(t - t_0)[A_i(t)\log(\pi) + \{1 - A_i(t)\}\log(1 - \pi)]dN_{i\cdot}(t).$$

An estimator of π is given by the maximizer of $\iota(\pi)$,

$$\hat{\pi}(t_0, \mathbf{z}_0) = \frac{\sum_{i=1}^{n} \mathbf{K}_h(\mathbf{Z}_{1i} - \mathbf{z}_{10}) I(\mathbf{Z}_{2i} = \mathbf{z}_{20}) \int K_h(t - t_0) A_i(t) dN_{i\cdot}(t)}{\sum_{i=1}^{n} \mathbf{K}_h(\mathbf{Z}_{1i} - \mathbf{z}_{10}) I(\mathbf{Z}_{2i} = \mathbf{z}_{20}) \int K_h(t - t_0) dN_{i\cdot}(t)}.$$

Next we give the derivation of $\hat{p}_k(t_0, \mathbf{z}_0)$. Under the MAR assumption, we can show that

$$\begin{aligned} p_k(t_0, \mathbf{z}_0) &\doteq \Pr\{\delta_{ik}(t_0) = 1 | A_i(t_0) = 0, dN_{i\cdot}(t_0) = 1, \mathbf{Z}_i = \mathbf{z}_0\} \\ &= \Pr\{\delta_{ik}(t_0) = 1 | A_i(t_0) = 1, dN_{i\cdot}(t_0) = 1, \mathbf{Z}_i = \mathbf{z}_0\}. \end{aligned}$$

Suppose $p_k(t, \mathbf{z}) \approx p_k(t_0, \mathbf{z}_0) \doteq p_k$ when $t \in (t_0 - h, t_0 + h)$, $\mathbf{z}_1 \in (\mathbf{z}_{10} - \mathbf{h}, \mathbf{z}_{10} + \mathbf{h})$ and $\mathbf{z}_2 = \mathbf{z}_{20}$. The likelihood function of p_k for subject i in $t \in (t_0 - h, t_0 + h)$, $\mathbf{z}_1 \in (\mathbf{z}_{10} - \mathbf{h}, \mathbf{z}_{10} + \mathbf{h})$ and $\mathbf{z}_2 = \mathbf{z}_{20}$ is

$$\begin{aligned} & \prod_{k=1}^{K-1} p_k^{\delta_{ik}(t)} \cdot \left(1 - \sum_{k=1}^{K-1} p_k\right)^{1 - \sum_{k=1}^{K-1} \delta_{ik}(t)} \\ = & \prod_{k=1}^{K-1} \left(\frac{p_k}{1 - \sum_{k=1}^{K-1} p_k}\right)^{\delta_{ik}(t)} \left(1 - \sum_{k=1}^{K-1} p_k\right). \end{aligned}$$

Define $\theta_k = \log\{p_k/(1 - \sum_{k=1}^{K-1} p_k)\}$. Then $p_k = \exp(\theta_k)/(1 + \sum_{k=1}^{K-1}\exp(\theta_k))$. The log-likelihood function of θ_k for subject i in $t \in (t_0 - h, t_0 + h)$, $\mathbf{z}_1 \in (\mathbf{z}_{10} - \mathbf{h}, \mathbf{z}_{10} + \mathbf{h})$ and

$\mathbf{z}_2 = \mathbf{z}_{20}$ is

$$l_i(t, t_0, \mathbf{z}, \mathbf{z}_0; \theta) = \sum_{k=1}^{K-1} \delta_{ik}(t)\theta_k - \log\left(1 + \sum_{k=1}^{K-1} \exp(\theta_k)\right).$$

The full local log-likelihood for $\theta = (\theta_1, \ldots, \theta_{K-1})$ is

$$l(\theta) = \sum_{i=1}^{n} \mathbf{K}_h(\mathbf{Z}_{1i} - \mathbf{z}_{10}) I(\mathbf{Z}_{2i} = \mathbf{z}_{20}) \int K_h(t - t_0) A_i(t) l_i(t, t_0, \mathbf{z}, \mathbf{z}_0; \theta) dN_{i\cdot}(t).$$

An estimator of θ is given by the maximizer of $l(\theta)$, which leads to

$$\hat{p}_k(t_0, \mathbf{z}_0) = \frac{\sum_{i=1}^{n} \mathbf{K}_h(\mathbf{Z}_{1i} - \mathbf{z}_{10}) I(\mathbf{Z}_{2i} = \mathbf{z}_{20}) \int K_h(t - t_0) A_i(t) \delta_{ik}(t) dN_{i\cdot}(t)}{\sum_{i=1}^{n} \mathbf{K}_h(\mathbf{Z}_{1i} - \mathbf{z}_{10}) I(\mathbf{Z}_{2i} = \mathbf{z}_{20}) \int K_h(t - t_0) A_i(t) dN_{i\cdot}(t)}.$$

S2. Proofs of Theoretical Results

S2.1. Notation and technical lemmas

Define

$$
\begin{aligned}
N_{ik}^{\text{IPW}}(t) &= \sum_{j=1}^{\infty} \frac{A_i^{(j)}}{\pi_i^{(j)}} \delta_{ik}^{(j)} I(L_i < T_i^{(j)} \leqslant R_i \wedge t), \\
N_{ik}^{\text{EEP}}(t) &= \sum_{j=1}^{\infty} \{A_i^{(j)} \delta_{ik}^{(j)} + (1 - A_i^{(j)}) p_{ik}^{(j)}\} I(L_i < T_i^{(j)} \leqslant R_i \wedge t), \\
\hat{G}_{0n}(t_0, \mathbf{z}_0) &= \frac{1}{n} \sum_{m=1}^{n} \mathbf{K}_h(\mathbf{Z}_{1m} - \mathbf{z}_{10}) I(\mathbf{Z}_{2m} = \mathbf{z}_{20}) \int K_h(t - t_0) dN_{m\cdot}(t), \\
\hat{G}_{1n}(t_0, \mathbf{z}_0) &= \frac{1}{n} \sum_{m=1}^{n} \mathbf{K}_h(\mathbf{Z}_{1m} - \mathbf{z}_{10}) I(\mathbf{Z}_{2m} = \mathbf{z}_{20}) \int K_h(t - t_0) A_m(t) dN_{m\cdot}(t), \\
\hat{G}_{2kn}(t_0, \mathbf{z}_0) &= \frac{1}{n} \sum_{m=1}^{n} \mathbf{K}_h(\mathbf{Z}_{1m} - \mathbf{z}_{10}) I(\mathbf{Z}_{2m} = \mathbf{z}_{20}) \int K_h(t - t_0) A_m(t) \delta_{mk}(t) dN_{m\cdot}(t).
\end{aligned}
$$

LEMMA 1: *Under the regularity conditions C6 and C7, we have*

$$
\begin{aligned}
\sup_{t,\mathbf{z}} \|\hat{\pi}(t, \mathbf{z}) - \pi(t, \mathbf{z})\| &= O_p\left(h^r + \left(\frac{\log n}{nh^{d+1}}\right)^{1/2}\right), \\
\sup_{t,\mathbf{z}} \|\hat{p}_k(t, \mathbf{z}) - p_k(t, \mathbf{z})\| &= O_p\left(h^r + \left(\frac{\log n}{nh^{d+1}}\right)^{1/2}\right),
\end{aligned}
$$

$$\sup_{t,\mathbf{z}} \|\hat{G}_{1n}(t,\mathbf{z}) - \pi(t,\mathbf{z})\hat{G}_{0n}(t,\mathbf{z})\| = O_p\left(h^r + \left(\frac{\log n}{nh^{d+1}}\right)^{1/2}\right),$$

$$\sup_{t,\mathbf{z}} \|\hat{G}_{2kn}(t,\mathbf{z}) - p_k(t,\mathbf{z})\hat{G}_{1n}(t,\mathbf{z})\| = O_p\left(h^r + \left(\frac{\log n}{nh^{d+1}}\right)^{1/2}\right).$$

Proof. Following the usual arguments in nonparametric regression (Mack and Silverman, 1982), we can show that

$$\sup_{t,\mathbf{z}} \left\|\hat{G}_{0n}(t,\mathbf{z}) - f(\mathbf{z})g_{\mathbf{Z}}(t)\right\| = O_p\left(h^r + \left(\frac{\log n}{nh^{d+1}}\right)^{1/2}\right),$$

$$\sup_{t,\mathbf{z}} \left\|\hat{G}_{1n}(t,\mathbf{z}) - f(\mathbf{z})\pi(t,\mathbf{z})g_{\mathbf{Z}}(t)\right\| = O_p\left(h^r + \left(\frac{\log n}{nh^{d+1}}\right)^{1/2}\right),$$

$$\sup_{t,\mathbf{z}} \left\|\hat{G}_{2kn}(t,\mathbf{z}) - f(\mathbf{z})\pi(t,\mathbf{z})p_k(t,\mathbf{z})g_{\mathbf{Z}}(t)\right\| = O_p\left(h^r + \left(\frac{\log n}{nh^{d+1}}\right)^{1/2}\right).$$

This immediately implies the desired results by following the fact $\hat{\pi}(t,\mathbf{z}) = \hat{G}_{1n}(t,\mathbf{z})/\hat{G}_{0n}(t,\mathbf{z})$ and $\hat{p}_k(t,\mathbf{z}) = \hat{G}_{2kn}(t,\mathbf{z})/\hat{G}_{1n}(t,\mathbf{z})$.

LEMMA 2: *Under the regularity conditions C1, C6 and C7, we have*

$$\sup_t \left\|\frac{1}{n}\sum_{i=1}^n \mathbf{X}_i \hat{N}_{ik}^{IPW}(t) - \frac{1}{n}\sum_{i=1}^n \mathbf{X}_i N_{ik}^{IPW}(t)\right\| = o_p(1), \tag{1}$$

and

$$\sup_t \left\|n^{-1/2}\sum_{i=1}^n \mathbf{X}_i \hat{N}_{ik}^{IPW}(t) - n^{-1/2}\sum_{i=1}^n \mathbf{X}_i N_{ik}^{AIPW}(t)\right\| = o_p(1). \tag{2}$$

Proof. By simple algebra,

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^n \mathbf{X}_i \hat{N}_{ik}^{\text{IPW}}(t) - \frac{1}{n}\sum_{i=1}^n \mathbf{X}_i N_{ik}^{\text{IPW}}(t) \\
= \ &\frac{1}{n}\sum_{i=1}^n \mathbf{X}_i \sum_{j=1}^\infty \left(\frac{1}{\hat{\pi}_i^{(j)}} - \frac{1}{\pi_i^{(j)}}\right) A_i^{(j)}\delta_{ik}^{(j)} I(L_i < T_i^{(j)} \leqslant R_i \wedge t).
\end{aligned}$$

Equation (1) follows immediately from the regularity conditions C1, C7 and Lemma 1.

