A KERNEL REGRESSION MODEL FOR PANEL COUNT DATA WITH TIME-VARING COEFFICIENTS

Yang Wang and Zhangsheng Yu Ph.D

Shanghai Jiao Tong University, Shanghai, China

Supplementary Material

In this supplementary note, we give some lemmas, which play a crucial role in the proof of Theorems 1−3. The detailed proofs of the theoretical results corresponding to the main document are presented. Meanwhile, we provide the cross-validation method and some additional notation for simplicity of presentation.

S1 Notation

In this section, the notation is the same as the main document. For simplicity of presentation, we introduce some additional notation, let $\boldsymbol{\alpha} = \boldsymbol{H}(\boldsymbol{\beta} - \boldsymbol{\beta}^*) = (\alpha_0, \ldots, \alpha_p)^{\mathrm{T}}$, where $\alpha_k = h^k\big\{\beta_k(t) - \beta^{(k)}(t)/k!\big\}$, $\boldsymbol{H} = diag(1, h, \ldots, h^p)$ is pth-order diagonal matrix and $\boldsymbol{\beta}^*$ is the true vector. $\tilde{z}_i(\boldsymbol{u}) = \boldsymbol{H}^{-1}z_i(\boldsymbol{u}) = z_i(1, (u-t)/h, \ldots, (u-t)^p/h^p)^{\mathrm{T}}$. For a matrix $\boldsymbol{A} = (a_{ij})$, $\|\boldsymbol{A}\| = \sup_{i,j}|a_{ij}|$. For a vector $\boldsymbol{a}$, $\|\boldsymbol{a}\| = \sup_i |a_i|$, and $|\boldsymbol{a}| = (\sum a_i^2)^{1/2}$. Some further definitions are:

For $j = 0, 1, 2$, set

$$S_{n,j}(u, \boldsymbol{\alpha}) = \frac{1}{n} \sum_{i=1}^{n} I(C_i \geq u) \exp(\boldsymbol{\alpha}^{\mathrm{T}} \tilde{z}_i(\boldsymbol{u}) + \boldsymbol{\beta}^{*\mathrm{T}} z_i(\boldsymbol{u})) o_i(u) z_i^j,$$

$$S_j(u, \boldsymbol{\alpha}) = E(p_1(u \mid z) p_2(u \mid z) \exp(\boldsymbol{\alpha}^{\mathrm{T}} \tilde{z}(\boldsymbol{u}) + \boldsymbol{\beta}^{*\mathrm{T}} z(\boldsymbol{u})) z^j).$$

For $j = 0, 1, 2$, set

$$\widetilde{S}_{n,j}(u) = \frac{1}{n} \sum_{i=1}^{n} I(C_i \geq u) \exp(\boldsymbol{\beta}^{*\mathrm{T}} z_i(\boldsymbol{u})) o_i(u) \tilde{z}_i(\boldsymbol{u})^{\otimes j},$$

$$\widetilde{S}_j(u) = E(p_1(u \mid z) p_2(u \mid z) \exp(\boldsymbol{\beta}^{*\mathrm{T}} z_i(\boldsymbol{u})) \tilde{z}(\boldsymbol{u})^{\otimes j}).$$

For $j = 0, 1$, put

$$\widetilde{S}^*_{n,j}(u) = \frac{1}{n} \sum_{i=1}^{n} I(C_i \geq u) exp(\beta(u) z_i) o_i(u) \tilde{z}_i^j(\boldsymbol{u}),$$

$$\widetilde{S}^*_j(u) = E(p_1(u \mid z) p_2(u \mid z) \exp(\beta(u) z) \tilde{z}^j(\boldsymbol{u})).$$

For $j = 0, 1, 2$, set

$$S_{n,j}(u, \boldsymbol{\beta}^*) = \frac{1}{n} \sum_{i=1}^{n} I(C_i \geq u) \exp(\boldsymbol{\beta}^{*\mathrm{T}} z_i(\boldsymbol{u})) o_i(u) z_i^j,$$

$$S_j(u, \boldsymbol{\beta}^*) = E(p_1(u \mid z) p_2(u \mid z) \exp(\boldsymbol{\beta}^{*\mathrm{T}} z_i(\boldsymbol{u})) z^j).$$

For $j = 0, 1, 2$, put

$$S^*_{n,j}(u, \beta(u)) = \frac{1}{n}\sum_{i=1}^{n} I(C_i \geq u)\exp(\beta(u)z_i)o_i(u)z_i^j,$$

$$S^*_j(u, \beta(u)) = E(p_1(u \mid z)p_2(u \mid z)\exp(\beta(u)z)z^j).$$

S2 Cross-validation method

In this part, we give the derived process of the approximations for $\widehat{\boldsymbol{\beta}}_{(-i)}$, as well as an asymptotic expression for the contribution $l_i(\boldsymbol{\beta})$ of individual i to the local partial likelihood. Then we can construct an alternative expression of cross-validation likelihood CVL.

S2.1 The expression of ith subject log-likelihood

Here, to facilitate notation, we omit the n in log-likelihood formula, and the local partial likelihood can be denote as:

$$\begin{aligned}\mathcal{L}(\boldsymbol{\beta}) = &\sum_{i=1}^{n}\int_0^{\tau} K_h(u-t)I\left(C_i \geq u\right)\Bigg[\boldsymbol{\beta}^{\mathrm{T}} z_i(\boldsymbol{u}) \\ &- \log\left\{\sum_{j=1}^{n} I\left(C_j \geq u\right)\exp\left(\boldsymbol{\beta}^{\mathrm{T}} z_j(\boldsymbol{u})\right) o_j(u)\right\}\Bigg] d\widetilde{N}_i(u),\end{aligned} \tag{S2.1}$$

which equivalent to the local partial likelihood defined in the main document. Analogous with (S2.1), we have the local partial likelihood which the

ith subject is left out:

$$\mathcal{L}_{(-i)}(\boldsymbol{\beta}) = \sum_{l\neq i}\int_0^{\tau} K_h(u-t) I\left(C_l \geq u\right)\Bigg[\boldsymbol{\beta}^{\mathrm{T}} z_l(\boldsymbol{u}) - \log\Bigg\{\sum_{j\neq i} I\left(C_j \geq u\right)\exp\left(\boldsymbol{\beta}^{\mathrm{T}} z_j(\boldsymbol{u})\right) o_j(u)\Bigg\}\Bigg] d\widetilde{N}_l(u). \tag{S2.2}$$

Then, from (S2.1) and (S2.2) yields,

$$\begin{aligned} l_i(\boldsymbol{\beta}) &= \mathcal{L}(\boldsymbol{\beta}) - \mathcal{L}_{(-i)}(\boldsymbol{\beta}) \\ &= \int_0^{\tau} K_h(u-t) I\left(C_i \geq u\right)\Bigg[\boldsymbol{\beta}^{\mathrm{T}} z_i(\boldsymbol{u}) \\ &- \log\Big\{\textstyle\sum_{j=1}^n I\left(C_j \geq u\right)\exp\left(\boldsymbol{\beta}^{\mathrm{T}} z_j(\boldsymbol{u})\right) o_j(u)\Big\}\Bigg] d\widetilde{N}_i(u) \\ &+ \textstyle\sum_{l\neq i}\int_0^{\tau} K_h(u-t) I\left(C_l \geq u\right)\log\left\{\frac{\sum_{j\neq i} I(C_j \geq u)\exp\left(\boldsymbol{\beta}^{\mathrm{T}} z_j(\boldsymbol{u})\right) o_j(u)}{\sum_{j=1}^n I(C_j \geq u)\exp(\boldsymbol{\beta}^{\mathrm{T}} z_j(\boldsymbol{u})) o_j(u)}\right\} d\widetilde{N}_l(u). \end{aligned} \tag{S2.3}$$

For the second term of right-hand side, the term

$$\begin{aligned} &\log\left\{\frac{\sum_{j\neq i} I\left(C_j \geq u\right)\exp\left(\boldsymbol{\beta}^{\mathrm{T}} z_j(\boldsymbol{u})\right) o_j(u)}{\sum_{j=1}^n I\left(C_j \geq u\right)\exp\left(\boldsymbol{\beta}^{\mathrm{T}} z_j(\boldsymbol{u})\right) o_j(u)}\right\} \\ &= \log\left\{1 - \frac{I\left(C_i \geq u\right)\exp\left(\boldsymbol{\beta}^{\mathrm{T}} z_i(\boldsymbol{u})\right) o_i(u)}{\sum_{j=1}^n I\left(C_j \geq u\right)\exp\left(\boldsymbol{\beta}^{\mathrm{T}} z_j(\boldsymbol{u})\right) o_j(u)}\right\}, \end{aligned}$$

approximately equal to zero, due to the latter term

$$\Big\{I\left(C_i \geq u\right)\exp\left(\boldsymbol{\beta}^{\mathrm{T}} z_i(\boldsymbol{u})\right) o_i(u) / \textstyle\sum_{j=1}^n I\left(C_j \geq u\right)\exp\left(\boldsymbol{\beta}^{\mathrm{T}} z_j(\boldsymbol{u})\right) o_j(u)\Big\}$$

is small.

Thus, we derive an alternative expression for $l_i(\boldsymbol{\beta})$ by

$$\begin{aligned} l_i(\boldsymbol{\beta}) = \int_0^{\tau} K_h(u-t) I\left(C_i \geq u\right) \Bigg[& \boldsymbol{\beta}^{\mathrm{T}} z_i(\boldsymbol{u}) - \log \Bigg\{ \sum_{j=1}^{n} I\left(C_j \geq u\right) \\ & \exp\left(\boldsymbol{\beta}^{\mathrm{T}} z_j(\boldsymbol{u})\right) o_j(u) \Bigg\} \Bigg] d\widetilde{N}_i(u), \end{aligned} \tag{S2.4}$$

which equivalent to the following log-likelihood,

$$\begin{aligned} l_i(\boldsymbol{\beta}) = & \frac{1}{n} \int_0^{\tau} K_h(u-t) I\left(C_i \geq u\right) \Bigg[\boldsymbol{\beta}^{\mathrm{T}} z_i(\boldsymbol{u}) - \log \Bigg\{ \frac{1}{n} \sum_{j=1}^{n} I\left(C_j \geq u\right) \\ & \exp\left(\boldsymbol{\beta}^{\mathrm{T}} z_j(\boldsymbol{u})\right) o_j(u) \Bigg\} \Bigg] d\widetilde{N}_i(u). \end{aligned} \tag{S2.5}$$

S2.2 Approximation of estimator

Here, we approximate $\widehat{\boldsymbol{\beta}}_{(-i)}$ using Taylor expansion. For $\mathcal{L}_{(-i)}(\boldsymbol{\beta}) = \mathcal{L}(\boldsymbol{\beta}) - l_i(\boldsymbol{\beta})$ defined in the main document. We have

$$\frac{\partial \mathcal{L}_{(-i)}}{\partial \boldsymbol{\beta}}(\boldsymbol{\beta}) = \frac{\partial \mathcal{L}}{\partial \boldsymbol{\beta}}(\boldsymbol{\beta}) - \frac{\partial l_i}{\partial \boldsymbol{\beta}}(\boldsymbol{\beta}). \tag{S2.6}$$

We approximate $\frac{\partial \mathcal{L}_{(-i)}}{\partial \boldsymbol{\beta}}(\boldsymbol{\beta})$ with first-order Taylor expansion at $\boldsymbol{\beta} = \widehat{\boldsymbol{\beta}}$, one has

$$\frac{\partial \mathcal{L}_{(-i)}}{\partial \boldsymbol{\beta}}(\boldsymbol{\beta}) = \frac{\partial \mathcal{L}_{(-i)}}{\partial \boldsymbol{\beta}}(\widehat{\boldsymbol{\beta}}) + \frac{\partial^2 \mathcal{L}_{(-i)}}{\partial \boldsymbol{\beta}^2}(\widehat{\boldsymbol{\beta}})(\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}}). \tag{S2.7}$$