Now, we prove equation (2). Note that

$$
\begin{aligned}
& n^{-1/2}\sum_{i=1}^{n}\mathbf{X}_i\hat{N}_{ik}^{\text{IPW}}(t) \\
= & n^{-1/2}\sum_{i=1}^{n}\mathbf{X}_i\sum_{j=1}^{\infty}\frac{A_i^{(j)}}{\pi_i^{(j)}}\delta_{ik}^{(j)}I(L_i<T_i^{(j)}\leqslant R_i\wedge t) \\
& + n^{-1/2}\sum_{i=1}^{n}\mathbf{X}_i\sum_{j=1}^{\infty}\left(\frac{1}{\hat{\pi}_i^{(j)}}-\frac{1}{\pi_i^{(j)}}\right)A_i^{(j)}\delta_{ik}^{(j)}I(L_i<T_i^{(j)}\leqslant R_i\wedge t) \\
\doteq & I_1+\sqrt{n}I_2,
\end{aligned}
$$

where

$$
\begin{aligned}
I_2 = & -\frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i\sum_{j=1}^{\infty}\frac{\hat{\pi}_i^{(j)}-\pi_i^{(j)}}{\pi_i^{(j)}}\frac{A_i^{(j)}}{\pi_i^{(j)}}\delta_{ik}^{(j)}I(L_i<T_i^{(j)}\leqslant R_i\wedge t) \\
& +\frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i\sum_{j=1}^{\infty}\frac{(\hat{\pi}_i^{(j)}-\pi_i^{(j)})^2}{\hat{\pi}_i^{(j)}\pi_i^{(j)}}\frac{A_i^{(j)}}{\pi_i^{(j)}}\delta_{ik}^{(j)}I(L_i<T_i^{(j)}\leqslant R_i\wedge t) \\
\doteq & -I_{21}+I_{22}.
\end{aligned}
$$

It follows from Lemma 1 that

$$
\sup_{i,j}\left\|\hat{\pi}_i^{(j)}-\pi_i^{(j)}\right\| = O_p\left(h^r+\left(\frac{\log n}{nh^{d+1}}\right)^{1/2}\right).
$$

This together with the regularity conditions C6 and C7 imply that

$$
\sup_t\|\sqrt{n}I_{22}\| = O_p\left(\sqrt{n}h^{2r}+\sqrt{\log n}h^{r-(d+1)/2}+\frac{\log n}{\sqrt{n}h^{d+1}}\right)=o_p(1).
$$

By simple algebra,

$$
\begin{aligned}
I_{21} = & \frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i\sum_{j=1}^{\infty}\frac{\hat{\pi}_i^{(j)}-\pi_i^{(j)}}{\pi_i^{(j)}}\frac{A_i^{(j)}}{\pi_i^{(j)}}\delta_{ik}^{(j)}I(L_i<T_i^{(j)}\leqslant R_i\wedge t) \\
= & \frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i\sum_{j=1}^{\infty}\frac{\hat{G}_{1n}(T_i^{(j)},\mathbf{Z}_i)-\pi_i^{(j)}\hat{G}_{0n}(T_i^{(j)},\mathbf{Z}_i)}{\hat{G}_{0n}(T_i^{(j)},\mathbf{Z}_i)}\frac{A_i^{(j)}}{(\pi_i^{(j)})^2}\delta_{ik}^{(j)}I(L_i<T_i^{(j)}\leqslant R_i\wedge t) \\
= & \frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i\sum_{j=1}^{\infty}\{\hat{G}_{1n}(T_i^{(j)},\mathbf{Z}_i)-\pi_i^{(j)}\hat{G}_{0n}(T_i^{(j)},\mathbf{Z}_i)\}f(\mathbf{Z}_i)^{-1}g_{\mathbf{Z}_i}(T_i^{(j)})^{-1}
\end{aligned}
$$

$$
\begin{aligned}
& (\pi_i^{(j)})^{-2} A_i^{(j)} \delta_{ik}^{(j)} I(L_i < T_i^{(j)} \leqslant R_i \wedge t) \\
& + \frac{1}{n} \sum_{i=1}^{n} \mathbf{X}_i \sum_{j=1}^{\infty} \{\hat{G}_{1n}(T_i^{(j)}, \mathbf{Z}_i) - \pi_i^{(j)} \hat{G}_{0n}(T_i^{(j)}, \mathbf{Z}_i)\}\{\hat{G}_{0n}(T_i^{(j)}, \mathbf{Z}_i)^{-1} - f(\mathbf{Z}_i)^{-1} g_{\mathbf{Z}_i}(T_i^{(j)})^{-1}\} \\
& (\pi_i^{(j)})^{-2} A_i^{(j)} \delta_{ik}^{(j)} I(L_i < T_i^{(j)} \leqslant R_i \wedge t) \\
= & \frac{1}{n^2} \sum_{i=1}^{n} \sum_{m=1}^{n} \mathbf{X}_i \sum_{j=1}^{\infty} \mathbf{K}_h(\mathbf{Z}_{1m} - \mathbf{Z}_{1i}) I(\mathbf{Z}_{2m} = \mathbf{Z}_{2i}) \int K_h(u - T_i^{(j)})\{A_m(u) - \pi_i^{(j)}\} dN_{m\cdot}(u) \\
& \cdot (\pi_i^{(j)})^{-2} f(\mathbf{Z}_i)^{-1} g_{\mathbf{Z}_i}(T_i^{(j)})^{-1} A_i^{(j)} \delta_{ik}^{(j)} I(L_i < T_i^{(j)} \leqslant R_i \wedge t) \\
& - \frac{1}{n} \sum_{i=1}^{n} \mathbf{X}_i \sum_{j=1}^{\infty} \{\hat{G}_{1n}(T_i^{(j)}, \mathbf{Z}_i) - \pi_i^{(j)} \hat{G}_{0n}(T_i^{(j)}, \mathbf{Z}_i)\} \left\{ \frac{\hat{G}_{0n}(T_i^{(j)}, \mathbf{Z}_i) - f(\mathbf{Z}_i) g_{\mathbf{Z}_i}(T_i^{(j)})}{\hat{G}_{0n}(T_i^{(j)}, \mathbf{Z}_i) f(\mathbf{Z}_i) g_{\mathbf{Z}_i}(T_i^{(j)})} \right\} \\
& (\pi_i^{(j)})^{-2} A_i^{(j)} \delta_{ik}^{(j)} I(L_i < T_i^{(j)} \leqslant R_i \wedge t) \\
\doteq & I_{211} - I_{212}.
\end{aligned}
$$

It follows from Lemma 1, the regularity conditions C6 and C7 that $\sup_t \|\sqrt{n} I_{212}\| = o_p(1)$. To calculate $\sqrt{n} I_{211}$, define

$$
\begin{aligned}
h(\mathbf{S}_i, \mathbf{S}_m) = & \mathbf{X}_i \sum_{j=1}^{\infty} \mathbf{K}_h(\mathbf{Z}_{1m} - \mathbf{Z}_{1i}) I(\mathbf{Z}_{2m} = \mathbf{Z}_{2i}) \int K_h(u - T_i^{(j)})\{A_m(u) - \pi_i^{(j)}\} dN_{m\cdot}(u) \\
& (\pi_i^{(j)})^{-2} f(\mathbf{Z}_i)^{-1} g_{\mathbf{Z}_i}(T_i^{(j)})^{-1} A_i^{(j)} \delta_{ik}^{(j)} I(L_i < T_i^{(j)} \leqslant R_i \wedge t),
\end{aligned}
$$

and $H(\mathbf{S}_i, \mathbf{S}_m) = 2^{-1}\{h(\mathbf{S}_i, \mathbf{S}_m) + h(\mathbf{S}_m, \mathbf{S}_i)\}$, where $\mathbf{S}_i = \{\mathbf{Z}_i, T_i^{(j)}, A_i^{(j)}, j = 1, 2, \ldots\}$. Then

$$
I_{211} = \frac{1}{n^2} \sum_{i=1}^{n} H(\mathbf{S}_i, \mathbf{S}_i) + \frac{1}{n^2} \sum_{i=1}^{n} \sum_{m \neq i} H(\mathbf{S}_i, \mathbf{S}_m)
$$

According to the mean zero property of $H(\mathbf{S}_i, \mathbf{S}_i)$ and the strong law of large numbers, we have $\sup_t \|n^{-3/2} \sum_{i=1}^{n} H(\mathbf{S}_i, \mathbf{S}_i)\| = o(1)$. It suffices to consider the U-process $n^{-2} \sum_{i=1}^{n} \sum_{m \neq i} H(\mathbf{S}_i, \mathbf{S}_m)$ only. Following the uniform convergence law of U-process stated in Theorem 7 of Nolan and Pollard (1987),

$$
\sup_t \left\| \frac{1}{n^{3/2}} \sum_{i=1}^{n} \sum_{m \neq i} H(\mathbf{S}_i, \mathbf{S}_m) - \frac{2}{\sqrt{n}} \sum_{i=1}^{n} E[H(\mathbf{S}_i, \mathbf{S}_m) | \mathbf{S}_i] \right\| = o(1).
$$

Following the usual arguments in nonparametric regression, we can show that

$$E[h(\mathbf{S}_i, \mathbf{S}_m)|\mathbf{S}_i] = O(h^r),$$

$$E[h(\mathbf{S}_m, \mathbf{S}_i)|\mathbf{S}_i] = \mathbf{X}_i \sum_{j=1}^{\infty} \left(\frac{A_i^{(j)}}{\pi_i^{(j)}} - 1 \right) p_{ik}^{(j)} I(L_i < T_i^{(j)} \leqslant R_i \wedge t) + O(h^r),$$

and further the convergence is uniform for t. Hence, we can conclude that

$$\sup_t \left\| \sqrt{n} I_{211} - \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \mathbf{X}_i \sum_{j=1}^{\infty} \left(\frac{A_i^{(j)}}{\pi_i^{(j)}} - 1 \right) p_{ik}^{(j)} I(L_i < T_i^{(j)} \leqslant R_i \wedge t) \right\| = o_p(1)$$

and

$$\begin{aligned} & \sup_t \left\| n^{-1/2} \sum_{i=1}^{n} \mathbf{X}_i \hat{N}_{ik}^{\text{IPW}}(t) - (I_1 - \sqrt{n} I_{211}) \right\| \\ = \; & \sup_t \left\| n^{-1/2} \sum_{i=1}^{n} \mathbf{X}_i \hat{N}_{ik}^{\text{IPW}}(t) - n^{-1/2} \sum_{i=1}^{n} \mathbf{X}_i N_{ik}^{\text{AIPW}}(t) \right\| = o_p(1). \end{aligned}$$

LEMMA 3: *Under the regularity conditions C1, C6 and C7, we have*

$$\sup_t \left\| \frac{1}{n} \sum_{i=1}^{n} \mathbf{X}_i \hat{N}_{ik}^{EEP}(t) - \frac{1}{n} \sum_{i=1}^{n} \mathbf{X}_i N_{ik}^{EEP}(t) \right\| = o_p(1), \tag{3}$$

and

$$\sup_t \left\| n^{-1/2} \sum_{i=1}^{n} \mathbf{X}_i \hat{N}_{ik}^{EEP}(t) - n^{-1/2} \sum_{i=1}^{n} \mathbf{X}_i N_{ik}^{AIPW}(t) \right\| = o_p(1). \tag{4}$$

Proof. Equation (3) holds by following the fact

$$\begin{aligned} & \frac{1}{n} \sum_{i=1}^{n} \mathbf{X}_i \hat{N}_{ik}^{\text{EEP}}(t) - \frac{1}{n} \sum_{i=1}^{n} \mathbf{X}_i N_{ik}^{\text{EEP}}(t) \\ = \; & \frac{1}{n} \sum_{i=1}^{n} \mathbf{X}_i \sum_{j=1}^{\infty} [A_i^{(j)} \delta_{ik}^{(j)} + (1 - A_i^{(j)})(\hat{p}_{ik}^{(j)} - p_{ik}^{(j)})] I(L_i < T_i^{(j)} \leqslant R_i \wedge t), \end{aligned}$$

the boundedness of $\mathbf{X}_i$ and $N_{ik}(t)$, and Lemma 1.

To prove (4), note that

$$\begin{aligned}
& n^{-1/2}\sum_{i=1}^{n}\mathbf{X}_i\hat{N}_{ik}^{\text{EEP}}(t) \\
= \;& n^{-1/2}\sum_{i=1}^{n}\mathbf{X}_i\sum_{j=1}^{\infty}[A_i^{(j)}\delta_{ik}^{(j)}+(1-A_i^{(j)})\hat{p}_{ik}^{(j)}]I(L_i<T_i^{(j)}\leqslant R_i\wedge t) \\
= \;& n^{-1/2}\sum_{i=1}^{n}\mathbf{X}_i\sum_{j=1}^{\infty}[A_i^{(j)}\delta_{ik}^{(j)}+(1-A_i^{(j)})p_{ik}^{(j)}]I(L_i<T_i^{(j)}\leqslant R_i\wedge t) \\
& +n^{-1/2}\sum_{i=1}^{n}\mathbf{X}_i\sum_{j=1}^{\infty}(1-A_i^{(j)})(\hat{p}_{ik}^{(j)}-p_{ik}^{(j)})I(L_i<T_i^{(j)}\leqslant R_i\wedge t) \\
\doteq \;& J_1+\sqrt{n}J_2,
\end{aligned}$$

where

$$\begin{aligned}
J_2 \;=\; & \frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i\sum_{j=1}^{\infty}\frac{\hat{G}_{2kn}(T_i^{(j)},\mathbf{Z}_i)-p_{ik}^{(j)}\hat{G}_{1n}(T_i^{(j)},\mathbf{Z}_i)}{\hat{G}_{1n}(T_i^{(j)},\mathbf{Z}_i)}(1-A_i^{(j)})I(L_i<T_i^{(j)}\leqslant R_i\wedge t) \\
= \;& \frac{1}{n^2}\sum_{i=1}^{n}\sum_{m=1}^{n}\mathbf{X}_i\sum_{j=1}^{\infty}\mathbf{K}_h(\mathbf{Z}_{1m}-\mathbf{Z}_{1i})I(\mathbf{Z}_{2m}=\mathbf{Z}_{2i}) \\
& \cdot\int K_h(u-T_i^{(j)})A_m(u)\{\delta_{mk}(u)-p_{ik}^{(j)}\}dN_{m\cdot}(u) \\
& \cdot f(\mathbf{Z}_i)^{-1}(\pi_i^{(j)})^{-1}g_{\mathbf{Z}_i}(T_i^{(j)})^{-1}(1-A_i^{(j)})I(L_i<T_i^{(j)}\leqslant R_i\wedge t) \\
& +\frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i\sum_{j=1}^{\infty}\{\hat{G}_{2kn}(T_i^{(j)},\mathbf{Z}_i)-p_{ik}^{(j)}\hat{G}_{1n}(T_i^{(j)},\mathbf{Z}_i)\} \\
& \cdot\left[\hat{G}_{1n}(T_i^{(j)},\mathbf{Z}_i)^{-1}-f(\mathbf{Z}_i)^{-1}(\pi_i^{(j)})^{-1}g_{\mathbf{Z}_i}(T_i^{(j)})^{-1}\right](1-A_i^{(j)})I(L_i<T_i^{(j)}\leqslant R_i\wedge t) \\
\doteq \;& J_{21}+J_{22}.
\end{aligned}$$

Simple algebra shows that

$$\begin{aligned}
J_{22} \;=\; & -\frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i\sum_{j=1}^{\infty}\{\hat{G}_{2kn}(T_i^{(j)},\mathbf{Z}_i)-p_{ik}^{(j)}\hat{G}_{1n}(T_i^{(j)},\mathbf{Z}_i)\} \\
& \cdot\left[\frac{\hat{G}_{1n}(T_i^{(j)},\mathbf{Z}_i)-f(\mathbf{Z}_i)\pi_i^{(j)}g_{\mathbf{Z}_i}(T_i^{(j)})}{\hat{G}_{1n}(T_i^{(j)},\mathbf{Z}_i)f(\mathbf{Z}_i)\pi_i^{(j)}g_{\mathbf{Z}_i}(T_i^{(j)})}\right](1-A_i^{(j)})I(L_i<T_i^{(j)}\leqslant R_i\wedge t).
\end{aligned}$$

It follows from Lemma 1 that

$$\sup_{i,j} \|\hat{G}_{2kn}(T_i^{(j)}, \mathbf{Z}_i) - p_{ik}^{(j)} \hat{G}_{1n}(T_i^{(j)}, \mathbf{Z}_i)\| = O_p\left(h^r + \left(\frac{\log n}{nh^{d+1}}\right)^{1/2}\right),$$

and

$$\sup_{i,j} \|\hat{G}_{1n}(T_i^{(j)}, \mathbf{Z}_i) - f(\mathbf{Z}_i)\pi_i^{(j)} g_{\mathbf{Z}_i}(T_i^{(j)})\| = O_p\left(h^r + \left(\frac{\log n}{nh^{d+1}}\right)^{1/2}\right).$$

By the regularity condition C6, we have

$$\sup_t \|\sqrt{n} J_{22}\| = O_p\left(\sqrt{n}h^{2r} + \sqrt{\log n}\, h^{r-(d+1)/2} + \frac{\log n}{\sqrt{n}h^{d+1}}\right) = o_p(1).$$

To calculate $\sqrt{n} J_{21}$, define

$$\begin{aligned} & \bar{h}(\mathbf{S}_i, \mathbf{S}_m) \\ \doteq\ & \mathbf{X}_i \sum_{j=1}^{\infty} \mathbf{K}_h(\mathbf{Z}_{1m} - \mathbf{Z}_{1i}) I(\mathbf{Z}_{2m} = \mathbf{Z}_{2i}) \int K_h(u - T_i^{(j)}) A_m(u) \{\delta_{mk}(u) - p_{ik}^{(j)}\} dN_{m\cdot}(u) \\ & \cdot f(\mathbf{Z}_i)^{-1} (\pi_i^{(j)})^{-1} g_{\mathbf{Z}_i}(T_i^{(j)})^{-1} (1 - A_i^{(j)}) I(L_i < T_i^{(j)} \leqslant R_i \wedge t), \end{aligned}$$

and $\bar{H}(\mathbf{S}_i, \mathbf{S}_m) = 2^{-1}\{\bar{h}(\mathbf{S}_i, \mathbf{S}_m) + \bar{h}(\mathbf{S}_m, \mathbf{S}_i)\}$. Then

$$J_{21} = \frac{1}{n^2} \sum_{i=1}^{n} \bar{H}(\mathbf{S}_i, \mathbf{S}_i) + \frac{1}{n^2} \sum_{i=1}^{n} \sum_{m \neq i} \bar{H}(\mathbf{S}_i, \mathbf{S}_m).$$

According to the mean zero property of $\bar{H}(\mathbf{S}_i, \mathbf{S}_i)$ and the strong law of large numbers, we have $\sup_t \|n^{-3/2} \sum_{i=1}^n \bar{H}(\mathbf{S}_i, \mathbf{S}_i)\| = o(1)$. It suffices to consider the U-process $n^{-2} \sum_{i=1}^n \sum_{m \neq i}$ $\bar{H}(\mathbf{S}_i, \mathbf{S}_m)$ only. Following the uniform convergence law of U-process stated in Theorem 7 of Nolan and Pollard (1987),

$$\sup_t \left\| \frac{1}{n^{3/2}} \sum_{i=1}^{n} \sum_{m \neq i} \bar{H}(\mathbf{S}_i, \mathbf{S}_m) - \frac{2}{\sqrt{n}} \sum_{i=1}^{n} E[\bar{H}(\mathbf{S}_i, \mathbf{S}_m) | \mathbf{S}_i] \right\| = o(1).$$

Following the usual arguments in nonparametric regression, we can show that

$$E[\bar{h}(\mathbf{S}_i, \mathbf{S}_m) | \mathbf{S}_i] = O\left(h^r\right),$$

$$E[\bar{h}(\mathbf{S}_m, \mathbf{S}_i) | \mathbf{S}_i] = \mathbf{X}_i \sum_{j=1}^{\infty} \frac{1 - \pi_i^{(j)}}{\pi_i^{(j)}} A_i^{(j)} [\delta_{ik}^{(j)} - p_{ik}^{(j)}] I(L_i < T_i^{(j)} \leqslant R_i \wedge t) + O\left(h^r\right),$$

and further the convergence is uniform for t. Hence, we can conclude that

$$\sup_t \left\| \sqrt{n} J_{21} - \frac{1}{\sqrt{n}} \sum_{i=1}^n \mathbf{X}_i \sum_{j=1}^{\infty} \frac{1-\pi_i^{(j)}}{\pi_i^{(j)}} A_i^{(j)} [\delta_{ik}^{(j)} - p_{ik}^{(j)}] I(L_i < T_i^{(j)} \leqslant R_i \wedge t) \right\| = o_p(1)$$

and

$$\begin{aligned} & \sup_t \left\| n^{-1/2} \sum_{i=1}^n \mathbf{X}_i \hat{N}_{ik}^{\text{EEP}}(t) - (J_1 + \sqrt{n} J_{21}) \right\| \\ = & \sup_t \left\| n^{-1/2} \sum_{i=1}^n \mathbf{X}_i \hat{N}_{ik}^{\text{EEP}}(t) - n^{-1/2} \sum_{i=1}^n \mathbf{X}_i N_{ik}^{\text{AIPW}}(t) \right\| = o_p(1). \end{aligned}$$

Lemma 4: *For any sequence* $\{\tilde{\boldsymbol{\beta}}_k(\tau), \tau \in (0, U]\}$ *that satisfies* $\sup_{\tau \in (0,U]} \|\boldsymbol{\upsilon}_k\{\tilde{\boldsymbol{\beta}}_k(\tau)\} - \boldsymbol{\upsilon}_k\{\boldsymbol{\beta}_{0k}(\tau)\}\| \xrightarrow{p} 0$, *we have*

$$\begin{aligned} \sup_{\tau \in (0,U]} \Bigg\| & n^{-1/2} \sum_{i=1}^n \mathbf{X}_i N_{ik}^{AIPW}\{\mathbf{X}_i^\top \tilde{\boldsymbol{\beta}}_k(\tau)\} - n^{-1/2} \sum_{i=1}^n \mathbf{X}_i N_{ik}^{AIPW}\{\mathbf{X}_i^\top \boldsymbol{\beta}_{0k}(\tau)\} \\ & - n^{1/2}[\boldsymbol{\upsilon}_k\{\tilde{\boldsymbol{\beta}}_k(\tau)\} - \boldsymbol{\upsilon}_k\{\boldsymbol{\beta}_{0k}(\tau)\}] \Bigg\| \xrightarrow{p} 0. \end{aligned}$$

Proof. The proof of Lemma 4 follows by that of Lemma B.1 in Peng and Huang (2008), and we omit it here.