Note that $\frac{\partial \mathcal{L}_{(-i)}}{\partial \boldsymbol{\beta}}(\widehat{\boldsymbol{\beta}}) = \frac{\partial \mathcal{L}}{\partial \boldsymbol{\beta}}(\widehat{\boldsymbol{\beta}}) - \frac{\partial l_i}{\partial \boldsymbol{\beta}}(\widehat{\boldsymbol{\beta}})$, and $\frac{\partial \mathcal{L}}{\partial \boldsymbol{\beta}}(\widehat{\boldsymbol{\beta}}) = 0$, we infer

$$\frac{\partial \mathcal{L}_{(-i)}}{\partial \boldsymbol{\beta}}(\boldsymbol{\beta}) = -\frac{\partial l_i}{\partial \boldsymbol{\beta}}(\widehat{\boldsymbol{\beta}}) + \frac{\partial^2 \mathcal{L}_{(-i)}}{\partial \boldsymbol{\beta}^2}(\widehat{\boldsymbol{\beta}})(\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}}). \tag{S2.8}$$

Substitute $\widehat{\boldsymbol{\beta}}_{(-i)}$ for $\boldsymbol{\beta}$ in (S2.8), note that $\frac{\partial \mathcal{L}_{(-i)}}{\partial \boldsymbol{\beta}}\left(\widehat{\boldsymbol{\beta}}_{(-i)}\right) = 0$, we obtain

$$0 = \frac{\partial \mathcal{L}_{(-i)}}{\partial \boldsymbol{\beta}}\left(\widehat{\boldsymbol{\beta}}_{(-i)}\right) = -\frac{\partial l_i}{\partial \boldsymbol{\beta}}(\widehat{\boldsymbol{\beta}}) + \frac{\partial^2 \mathcal{L}_{(-i)}}{\partial \boldsymbol{\beta}^2}(\widehat{\boldsymbol{\beta}})\left(\widehat{\boldsymbol{\beta}}_{(-i)} - \widehat{\boldsymbol{\beta}}\right). \tag{S2.9}$$

By direct calculate, we establish

$$\widehat{\boldsymbol{\beta}}_{(-i)} = \widehat{\boldsymbol{\beta}} + \left\{\frac{\partial^2 \mathcal{L}_{(-i)}}{\partial \boldsymbol{\beta}^2}(\widehat{\boldsymbol{\beta}})\right\}^{-1} \frac{\partial l_i}{\partial \boldsymbol{\beta}}(\widehat{\boldsymbol{\beta}}). \tag{S2.10}$$

From (S2.6), we have

$$\frac{\partial^2 \mathcal{L}_{(-i)}}{\partial \boldsymbol{\beta}^2}(\widehat{\boldsymbol{\beta}}) = \frac{\partial^2 \mathcal{L}}{\partial \boldsymbol{\beta}^2}(\widehat{\boldsymbol{\beta}}) - \frac{\partial^2 l_i}{\partial \boldsymbol{\beta}^2}(\widehat{\boldsymbol{\beta}}), \tag{S2.11}$$

and from (S2.5) yields

$$\begin{aligned}
&\frac{\partial^2 l_i}{\partial \boldsymbol{\beta}^2}(\boldsymbol{\beta}) = \\
&-\frac{1}{n}\int_0^{\tau} K_h(u-t) I\left(C_i \geq u\right)\left[\frac{\sum_{j=1}^n I\left(C_j \geq u\right)\exp\left(\boldsymbol{\beta}^{\mathrm{T}} z_j(\boldsymbol{u})\right) o_j(u) z_j(\boldsymbol{u})^{\otimes 2}}{\sum_{j=1}^n I\left(C_j \geq u\right)\exp\left(\boldsymbol{\beta}^{\mathrm{T}} z_j(\boldsymbol{u})\right) o_j(u)}\right. \\
&\left.-\left\{\frac{\sum_{j=1}^n I\left(C_j \geq u\right)\exp\left(\boldsymbol{\beta}^{\mathrm{T}} z_j(\boldsymbol{u})\right) o_j(u) z_j(\boldsymbol{u})}{\sum_{j=1}^n I\left(C_j \geq u\right)\exp\left(\boldsymbol{\beta}^{\mathrm{T}} z_j(\boldsymbol{u})\right) o_j(u)}\right\}^{\otimes 2}\right] d\widetilde{N}_i(u).
\end{aligned} \tag{S2.12}$$

Since the calculation of the above second derivation can increase computation burden. Omitting this term leads to, combined with (S2.10) and (S2.11),

$$\widehat{\boldsymbol{\beta}}_{(-i)} = \widehat{\boldsymbol{\beta}} + \left\{\frac{\partial^2 \mathcal{L}}{\partial \boldsymbol{\beta}^2}(\widehat{\boldsymbol{\beta}})\right\}^{-1} \frac{\partial l_i}{\partial \boldsymbol{\beta}}(\widehat{\boldsymbol{\beta}}). \tag{S2.13}$$

Then, we derive the approximation of the estimator $\widehat{\boldsymbol{\beta}}_{(-i)}$, which is the function of $\widehat{\boldsymbol{\beta}}$.