S2.2. Proofs of Theorems 1–2

Proof of Theorem 1: Simple algebra shows that

$$\begin{aligned} & \frac{1}{n} \sum_{i=1}^n \mathbf{X}_i \hat{N}_{ik}^{\text{L}}(\exp\{\mathbf{X}_i^\top \hat{\boldsymbol{\beta}}_k^{\text{L}}(\tau_j)\}) - E[\mathbf{X}_i N_{ik}(\exp\{\mathbf{X}_i^\top \boldsymbol{\beta}_{0k}(\tau_j)\})] \\ = & \frac{1}{n} \sum_{i=1}^n \mathbf{X}_i \hat{N}_{ik}^{\text{L}}(\exp\{\mathbf{X}_i^\top \hat{\boldsymbol{\beta}}_k^{\text{L}}(\tau_j)\}) - \frac{1}{n} \sum_{i=1}^n \mathbf{X}_i N_{ik}^{\text{L}}(\exp\{\mathbf{X}_i^\top \hat{\boldsymbol{\beta}}_k^{\text{L}}(\tau_j)\}) \\ & + \frac{1}{n} \sum_{i=1}^n \mathbf{X}_i N_{ik}^{\text{L}}(\exp\{\mathbf{X}_i^\top \hat{\boldsymbol{\beta}}_k^{\text{L}}(\tau_j)\}) - \boldsymbol{\upsilon}_k\{\hat{\boldsymbol{\beta}}_k^{\text{L}}(\tau_j)\} + \boldsymbol{\upsilon}_k\{\hat{\boldsymbol{\beta}}_k^{\text{L}}(\tau_j)\} - \boldsymbol{\upsilon}_k\{\boldsymbol{\beta}_{0k}(\tau_j)\} \quad (5) \end{aligned}$$

for L = IPW or EEP, and

$$\frac{1}{n} \sum_{i=1}^n \int_0^{\tau_j} \mathbf{X}_i Y_i(\exp\{\mathbf{X}_i^\top \hat{\boldsymbol{\beta}}_k^{\text{L}}(u)\}) g(u) du - E\left[\int_0^{\tau_j} \mathbf{X}_i Y_i(\exp\{\mathbf{X}_i^\top \boldsymbol{\beta}_{0k}(u)\}) g(u) du\right]$$

$$= \frac{1}{n}\sum_{i=1}^{n}\int_0^{\tau_j}\mathbf{X}_i Y_i(\exp\{\mathbf{X}_i^\top\hat{\boldsymbol\beta}_k^{\text{L}}(u)\})g(u)du - \int_0^{\tau_j}\tilde{\boldsymbol\upsilon}\{\hat{\boldsymbol\beta}_k^{\text{L}}(u)\}g(u)du$$
$$+\int_0^{\tau_j}[\tilde{\boldsymbol\upsilon}\{\hat{\boldsymbol\beta}_k^{\text{L}}(u)\} - \tilde{\boldsymbol\upsilon}\{\boldsymbol\beta_{0k}(u)\}]g(u)du. \quad (6)$$

According to the estimation procedure proposed in Section 2 and the definition of a generalized solution,

$$\frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i\hat{N}_{ik}^{\text{L}}(\exp\{\mathbf{X}_i^\top\hat{\boldsymbol\beta}_k^{\text{L}}(\tau_j)\}) = \frac{1}{n}\sum_{i=1}^{n}\int_0^{\tau_j}\mathbf{X}_i Y_i(\exp\{\mathbf{X}_i^\top\hat{\boldsymbol\beta}_k^{\text{L}}(u)\})g(u)du + O_p\left(\frac{1}{n}\right). \quad (7)$$

On the other hand,

$$E[\mathbf{X}_i N_{ik}(\exp\{\mathbf{X}_i^\top\boldsymbol\beta_{0k}(\tau_j)\})] - E\left[\int_0^{\tau_j}\mathbf{X}_i Y_i(\exp\{\mathbf{X}_i^\top\boldsymbol\beta_{0k}(u)\})g(u)du\right] = 0. \quad (8)$$

It is seen from the foregoing four equations (5)–(8) that

$$\begin{aligned}
&\boldsymbol\upsilon_k\{\hat{\boldsymbol\beta}_k^{\text{L}}(\tau_j)\} - \boldsymbol\upsilon_k\{\boldsymbol\beta_{0k}(\tau_j)\}\\
=\;& -\left\{\frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i\hat{N}_{ik}^{\text{L}}(\exp\{\mathbf{X}_i^\top\hat{\boldsymbol\beta}_k^{\text{L}}(\tau_j)\}) - \frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i N_{ik}^{\text{L}}(\exp\{\mathbf{X}_i^\top\hat{\boldsymbol\beta}_k^{\text{L}}(\tau_j)\})\right\}\\
& -\left\{\frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i N_{ik}^{\text{L}}(\exp\{\mathbf{X}_i^\top\hat{\boldsymbol\beta}_k^{\text{L}}(\tau_j)\}) - \boldsymbol\upsilon_k\{\hat{\boldsymbol\beta}_k^{\text{L}}(\tau_j)\}\right\}\\
& +\frac{1}{n}\sum_{i=1}^{n}\int_0^{\tau_j}\mathbf{X}_i Y_i(\exp\{\mathbf{X}_i^\top\hat{\boldsymbol\beta}_k^{\text{L}}(u)\})g(u)du - \int_0^{\tau_j}\tilde{\boldsymbol\upsilon}\{\hat{\boldsymbol\beta}_k^{\text{L}}(u)\}g(u)du\\
& +\sum_{l=1}^{j}\int_{\tau_{l-1}}^{\tau_l}[\tilde{\boldsymbol\upsilon}\{\hat{\boldsymbol\beta}_k^{\text{L}}(u)\} - \tilde{\boldsymbol\upsilon}\{\boldsymbol\beta_{0k}(u)\}]g(u)du + O_p\left(\frac{1}{n}\right).
\end{aligned}$$

For any given $C_0 > 0$ and $\delta > 0$, according to Lemmas 2 and 3, there exists a constant N_0 such that $\Pr\left(\sup_j \|n^{-1}\sum_{i=1}^n \mathbf{X}_i\hat{N}_{ik}^{\text{L}}(\exp\{\mathbf{X}_i^\top\hat{\boldsymbol\beta}_k^{\text{L}}(\tau_j)\}) - n^{-1}\sum_{i=1}^n \mathbf{X}_i N_{ik}^{\text{L}}(\exp\{\mathbf{X}_i^\top\hat{\boldsymbol\beta}_k^{\text{L}}(\tau_j)\})\| < C_0\right) > 1-\delta$ when $n \geqslant N_0$. Consider $\mathcal{G}_1 = \{\mathbf{X}_i N_{ik}^{\text{L}}\{\exp(\mathbf{X}_i^\top\mathbf{b})\} : \mathbf{b}\in R^p\}$ and $\mathcal{G}_2 = \{\mathbf{X}_i Y_i\{\exp(\mathbf{X}_i^\top\mathbf{b})\} : \mathbf{b}\in R^p\}$. Both $\mathcal{G}_1$ and $\mathcal{G}_2$ are Glivenko–Cantelli class (van der Vaart and Wellner, 1996) under the regularity conditions C1 and C7 because the class of uniformly bounded monotone functions is Donsker class and thus Glivenko-Cantelli class. It follows from the Glivenko-Cantelli theorem that $\sup_{\mathbf{b}}\|n^{-1}\sum_{i=1}^n \mathbf{X}_i N_{ik}^{\text{L}}\{\exp(\mathbf{X}_i^\top\mathbf{b})\} - \boldsymbol\upsilon_k(\mathbf{b})\| \to 0$ almost surely, and

$\sup_{\mathbf{b}} \|n^{-1}\sum_{i=1}^{n} \mathbf{X}_i Y_i \{\exp(\mathbf{X}_i^\top \mathbf{b})\} - \tilde{\boldsymbol{\upsilon}}(\mathbf{b})\| \to 0$ almost surely. Consequently, for any given $C_1 > 0$,

$$\sup_j \Bigg\| -\frac{1}{n}\sum_{i=1}^{n} \mathbf{X}_i N_{ik}^{\text{L}}(\exp\{\mathbf{X}_i^\top \hat{\boldsymbol{\beta}}_k^{\text{L}}(\tau_j)\}) + \boldsymbol{\upsilon}_k\{\hat{\boldsymbol{\beta}}_k^{\text{L}}(\tau_j)\}$$
$$+\frac{1}{n}\sum_{i=1}^{n}\int_0^{\tau_j} \mathbf{X}_i Y_i(\exp\{\mathbf{X}_i^\top \hat{\boldsymbol{\beta}}_k^{\text{L}}(u)\}) g(u)du - \int_0^{\tau_j} \tilde{\boldsymbol{\upsilon}}\{\hat{\boldsymbol{\beta}}_k^{\text{L}}(u)\} g(u)du \Bigg\| < C_1$$

almost surely when n is sufficiently large.

Consider the case $\sup_j \|n^{-1}\sum_{i=1}^{n} \mathbf{X}_i \hat{N}_{ik}^{\text{L}}(\exp\{\mathbf{X}_i^\top \hat{\boldsymbol{\beta}}_k^{\text{L}}(\tau_j)\}) - n^{-1}\sum_{i=1}^{n} \mathbf{X}_i N_{ik}^{\text{L}}(\exp\{\mathbf{X}_i^\top \hat{\boldsymbol{\beta}}_k^{\text{L}}(\tau_j)\})\| < C_0$. Define $\epsilon_0 = C_3 a_n$, $\epsilon_l = (1 + C_4 a_n)^{l-1}(C_0 + C_1 + \epsilon_0 C_4 a_n + C_2 n^{-1} + C_3 a_n)$ for $l = 1, \ldots, L(n)$, where $a_n = \|\mathcal{S}_{L(n)}\|$, $C_i (i = 2, 3, 4)$ are some positive constants. Under conditions C1–C3, following the same arguments as in Peng and Huang (2008), we can show that $\sup_{u\in[u_{l-1},u_l]} \|\boldsymbol{\upsilon}_k\{\hat{\boldsymbol{\beta}}_k^{\text{L}}(u)\} - \boldsymbol{\upsilon}_k\{\boldsymbol{\beta}_{0k}(u)\}\| \leqslant \epsilon_l$. Suppose $\mathcal{S}_{L(n)}$ is an equally spaced grid, then $L(n) = U/a_n$. Given $\lim_{n\to\infty} a_n = 0$, we have $\lim_{n\to\infty}(1 + C_4 a_n)^{L(n)-1} = \exp(C_4 U)$. Because $a_n \to 0$ and C_0, C_1 can be arbitrarily small as n increases, we have

$$\sup_{u\in(0,U]} \|\boldsymbol{\upsilon}_k\{\hat{\boldsymbol{\beta}}_k^{\text{L}}(u)\} - \boldsymbol{\upsilon}_k\{\boldsymbol{\beta}_{0k}(u)\}\| \xrightarrow{p} 0. \tag{9}$$

Similar to Peng and Huang (2008) and Sun et al. (2016), it follows from the regularity condition C3 that $\boldsymbol{\upsilon}_k(\cdot)$ is a one-to-one map from $\mathcal{B}_k(d_0)$ to $\mathcal{A}_k(d_0) \doteq \{\boldsymbol{\upsilon}_k(\mathbf{b}) : \mathbf{b} \in \mathcal{B}(d_0)\}$. There exists an inverse function $\boldsymbol{\kappa}_k(\cdot)$ from $\mathcal{A}_k(d_0)$ to $\mathcal{B}(d_0)$ such that $\boldsymbol{\kappa}_k\{\boldsymbol{\upsilon}_k(\mathbf{b})\} = \mathbf{b}$. Using Taylor expansion of $\boldsymbol{\kappa}_k\{\boldsymbol{\upsilon}_k\{\hat{\boldsymbol{\beta}}_k^{\text{L}}(u)\}\}$ around $\boldsymbol{\upsilon}_k\{\boldsymbol{\beta}_{0k}(u)\}$, $\sup_{u\in[v,U]} \|\hat{\boldsymbol{\beta}}_k^{\text{L}}(u) - \boldsymbol{\beta}_{0k}(u)\| \xrightarrow{p} 0$ follows under the regularity condition C5.

Proof of Theorem 2: From (9) in the proofs of Theorem 1 and Lemma 4, we have

$$\sup_{\tau\in(0,U]} \Bigg\| n^{-1/2}\sum_{i=1}^{n} \mathbf{X}_i N_{ik}^{\text{AIPW}}\{\mathbf{X}_i^\top \hat{\boldsymbol{\beta}}_k^{\text{L}}(\tau)\} - n^{-1/2}\sum_{i=1}^{n} \mathbf{X}_i N_{ik}^{\text{AIPW}}\{\mathbf{X}_i^\top \boldsymbol{\beta}_{0k}(\tau)\}$$
$$- n^{1/2}[\boldsymbol{\upsilon}_k\{\hat{\boldsymbol{\beta}}_k^{\text{L}}(\tau)\} - \boldsymbol{\upsilon}_k\{\boldsymbol{\beta}_{0k}(\tau)\}] \Bigg\| \xrightarrow{p} 0. \tag{10}$$

Similarly, we can get

$$\sup_{\tau\in(0,U]}\Bigg\| n^{-1/2}\sum_{i=1}^{n}\mathbf{X}_iY_i(\exp\{\mathbf{X}_i^\top\hat{\boldsymbol\beta}_k^{\text{L}}(\tau)\}) - n^{-1/2}\sum_{i=1}^{n}\mathbf{X}_iY_i(\exp\{\mathbf{X}_i^\top\boldsymbol\beta_{0k}(\tau)\}) - n^{1/2}[\tilde{\boldsymbol\upsilon}\{\hat{\boldsymbol\beta}_k^{\text{L}}(\tau)\} - \tilde{\boldsymbol\upsilon}\{\boldsymbol\beta_{0k}(\tau)\}]\Bigg\| \xrightarrow{p} 0. \quad (11)$$

It follows from Lemmas 2 and 3 that

$$\sup_{\tau\in(0,U]}\Bigg\| n^{-1/2}\sum_{i=1}^{n}\mathbf{X}_i\hat{N}_{ik}^{\text{L}}\{\mathbf{X}_i^\top\hat{\boldsymbol\beta}_k^{\text{L}}(\tau)\} - n^{-1/2}\sum_{i=1}^{n}\mathbf{X}_iN_{ik}^{\text{AIPW}}\{\mathbf{X}_i^\top\hat{\boldsymbol\beta}_k^{\text{L}}(\tau)\}\Bigg\| \xrightarrow{p} 0 \quad (12)$$

for L = IPW and EEP. Given $n^{1/2}\|\mathcal{S}_{L(n)}\| \to 0$, we can show that

$$n^{-1/2}\sum_{i=1}^{n}\mathbf{X}_i\left\{\hat{N}_{ik}^{\text{L}}\{\mathbf{X}_i^\top\hat{\boldsymbol\beta}_k^{\text{L}}(\tau)\} - \int_0^\tau Y_i(\exp\{\mathbf{X}_i^\top\hat{\boldsymbol\beta}_k^{\text{L}}(u)\})g(u)du\right\} = o_{(0,U]}(1) \quad (13)$$

for L = IPW or EEP, where $o_{(0,U]}(1)$ denotes a term that converges uniformly to 0 in $u \in (0,U]$. Then it follows from (10)–(13) that

$$\begin{aligned}
&-\left\{n^{-1/2}\sum_{i=1}^{n}\mathbf{X}_i\left\{N_{ik}^{\text{AIPW}}\{\mathbf{X}_i^\top\boldsymbol\beta_{0k}(\tau)\} - \int_0^\tau Y_i(\exp\{\mathbf{X}_i^\top\boldsymbol\beta_{0k}(u)\})g(u)du\right\}\right\}\\
&= n^{1/2}[\boldsymbol\upsilon_k\{\hat{\boldsymbol\beta}_k^{\text{L}}(\tau)\} - \boldsymbol\upsilon_k\{\boldsymbol\beta_{0k}(\tau)\}] - \int_0^\tau n^{1/2}[\tilde{\boldsymbol\upsilon}\{\hat{\boldsymbol\beta}_k^{\text{L}}(u)\} - \tilde{\boldsymbol\upsilon}\{\boldsymbol\beta_{0k}(u)\}]g(u)du + o_{(0,U]}(1)\\
&= n^{1/2}[\boldsymbol\upsilon_k\{\hat{\boldsymbol\beta}_k^{\text{L}}(\tau)\} - \boldsymbol\upsilon_k\{\boldsymbol\beta_{0k}(\tau)\}]\\
&\quad - \int_0^\tau \mathbf{J}\{\boldsymbol\beta_{0k}(u)\}\mathbf{B}_k\{\boldsymbol\beta_{0k}(u)\}^{-1}n^{1/2}[\boldsymbol\upsilon_k\{\hat{\boldsymbol\beta}_k^{\text{L}}(u)\} - \boldsymbol\upsilon_k\{\boldsymbol\beta_{0k}(u)\}]g(u)du + o_{(0,U]}(1).
\end{aligned}$$

Viewing the above equation as a stochastic differential equation for $n^{1/2}[\boldsymbol\upsilon_k\{\hat{\boldsymbol\beta}_k^{\text{L}}(u)\} - \boldsymbol\upsilon_k\{\boldsymbol\beta_{0k}(u)\}]$, and using product integration theory (Andersen et al., 1993), we get

$$n^{1/2}[\boldsymbol\upsilon_k\{\hat{\boldsymbol\beta}_k^{\text{L}}(\tau)\} - \boldsymbol\upsilon_k\{\boldsymbol\beta_{0k}(\tau)\}] = \boldsymbol\phi\{-n^{1/2}\mathbf{S}_n\{\boldsymbol\beta_{0k}(\tau)\}\} + o_{(0,U]}(1),$$

where $\mathbf{S}_n\{\boldsymbol\beta_{0k}(\tau)\} = n^{-1}\sum_{i=1}^{n}\mathbf{X}_i\left\{N_{ik}^{\text{AIPW}}\{\mathbf{X}_i^\top\boldsymbol\beta_{0k}(\tau)\} - \int_0^\tau Y_i(\exp\{\mathbf{X}_i^\top\boldsymbol\beta_{0k}(u)\})g(u)du\right\}$, $\boldsymbol\phi(\mathbf{w})(u) \doteq \int_0^u \mathcal{I}(s,u)d\mathbf{w}(s)$ is a linear operator on $\mathcal{F} = \{\mathbf{w} : [0,U] \to R^p, \mathbf{w}$ is left-continuous with right limit, $\mathbf{w}(0) = 0\}$, and

$$\mathcal{I}(s,t) = \prod_{u\in(s,t]} [\mathbf{I}_p + \mathbf{J}\{\boldsymbol{\beta}_{0k}(u)\}\mathbf{B}_k\{\boldsymbol{\beta}_{0k}(u)\}^{-1}g(u)du].$$

Consider $\mathcal{G}_3 = \left\{\mathbf{X}_i[N_{ik}^{\text{AIPW}}(\mathbf{X}_i^\top\boldsymbol{\beta}_{0k}(\tau)) - \int_0^\tau Y_i(\exp\{\mathbf{X}_i^\top\boldsymbol{\beta}_{0k}(u)\})g(u)du], \tau \in [v, \tau_U]\right\}$. Under the regularity conditions C1 and C7, $\mathcal{G}_3$ is a Donsker class (van der Vaart and Wellner, 1996). By the Donsker theorem and the linear property of $\boldsymbol{\phi}$, $\boldsymbol{\phi}\{-n^{1/2}\mathbf{S}_n\{\boldsymbol{\beta}_{0k}(\tau)\}\}$ converges weakly to a mean zero Gaussian process. It follows from Taylor expansions and the uniform consistency of $\hat{\boldsymbol{\beta}}_k^{\text{L}}(u)$ stated in Theorem 1 that

$$n^{1/2}\{\hat{\boldsymbol{\beta}}_k^{\text{L}}(\tau) - \boldsymbol{\beta}_{0k}(\tau)\} = n^{-1/2}\sum_{i=1}^n \boldsymbol{\eta}_{ik}(u) + o_{(v,U]}(1),$$

where $\boldsymbol{\eta}_{ik}(u) = \mathbf{B}_k\{\boldsymbol{\beta}_{0k}(u)\}^{-1}\boldsymbol{\phi}(\boldsymbol{\xi}_{ik})$ with

$$\boldsymbol{\xi}_{ik}(\tau) = \mathbf{X}_i\left\{N_{ik}^{\text{AIPW}}\{\mathbf{X}_i^\top\boldsymbol{\beta}_{0k}(\tau)\} - \int_0^\tau Y_i(\exp\{\mathbf{X}_i^\top\boldsymbol{\beta}_{0k}(u)\})g(u)du\right\}.$$

This implies $n^{1/2}\{\hat{\boldsymbol{\beta}}_k^{\text{L}}(\tau) - \boldsymbol{\beta}_0(\tau)\}$ converges weakly to a mean zero Gaussian process with the covariance matrix $\boldsymbol{\Sigma}_k(s,t) = E[\boldsymbol{\eta}_{ik}(s)\boldsymbol{\eta}_{ik}(t)^\top]$.

S3. Additional Simulation Results

We present the additional simulation results in this Section. Figure S1 shows the results for type-2 event coefficients in Case 2.

[Figure 1 about here.]

The coefficients estimates for type-1 event and type-2 event in Case 1 are depicted in Figures S2 and S3 respectively.

[Figure 2 about here.]

[Figure 3 about here.]

Figure S4 presents the comparison of type-2 event coefficient estimates in Case 2 with different values of h.

[Figure 4 about here.]

In addition, we present the four estimators' empirical standard derivations under four different situations in Figures S5–S8.

[Figure 5 about here.]

[Figure 6 about here.]

[Figure 7 about here.]

[Figure 8 about here.]

S4. Sensitivity Analysis of the MAR assumption in CFFPR Data

In this Section, we investigate the sensitivity analysis of the MAR assumption in CFFPR data. When calculating the IPW or EEP estimator in the CFFPR data, the key task is to estimate the conditional probability of event type being observed or the missing event type being a specific type given $\mathbf{Z}_i$ and $T_i^{(j)}$. Different choice of $\mathbf{Z}_i$ may lead to different results. In the manuscript, we set $\mathbf{Z}_i$ =(*Sex*, *MI*, *NewScreen*, *FamilyHis*)$^\top$. Here we choose $\mathbf{Z}_i$ =(*NewScreen*, *FamilyHis*)$^\top$. The corresponding results for nonmucoid and mucoid event types are depicted in Figure S9 and Figure S10 respectively. We find that the results are very similar, which shows the robustness of our method.

[Figure 9 about here.]

[Figure 10 about here.]

S5. Sample Code

We now provide a link to a webpage where readers can access our sample code for implementing the proposed method; see http://web1.sph.emory.edu/users/lpeng/Rpackage.html

References

Andersen, P., Borgan, O., Gill, R., and Keiding, N. (1993). *Statistical Models Based on Counting Processes*. Springer, New York.

Nolan, D. and Pollard, D. (1987). U-processes: rates of convergence. *The Annals of Statistics* **15,** 780–799.

Peng, L. and Huang, Y. (2008). Survival analysis with quantile regression models. *Journal of the American Statistical Association* **103,** 637–649.

Sun, X., Peng, L., Huang, Y., and Lai, H. (2016). Generalizing quantile regression for counting processes with applications to recurrent events. *Journal of the American Statistical Association* **111,** 145–156.

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

Received April 2017. *Revised* ? ?.

Accepted ? ?.