We are now prepared to approximate CVL, $CVL(h) = \sum_{i=1}^{n} l_i\left(\widehat{\boldsymbol{\beta}}_{(-i)}\right)$, defined in the main documents. A first-order Taylor approximation for $l_i(\boldsymbol{\beta})$, coupled with (S2.5) and (S2.13), yields

$$\begin{aligned} l_i\big(\widehat{\boldsymbol{\beta}}_{(-i)}\big) &= l_i\Big(\widehat{\boldsymbol{\beta}} + \left\{\frac{\partial^2 \mathcal{L}}{\partial \boldsymbol{\beta}^2}(\widehat{\boldsymbol{\beta}})\right\}^{-1} \frac{\partial l_i}{\partial \boldsymbol{\beta}}(\widehat{\boldsymbol{\beta}})\Big) \\ &= l_i(\widehat{\boldsymbol{\beta}}) + \left\{\frac{\partial l_i}{\partial \boldsymbol{\beta}}(\widehat{\boldsymbol{\beta}})\right\}^{\mathrm{T}} \left\{\frac{\partial^2 \mathcal{L}}{\partial \boldsymbol{\beta}^2}(\widehat{\boldsymbol{\beta}})\right\}^{-1} \frac{\partial l_i}{\partial \boldsymbol{\beta}}(\widehat{\boldsymbol{\beta}}) \\ &= l_i(\widehat{\boldsymbol{\beta}}) + \mathrm{tr}\left[\left\{\frac{\partial^2 \mathcal{L}}{\partial \boldsymbol{\beta}^2}(\widehat{\boldsymbol{\beta}})\right\}^{-1} \left\{\frac{\partial l_i}{\partial \boldsymbol{\beta}}(\widehat{\boldsymbol{\beta}})\right\}\left\{\frac{\partial l_i}{\partial \boldsymbol{\beta}}(\widehat{\boldsymbol{\beta}})\right\}^{\mathrm{T}}\right]. \end{aligned} \tag{S2.14}$$

Hence, which from (S2.14) gives

$$\begin{aligned} CVL(h) &= \sum_{i=1}^{n} l_i(\widehat{\boldsymbol{\beta}}) + \mathrm{tr}\left[\left\{\frac{\partial^2 \mathcal{L}}{\partial \boldsymbol{\beta}^2}(\widehat{\boldsymbol{\beta}})\right\}^{-1} \sum_{i=1}^{n} \left\{\frac{\partial l_i}{\partial \boldsymbol{\beta}}(\widehat{\boldsymbol{\beta}})\right\}\left\{\frac{\partial l_i}{\partial \boldsymbol{\beta}}(\widehat{\boldsymbol{\beta}})\right\}^{\mathrm{T}}\right] \\ &= \mathcal{L}(\widehat{\boldsymbol{\beta}}) + \mathrm{tr}\left[\left\{\frac{\partial^2 \mathcal{L}}{\partial \boldsymbol{\beta}^2}(\widehat{\boldsymbol{\beta}})\right\}^{-1} \sum_{i=1}^{n} \left\{\frac{\partial l_i}{\partial \boldsymbol{\beta}}(\widehat{\boldsymbol{\beta}})\right\}\left\{\frac{\partial l_i}{\partial \boldsymbol{\beta}}(\widehat{\boldsymbol{\beta}})\right\}^{\mathrm{T}}\right]. \end{aligned} \tag{S2.15}$$

Thus the derived process is complete.

S3 Lemmas

Before the proof of the theoretical results, we show two main conclusions which are conducive to the proofs of Theorems 1−3.

Lemma 1. *Let*

$$c_n(u) = \frac{1}{n}\sum_{i=1}^{n} I(C_i \geq u)o_i(u)g(u,z_i) \quad and \quad c(u) = E\left(p_1(u \mid z)p_2(u \mid z)g(u,z)\right),$$

if $g(u, z_i)$ *is bounded variation, then*

$$\sup_{u\in T} \|c_n(u) - c(u)\| = O_p(n^{-1/2}). \tag{S3.16}$$

Proof. Given $g(u, z_i)$ is bounded variation, under $C2$ and $C4$, we have $o_i(u)g(u, z_i)$ is bounded variation, then, we can write $o_i(u)g(u, z_i) = g_1(u, z_i) - g_2(u, z_i)$, where both $g_1(u, z_i)$ and $g_2(u, z_i)$ are nonnegative and nondecreasing. Thus

$$c_n(u) = \frac{1}{n}\sum_{i=1}^{n} \Big\{I(C_i \geq u)g_1(u,z_i) - I(C_i \geq u)g_2(u,z_i)\Big\}, \tag{S3.17}$$

and $I(C_i \geq u)$, for each i, are non-increasing in u, then by lemma A.2 of Bilias et al. (1997), $\{I(C_i \geq u), u \in T\}, \{g_j(u, z_i), u \in T\}_{j=1,2}$ have pseudodimension at most 1. By lemma 5.1 of Pollard (1990) combined with (S3.17), $\{I(C_i \geq u)o_i(u)g(u, z_i), u \in T\}$ has pseudodimension at most 10. Therefore, it must be Euclidean and certainly manageable according to theorem 4.8 of Pollard (1990). In view of $C2$, we choose envelops as $B_1/\sqrt{n}$, for some constant B_1. Then by theorem 8.3 (the uniform laws of large numbers) of Pollard (1990), we have $sup_{u\in T}\|c_n(u)-c(u)\| = O_p(n^{-1/2})$. □

Lemma 2. *Let* $T = [a, b] \subset R$*, suppose that*

$$\lim_{n\to\infty} \sup_{s\in T} \Big\{|h_n(s) - h(s)| + |J_n(s) - J(s)|\Big\} = 0, \tag{S3.18}$$

where $h_n(\cdot)$*,* $h(\cdot)$ *are continuous on* T*, and* $J_n(\cdot)$*,* $J(\cdot)$ *are right continuous with bounded variations on* T*. Then*

$$\lim_{n\to\infty} \sup_{s\in T} \Big\{|\int_a^s h_n(u)J_n(du) - \int_a^s h(u)J(du)|\Big\} = 0, \tag{S3.19}$$

$$\lim_{n\to\infty} \sup_{s\in T} \Big\{|\int_a^s h_n(u)J_n(du) - \int_a^s h_n(u)J(du)|\Big\} = 0. \tag{S3.20}$$

Proof. First, since h_n uniform converges to h, and J_n ,J are bounded variation functions with total variations bounded B_2, for some constant B_2.