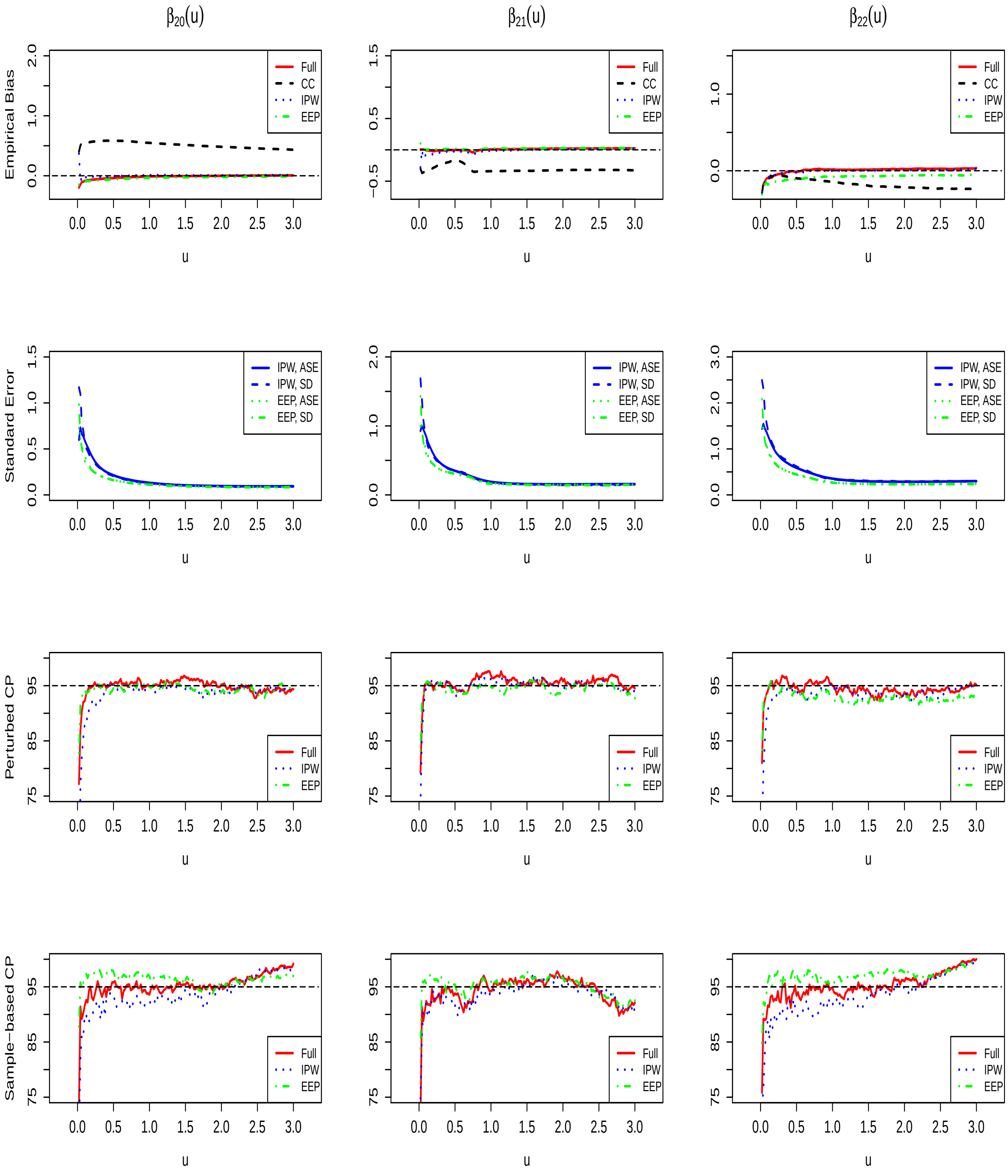

Figure S1. The simulation results for type-2 event coefficients in Case 2. IPW, the inverse probability weighting estimator; EEP, the estimating equation projection estimator; CC, the complete-case estimator; Full, the full data estimator. SD, the empirical standard derivation. ASE, the average standard error. CP, the coverage probability.

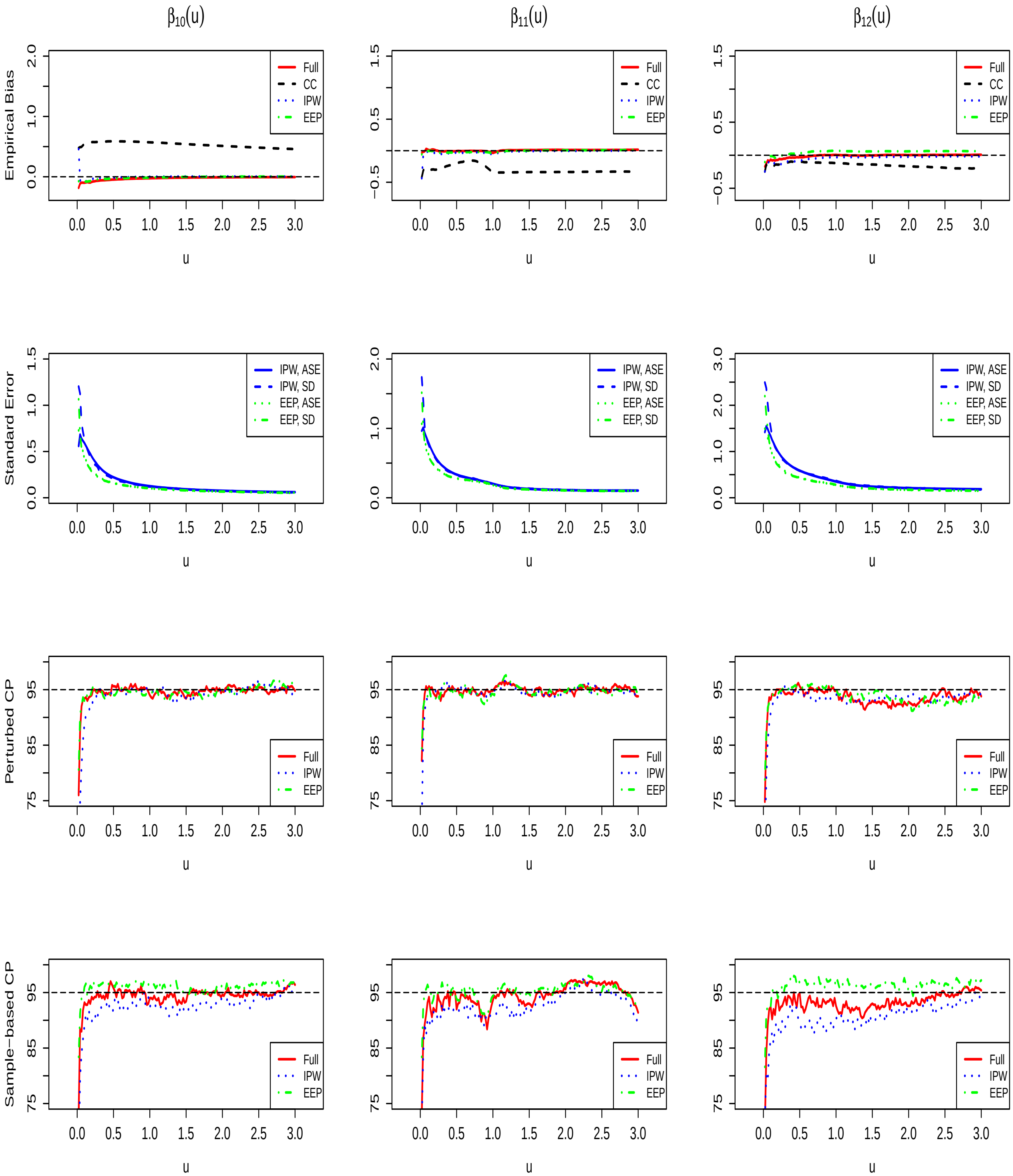

Figure S2. The simulation results for type-1 event coefficients in Case 1. IPW, the inverse probability weighting estimator; EEP, the estimating equation projection estimator; CC, the complete-case estimator; Full, the full data estimator. SD, the empirical standard derivation. ASE, the average standard error. CP, the coverage probability.

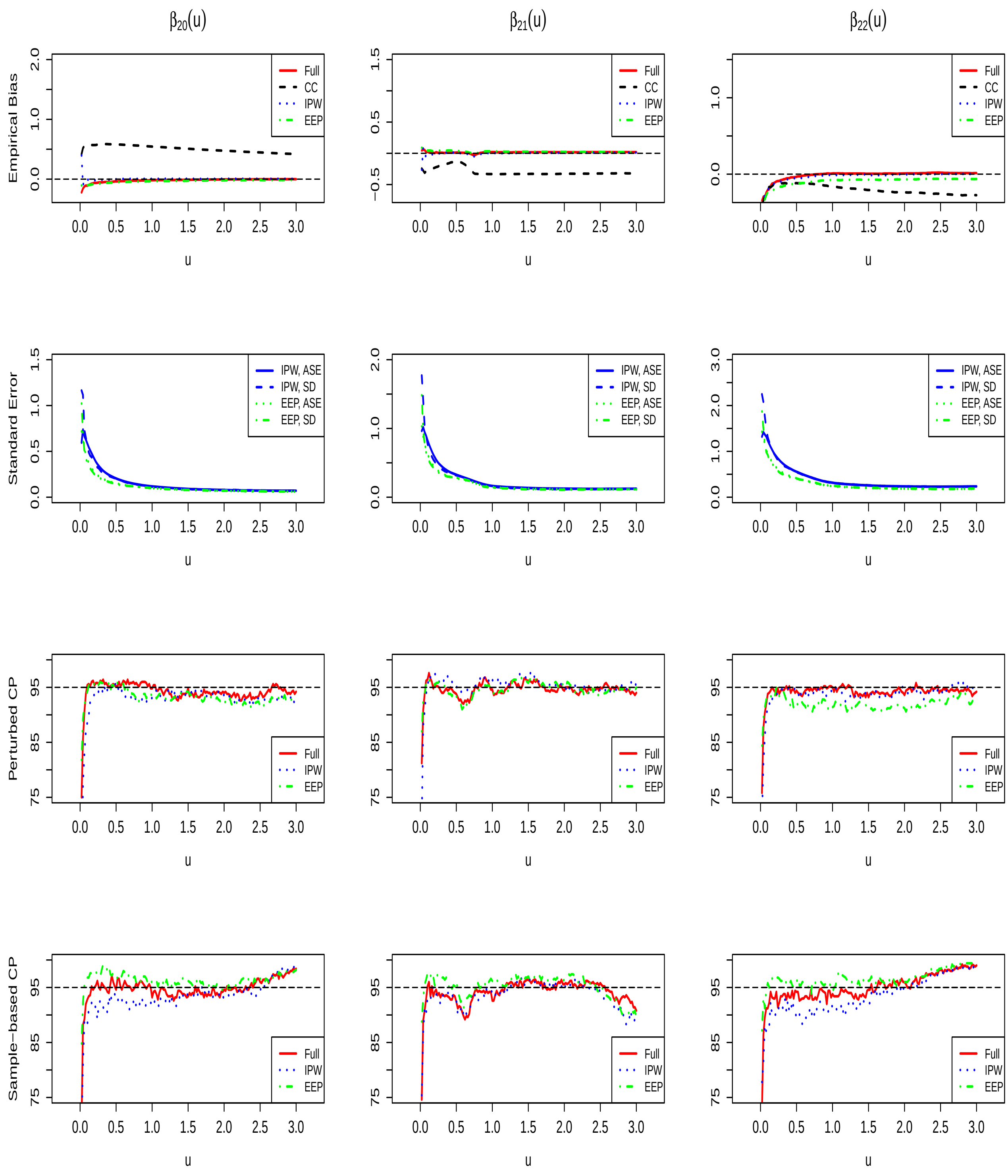

Figure S3. The simulation results for type-2 event coefficients in Case 1. IPW, the inverse probability weighting estimator; EEP, the estimating equation projection estimator; CC, the complete-case estimator; Full, the full data estimator. SD, the empirical standard derivation. ASE, the average standard error. CP, the coverage probability.