Then

$$\lim_{n\to\infty}\sup_{s\in T}\Big\{|\int_a^s h_n(u)J_n(du)-\int_a^s h(u)J_n(du)|\Big\}=0, \tag{S3.21}$$

$$\lim_{n\to\infty}\sup_{s\in T}\Big\{|\int_a^s h_n(u)J(du)-\int_a^s h(u)J(du)|\Big\}=0. \tag{S3.22}$$

Since

$$\begin{aligned}&|\int_a^s h_n(u)J_n(du)-\int_a^s h_n(u)J(du)|\\ &\leq|\int_a^s h_n(u)J_n(du)-\int_a^s h(u)J(du)|+|\int_a^s h(u)J(du)-\int_a^s h_n(u)J(du)|.\end{aligned} \tag{S3.23}$$

Thus, from (S3.22) and (S3.23), we know that (S3.19) implies (S3.20). And

$$\begin{aligned}&|\int_a^s h_n(u)J_n(du)-\int_a^s h(u)J(du)|\\ &\leq|\int_a^s h_n(u)J_n(du)-\int_a^s h(u)J_n(du)|+|\int_a^s h(u)J_n(du)-\int_a^s h(u)J(du)|.\end{aligned} \tag{S3.24}$$

For the second term of the right-hand side in (S3.24), since $h(\cdot)$ is continuous, we can partition T by $a=s_0<\ldots<s_{n_0}=b$, and take constant

$h_j (= h(s_j))$ such that the simple function:

$$h_\varepsilon(s) = \sum_{j=0}^{n_0-1} h_j I\left(s \in [s_j, s_{j+1})\right), \tag{S3.25}$$

satisfies

$$\sup_{s\in T} |h_\varepsilon(s) - h(s)| < \varepsilon. \tag{S3.26}$$

Thus

$$\begin{aligned}
&|\int_a^s h(u)J_n(du) - \int_a^s h(u)J(du)| \\
&\leq |\int_a^s \big\{h(u) - h_\varepsilon(u)\big\} J_n(du)| + |\int_a^s h_\varepsilon(u)\big\{J_n(du) - J(du)\big\}| + |\int_a^s \big\{h(u) \\
&- h_\varepsilon(u)\big\} J(du)| \\
&\leq 2\varepsilon B_2 + |\int_a^s \sum_{j=0}^{n_0-1} h_j I(u \in [s_j - s_{j+1}))\big\{J_n(du) - J(du)\big\}| \\
&\leq 2\varepsilon B_2 + \sum_{j=0}^{n_0-1} |h_j||J_n(s_{j+1}) - J(s_{j+1}) - J_n(s_j) + J(s_j)| \\
&\leq 2\varepsilon B_2 + 2\sum_{j=0}^{n_0-1} |h_j| \sup_{s\in T} |J_n(s) - J(s)| \\
&\to 2\varepsilon B_2, \quad as \quad n \to \infty.
\end{aligned}$$

This in conjunction with (S3.21) and (S3.24), we obtain (S3.19). And from (S3.19) and (S3.22), then (S3.20) holds. □

S4 Detailed techniques for the main results proofs

S4.1 The detailed proof of Theorem 1

The proof of Theorem 1 is basically same as the proof of Lemma 2.2 of Hardle et al. (1988) and Theorem 2.1 of Zhao (1994). The major difference is that we have to treat a vector parameter $\boldsymbol{\beta}^{*} = (\beta(t), \beta'(t), \ldots, \beta^{(p)}/p!)^{\mathrm{T}}$ due to the local polynomial estimation. Next, we will show detailed proof procedure by the below two lemmas. Introduce some notation as follows:

$$G_{\alpha_k n1}(t, t+s) = \frac{1}{n}\sum_{i=1}^{n}\int_0^{\tau} I(C_i \geq u)I(t < u < t+s)\big\{(u-t)/h\big\}^k z_i d\widetilde{N}_i(u), \tag{S4.1}$$

and

$$G_{\alpha_k 1}(t, t+s) = E(G_{\alpha_k n1}(t, t+s)).$$

$$G_{\alpha_k n2}(t, t+s) = \frac{1}{n}\sum_{i=1}^{n}\int_0^{\tau} I(C_i \geq u)I(t < u < t+s)(\frac{u-t}{h})^k \frac{S_{n,1}(u, \boldsymbol{\alpha})}{S_{n,0}(u, \boldsymbol{\alpha})} d\widetilde{N}_i(u), \tag{S4.2}$$

and

$$G_{\alpha_k 2}(t, t+s) = E(G_{\alpha_k n2}(t, t+s)).$$

For $c > 0$,

$$V_{\alpha_k n1}(t,c) = \sup_{|s|\le c} |G_{\alpha_k n1}(t,t+s) - G_{\alpha_k 1}(t,t+s)|, \tag{S4.3}$$

$$V_{\alpha_k n2}(t,c) = \sup_{|s|\le c} |G_{\alpha_k n2}(t,t+s) - G_{\alpha_k 2}(t,t+s)|, \tag{S4.4}$$

where α_k is the kth component of $\boldsymbol{\alpha}$, and denote $\sup\{\alpha_k\} = \bar{\alpha}_k$, and $\inf\{\alpha_k\} = \underline{\alpha}_k$.

Lemma 3. *Let $0 < c_n \to 0$, as $n \to \infty$, and $1 < c_n^{-1} \le (n/\log n)^{1-2/\lambda}$, then almost surely (a.s.),*

$$V_{n1} = \sup_{t\in T} \sup_{\alpha_k \in \mathcal{N}_0} V_{\alpha_k n1}(t,c_n) = O\left(n^{-1/2}(c_n \log n)^{1/2}\right), \quad as \quad n \to \infty, \tag{S4.5}$$

$$V_{n2} = \sup_{t\in T} \sup_{\alpha_k \in \mathcal{N}_0} V_{\alpha_k n2}(t,c_n) = O\left(n^{-1/2}(c_n \log n)^{1/2}\right), \quad as \quad n \to \infty, \tag{S4.6}$$

where $\mathcal{N}_0 := \{\alpha_k : |\alpha_k - 0| < \epsilon\}$.