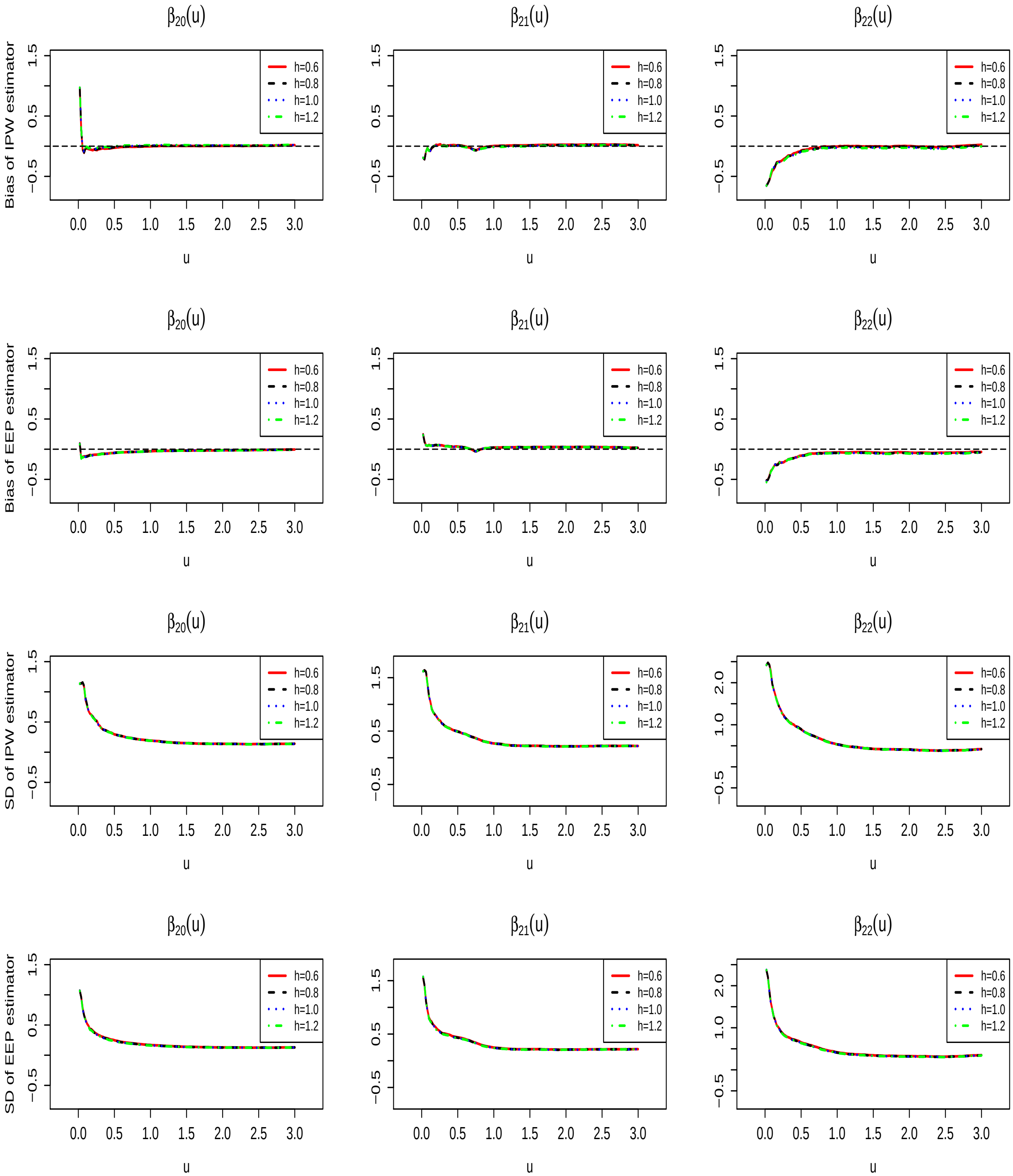

Figure S4. The comparison of type-2 event coefficient estimates in Case 2 with different values of h. IPW, the inverse probability weighting estimator; EEP, the estimating equation projection estimator. SD, the empirical standard derivation.

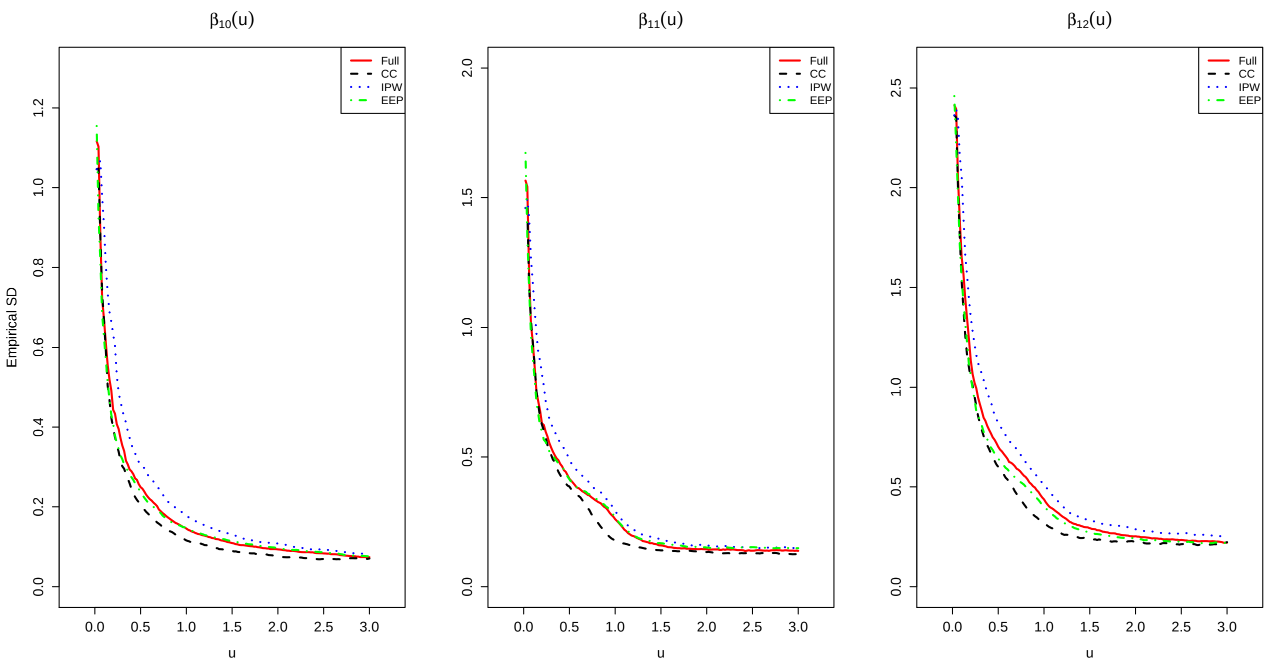

Figure S5. The empirical standard deviations for type-1 event coefficients in Case 1. IPW, the inverse probability weighting estimator; EEP, the estimating equation projection estimator; CC, the complete-case estimator; Full, the full data estimator. SD, the empirical standard derivation.

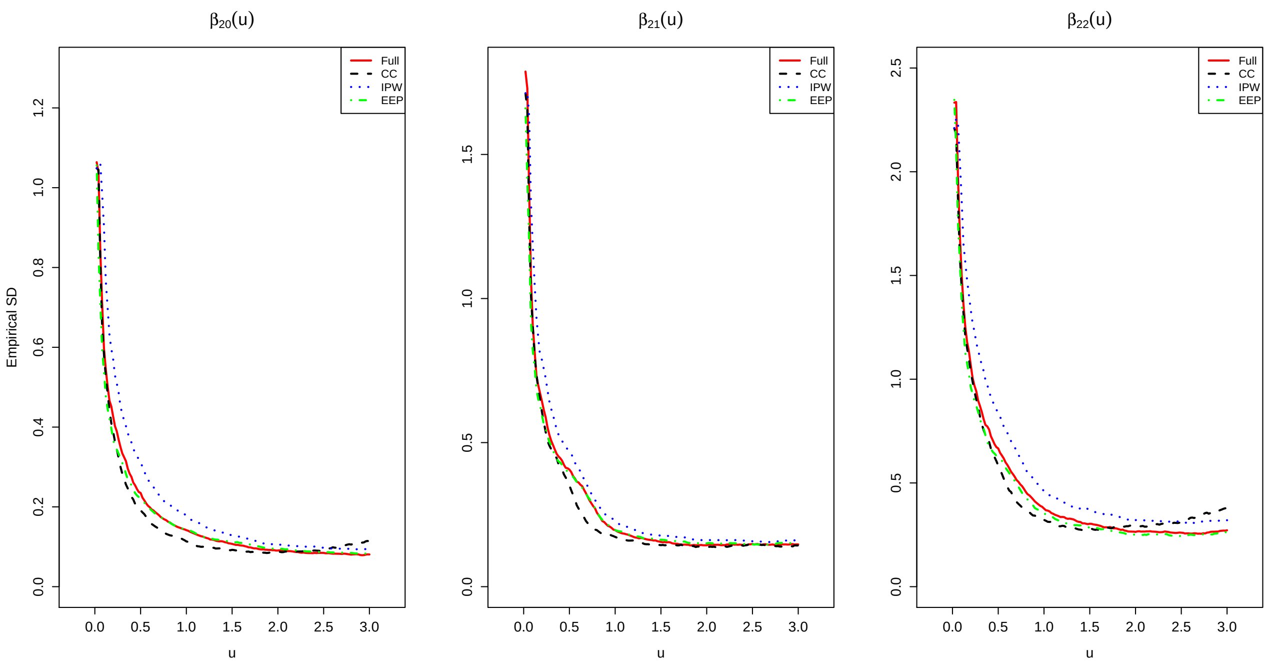

Figure S6. The empirical standard deviations for type-2 event coefficients in Case 1. IPW, the inverse probability weighting estimator; EEP, the estimating equation projection estimator; CC, the complete-case estimator; Full, the full data estimator. SD, the empirical standard derivation.

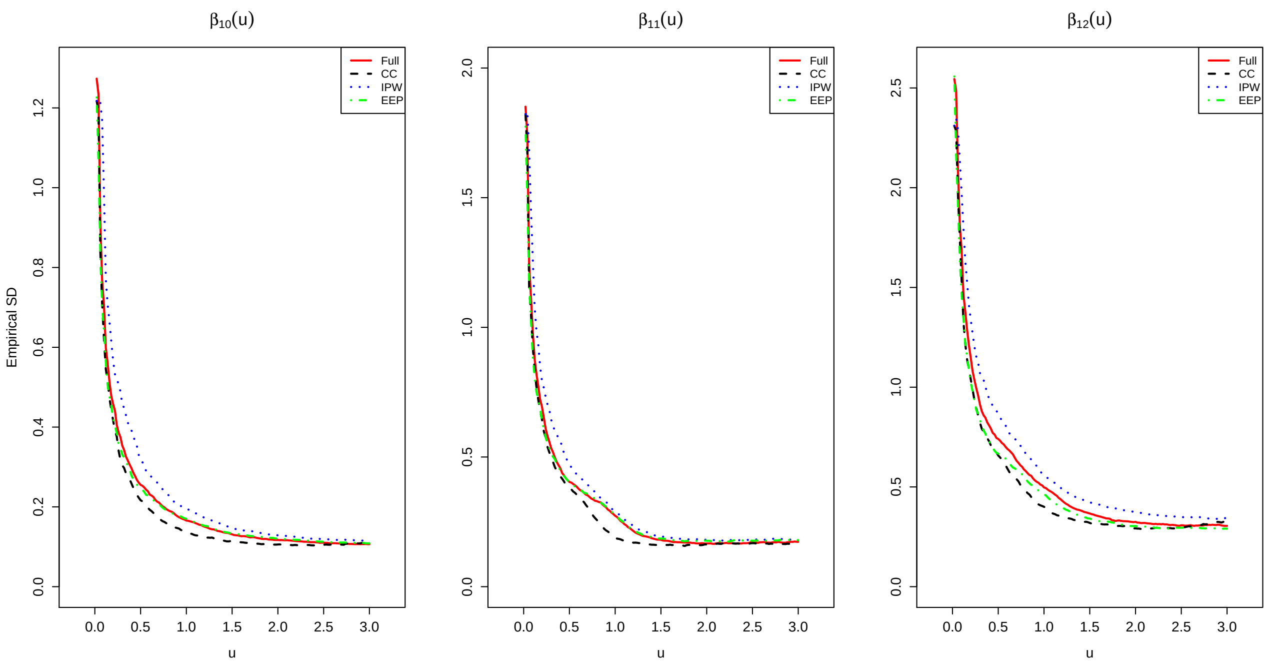

Figure S7. The empirical standard deviations for type-1 event coefficients in Case 2. IPW, the inverse probability weighting estimator; EEP, the estimating equation projection estimator; CC, the complete-case estimator; Full, the full data estimator. SD, the empirical standard derivation.

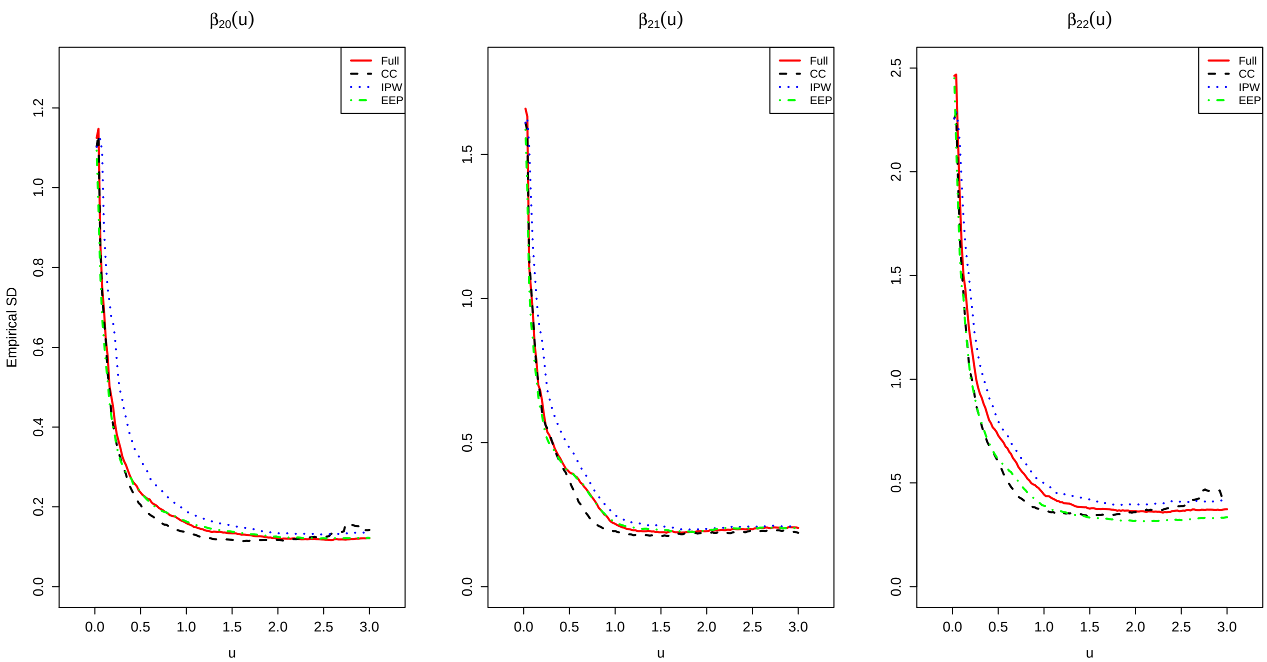

Figure S8. The empirical standard deviations for type-2 event coefficients in Case 2. IPW, the inverse probability weighting estimator; EEP, the estimating equation projection estimator; CC, the complete-case estimator; Full, the full data estimator. SD, the empirical standard derivation.

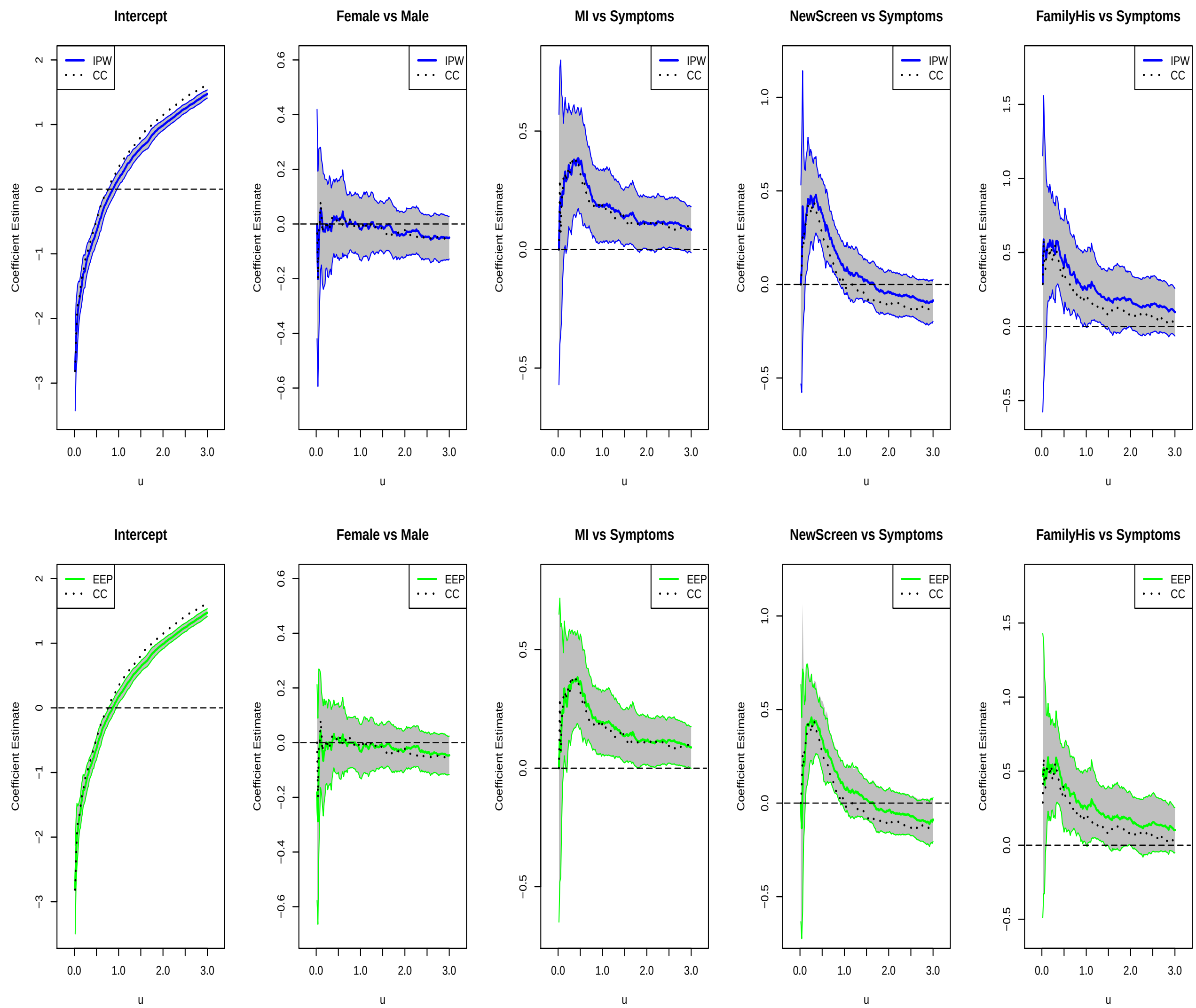

Figure S9. CFFPR data example: the proposed IPW coefficient estimates (blue solid lines) and their corresponding 95% pointwise confidence intervals, the proposed EEP coefficient estimates (green solid lines) and their corresponding 95% pointwise confidence intervals, along with the complete-case (CC) coefficient estimates (black dotted lines) for nonmucoid PA infection.

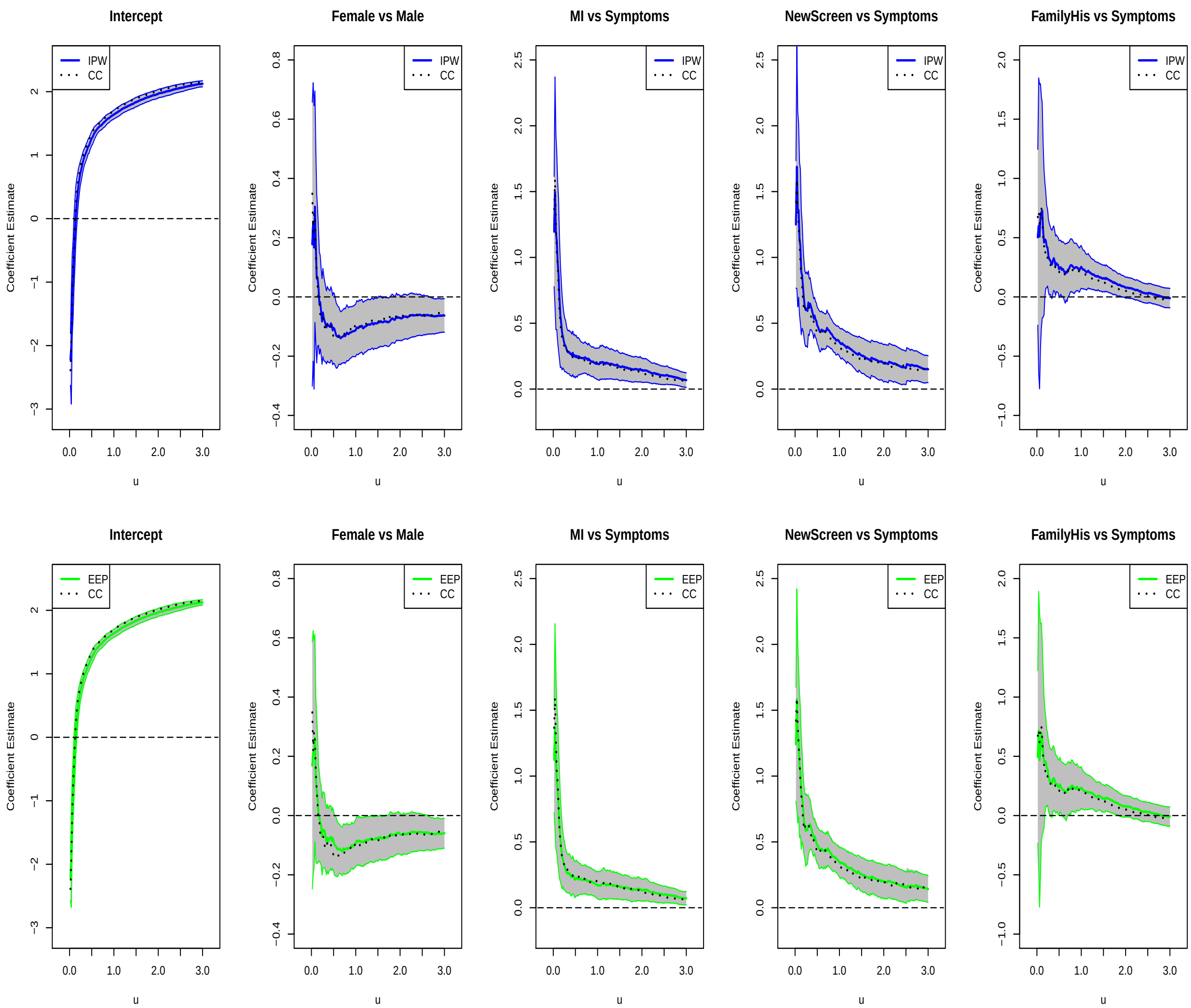

Figure S10. CFFPR data example: the proposed IPW coefficient estimates (blue solid lines) and their corresponding 95% pointwise confidence intervals, the proposed EEP coefficient estimates (green solid lines) and their corresponding 95% pointwise confidence intervals, along with the complete-case (CC) coefficient estimates (black dotted lines) for mucoid PA infection.