Proof. Since $V_{\alpha_k n1}$ is a special case of $V_{\alpha_k n2}$, when substituted $\frac{S_{n,1}(u,\boldsymbol{\alpha})}{S_{n,0}(u,\boldsymbol{\alpha})}$ by z_i. We only need to prove (S4.6). Put

$$a_n = n^{-1/2}(c_n \log n)^{1/2}.$$

Since we can treat the positive and negative part of z_i, separately, we assume that z_i is nonnegative. First, we reduce $\sup_{\alpha_k \in \mathcal{N}_0}$ in (S4.6) to a maximum on a finite set. We use finite points $b_1 < b_2 < \ldots < b_{N_n}$ to partition $\mathcal{N}_0$, such that $b_1 - \underline{\alpha}_k \leq a_n$, $\bar{\alpha}_k - b_{N_n} \leq a_n$, and $b_j - b_{j-1} \leq a_n$, for $2 \leq j \leq N_n$. Further, we assume that

$$N_n \leq 2(\bar{\alpha}_k - \underline{\alpha}_k)/a_n, \tag{S4.7}$$

and for any $t \in T$, and $|s| \leq c_n$, by Cauchy−Schwarz inequality, the functions $G_{\alpha_k n 2}(t, t+s)$ and $G_{\alpha_k 2}(t, t+s)$ are monotone in α_k. Letting J_n denote the set $\{\underline{\alpha}_k, b_1, \ldots, b_{N_n}, \bar{\alpha}_k\}$, and J_n^* denote the set $\{(\underline{\alpha}_k, b_1), (b_1, b_2), \ldots, (b_{N_n}, \bar{\alpha}_k)\}$. Hence, we have, for any $\alpha_k \in \mathcal{N}_0$,

$$\begin{aligned}
&G_{b_k n2}(t, t+s) - G_{b_k 2}(t, t+s) + G_{b_k 2}(t, t+s) - G_{b_{k+1} 2}(t, t+s) \\
&\leq G_{\alpha_k n2}(t, t+s) - G_{\alpha_k 2}(t, t+s) \\
&\leq G_{b_{k+1} n2}(t, t+s) - G_{b_{k+1} 2}(t, t+s) + G_{b_{k+1} 2}(t, t+s) - G_{b_k 2}(t, t+s).
\end{aligned}$$

Thus

$$|G_{\alpha_k n2}(t,t+s) - G_{\alpha_k 2}(t,t+s)|$$
$$\leq \max_{\alpha_k \in J_n} |G_{\alpha_k n2}(t,t+s) - G_{\alpha_k 2}(t,t+s)| + \max_{(\alpha'_k,\alpha''_k)\in J_n^*} |G_{\alpha''_k 2}(t,t+s)$$
$$- G_{\alpha'_k 2}(t,t+s)|.$$

For $\alpha'_k < \alpha''_k$,

$$G_{\alpha''_k 2}(t,t+s) - G_{\alpha'_k 2}(t,t+s)|$$
$$= |\int_0^\tau I(t<u<t+s)((u-t)/h)^k S_0^*(u,\beta(u))\Big\{\frac{S_2(u,\boldsymbol{\alpha})}{S_0(u,\boldsymbol{\alpha})} - (\frac{S_1(u,\boldsymbol{\alpha})}{S_0(u,\boldsymbol{\alpha})})^2\Big\}$$
$$(\alpha''_k - \alpha'_k)du|$$
$$\leq |\int_0^\tau I(t<u<t+s)((u-t)/h)^k M_0(\alpha''_k - \alpha'_k)du|$$
$$\leq M_0 a_n,$$

there exists some positive constant M_0 satisfied the upper inequality. Hence

$$V_{n2} \leq \sup_{t\in T} \max_{\alpha_k \in J_n} V_{\alpha_k n2}(t,c_n) + M_0 a_n. \tag{S4.8}$$

Next, we reduce $\sup_{t\in T}$ to a maximum on a finite set. Now, we partition T by an equally-spaced grid $I_n := \big\{t_k : t_k = kc_n, k = 0,\dots,[\tau/c_n]\big\}$, with $t_{[\tau/c_n]+1} = \tau$, where $[\cdot]$ denote the greatest integer part. For any $t \in T$ and

$|s| \leq c_n$, there exists a grid point t_k, such that both t and $t+s$ are between t_k and t_{k+1}. And

$$\begin{aligned}
&|G_{\alpha_k n2}(t,t+s) - G_{\alpha_k 2}(t,t+s)| \\
&\leq |G_{\alpha_k n2}(t_k,t+s) - G_{\alpha_k 2}(t_k,t+s)| + |G_{\alpha_k n2}(t_k,t) - G_{\alpha_k 2}(t_k,t)|.
\end{aligned}$$

Then, we obtain

$$|G_{\alpha_k n2}(t,t+s) - G_{\alpha_k 2}(t,t+s)| \leq 2\max_{t\in I_n} V_{\alpha_k n2}(t,c_n).$$

Thus

$$V_{n2} \leq 2\max_{t\in I_n}\max_{\alpha_k\in J_n} V_{\alpha_k n2}(t,c_n) + 2M_0 a_n. \tag{S4.9}$$

In order to apply Bernstein's inequality, we truncate $\{z_i\}$ by some value, and define $V^*_{\alpha_k n2}(t,c_n)$ similar to $V_{\alpha_k n2}(t,c_n)$. Put

$$Q_n = c_n/a_n,$$

and

$$\begin{aligned}
G^*_{\alpha_k n2}(t,t+s) = \frac{1}{n}\sum_{i=1}^{n}\int_0^\tau I(C_i\geq u)((u-t)/h)^k I(t<u<t+s)\Big\{\sum_{j=1}^{n} \\
I(C_j\geq u)\exp(\boldsymbol{\alpha}^{\mathrm{T}}\tilde{z}_j(\boldsymbol{u}) + \boldsymbol{\beta}^{*\mathrm{T}} z_j(\boldsymbol{u}))z_j I(z_j\leq Q_n)o_j(u)/S_{n,0}(u,\boldsymbol{\alpha})\Big\}d\widetilde{N}_i(u),
\end{aligned}$$

and

$$G^*_{\alpha_k 2}(t,t+s) = E\left(G^*_{\alpha_k n2}(t,t+s)\right).$$

Likewise, we have

$$V^*_{\alpha_k n2}(t, c_n) = \sup_{|s|\le c_n} |G^*_{\alpha_k n2}(t, t+s) - G^*_{\alpha_k 2}(t, t+s)|,$$

$$V^*_{n2} = \max_{t\in I_n} \max_{\alpha_k\in J_n} V^*_{\alpha_k n2}(t, c_n).$$

Thus

$$V_{n2} \le V^*_{n2} + 2M_0 a_n + 2A_{n1} + 2A_{n2}, \tag{S4.10}$$

where

$$A_{n1} = \sup_{t\in I_n} \sup_{\alpha_k\in J_n} \sup_{|s|\le c_n} \left| G_{\alpha_k n2}(t, t+s) - G^*_{\alpha_k n2}(t, t+s)\right|,$$

$$A_{n2} = \sup_{t\in I_n} \sup_{\alpha_k\in J_n} \sup_{|s|\le c_n} \left| G_{\alpha_k 2}(t, t+s) - G^*_{\alpha_k 2}(t, t+s)\right|.$$

For

$$
\begin{aligned}
&G_{\alpha_k n2}(t,t+s) - G^*_{\alpha_k n2}(t,t+s) \\
&= \frac{1}{n}\sum_{i=1}^{n}\int_0^{\tau} I(C_i \ge u)\big\{(u-t)/h\big\}^k I(t<u<t+s)\Big\{\sum_{j=1}^{n} I(C_j \ge u)\exp(\\
&\boldsymbol{\alpha}^{\mathrm{T}}\tilde{z}_j(\boldsymbol{u}) + \boldsymbol{\beta}^{*\mathrm{T}} z_j(\boldsymbol{u}))z_j I(z_j > Q_n) o_j(u)/S_{n,0}(u,\boldsymbol{\alpha})\Big\} d\widetilde{N}_i(u) \\
&\le Q_n^{1-\lambda}\frac{1}{n}\sum_{i=1}^{n}\int_0^{\tau} I(C_i \ge u)\Big\{\sum_{j=1}^{n} I(C_j \ge u)\exp(\boldsymbol{\alpha}^{\mathrm{T}}\tilde{z}_j(\boldsymbol{u}) + \boldsymbol{\beta}^{*\mathrm{T}} z_j(\boldsymbol{u}))z_j^{\lambda} \\
&o_j(u)/S_{n,0}(u,\boldsymbol{\alpha})\Big\} d\widetilde{N}_i(u).
\end{aligned}
\tag{S4.11}
$$

We have, by the classical strong low of large numbers and Lemma 1,

$$
\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n}\int_0^{\tau} I(C_i \ge u)\Big\{\sum_{j=1}^{n} I(C_j \ge u)\exp(\boldsymbol{\alpha}^{\mathrm{T}}\tilde{z}_j(\boldsymbol{u}) + \boldsymbol{\beta}^{*\mathrm{T}} z_j(\boldsymbol{u}))z_j^{\lambda} o_j(u)/ \\
&S_{n,0}(u,\boldsymbol{\alpha})\Big\} d\widetilde{N}_i(u) \to \int_0^{\tau} S_0^*(u,\beta(u))E\big(p_1(u \mid z)p_2(u \mid z)\exp(\boldsymbol{\alpha}^{\mathrm{T}}\tilde{z}(\boldsymbol{u}) + \\
&\boldsymbol{\beta}^{*\mathrm{T}} z(\boldsymbol{u}))z^{\lambda}\big)/S_0(u,\boldsymbol{\alpha})du < \infty, \quad a.s.
\end{aligned}
\tag{S4.12}
$$

Noting that

$$
a_n^{-1}Q_n^{1-\lambda} = \big\{c_n^{-1}(\log n/n)^{1-2/\lambda}\big\}^{\lambda/2} = o(1). \tag{S4.13}
$$

From (S4.11), (S4.12) and (S4.13), we have, as $n \to \infty$,

$$a_n^{-1} A_{n1} \to 0, \quad a.s. \tag{S4.14}$$

From (S4.11), (S4.12), (S4.13) and $A_{n2} \leq E(A_{n1})$, then, as $n \to \infty$,

$$a_n^{-1} A_{n2} \to 0, \quad a.s. \tag{S4.15}$$

Then, combining (S4.10), (S4.14) and (S4.15), it suffices for (S4.6) to show

$$V_{n2}^{*} = O(a_n), \quad a.s. \tag{S4.16}$$

Next, we will find a suitable upper bound for $pr\left(V_{n2}^{*} \geq B_0 a_n\right)$ by appropriate choice of B_0. Now, we perform a further partition for $V_{\alpha_k n2}^{*}(t, c_n)$ at a fixed $t \in I_n$. Set $w_n = [(Q_n c_n / a_n) + 1]$, and $s_r = r c_n / w_n$, for $r = -w_n, -w_n + 1, \ldots, w_n$. Since $G_{\alpha_k n2}^{*}(t, t+s)$ and $G_{\alpha_k 2}^{*}(t, t+s)$ are monotone in $|s|$, suppose that $0 \leq s_r \leq s \leq s_{r+1}$, then

$$\begin{aligned}
& G_{\alpha_k n2}^{*}(t, t+s_r) - G_{\alpha_k 2}^{*}(t, t+s_r) + G_{\alpha_k 2}^{*}(t, t+s_r) - G_{\alpha_k 2}^{*}(t, t+s_{r+1}) \\
& \leq G_{\alpha_k n2}^{*}(t, t+s) - G_{\alpha_k 2}^{*}(t, t+s) \\
& \leq G_{\alpha_k n2}^{*}(t, t+s_{r+1}) - G_{\alpha_k 2}^{*}(t, t+s_{r+1}) + G_{\alpha_k 2}^{*}(t, t+s_{r+1}) - G_{\alpha_k 2}^{*}(t, t+s_r),
\end{aligned}$$

from which we obtain

$$|G^*_{\alpha_k n2}(t, t+s) - G^*_{\alpha_k 2}(t, t+s)| \leq \max\{\xi_{n,r}, \xi_{n,r+1}\} + G^*_{\alpha_k 2}(t+s_r, t+s_{r+1}),$$

where

$$\xi_{n,r} = |G^*_{\alpha_k n2}(t, t+s_r) - G^*_{\alpha_k 2}(t, t+s_r)|.$$

The same holds for $s_r \leq s \leq s_{r+1} \leq 0$. Therefore

$$V^*_{\alpha_k n2}(t, c_n) \leq \max_{-w_n \leq r \leq w_n} \xi_{n,r} + \max_{-w_n \leq r \leq w_n - 1} G^*_{\alpha_k 2}(t+s_r, t+s_{r+1}). \quad \text{(S4.17)}$$

For all r, under $C5$,

$$G^*_{\alpha_k 2}(t+s_r, t+s_{r+1}) \leq \int_{t+s_r}^{t+s_{r+1}} q_0(u) Q_n du \leq M_3 Q_n (s_{r+1} - s_r) \leq M_3 a_n,$$

so that

$$pr\left(V^*_{\alpha_k n2}(t, c_n) \geq B_0 a_n\right) \leq pr\left(\max_{-w_n \leq r \leq w_n} \xi_{n,r} \geq (B_0 - M_3) a_n\right). \quad \text{(S4.18)}$$

Now, let

$$X_i = \int_0^\tau I(C_i \geq u)\big\{(u-t)/h\big\}^k I(t < u < t+s)\Big\{\sum_{j=1}^n I(C_j \geq u)\exp(\boldsymbol{\alpha}^{\mathrm{T}}\tilde{z}_j(\boldsymbol{u})$$
$$+\boldsymbol{\beta}^{*\mathrm{T}} z_j(\boldsymbol{u})) z_j I(z_j \geq Q_n) o_j(u)/S_{n,0}(u, \boldsymbol{\alpha})\Big\} d\widetilde{N}_i(u),$$

then

$\xi_{nr} = |\frac{1}{n}\sum_{i=1}^{n}\{X_i - E(X_i)\}|.$

For

$|X_i - E(X_i)| \leq |\int_0^\tau I(C_i \geq u)((u-t)/h)^k I(t < u < t+s)Q_n d\widetilde{N}_i(u)| \leq \bar{N}Q_n,$

where $\bar{N} = \tau \sup_{u\in T} N_i(u)$.

And, for some constant M_4, we have

$$\sum_{i=1}^{n} var(X_i) \leq \sum_{i=1}^{n} E(X_i^2)$$
$$\leq \sum_{i=1}^{n}\int_0^\tau I(t \leq u \leq t+s_r)\big\{(u-t)/h\big\}^k E(p_1(u \mid z)p_2(u \mid z)E(N^2(u) \mid z))$$
$$\Big\{E(p_1(u \mid z)p_2(u \mid z)\exp(\boldsymbol{\alpha}^{\mathrm{T}}\tilde{z}(\boldsymbol{u}) + \boldsymbol{\beta}^{*\mathrm{T}}z(\boldsymbol{u}))zI(z \leq Q_n))/S_0(u,\boldsymbol{\alpha})\Big\}^2 du$$
$$\leq \sum_{i=1}^{n}\int_{t+s_r}^{t+s_{r+1}} M_4 du \leq nM_4c_n.$$

Then, by Bernstein's inequality,

$$pr\,(\xi \geq (B_0 - M_3)a_n)$$
$$\leq \exp\Big\{-((B_0 - M_3)na_n)^2/2(\sum_{i=1}^{n} var(X_i) + 3^{-1}(B_0 - M_0)\bar{N}Q_n na_n)\Big\}$$
$$\leq \exp\Big\{-((B_0 - M_3)na_n)^2/2(M_4 nc_n + 3^{-1}(B_0 - M_0)\bar{N}Q_n na_n)\Big\} \leq n^{-B_0^*},$$

where

$$B_0^* = (B_0 - M_3)^2/2\big\{M_4 + 3^{-1}(B_0 - M_3)\bar{N}\big\}. \qquad \text{(S4.19)}$$

By (S4.18) and Boole's inequality,

$$pr\left(\sup_{t\in I_n}\sup_{\alpha_k\in J_n} V^*_{\alpha_k n2}(t,c_n)\geq B_0 a_n\right)\leq (N_n{+}2)([\tau/c_n]{+}1)2[(Q_n c_n/a_n){+}1]n^{-B_0^*}, \tag{S4.20}$$

where $[\cdot]$ denote the greatest integer part.

From (S4.7), we obtain

$$N_n+2\leq 2(\bar{\alpha}_k-\underline{\alpha}_k)a_n^{-1}+2.$$

And, obviously,

$$[\tau/c_n]+1\leq(\tau+1)c_n^{-1}.$$

Also,

$$2[(Q_n c_n/a_n)+1]\leq(2Q_n c_n/a_n)+2\leq 2\big\{(c_n a_n^{-1})^2+1\big\},$$

since

$$(c_n a_n^{-1})^2=c_n n/\log n\geq c_n^{-2/(\lambda-2)}\geq 1,$$

then, we have,

$$2[(Q_n c_n/a_n)+1]\leq 3c_n^2 a_n^{-2}.$$

Hence

$$pr\,(V_{n2}^{*} \geq B_0 a_n) \leq 2(\bar{\alpha}_k - \underline{\alpha}_k + 1)(\tau + 1)3c_n a_n^{-3} n^{-B_0^*} \leq \bar{M}_0 (n/\log n)^{(2\lambda - 1/\lambda)} n^{-B_0^*}, \tag{S4.21}$$

for some constant $\bar{M}_0$.

Given λ and real $\kappa > 0$, we choose a suitable B_0 denoted as $B_{\kappa,\lambda}$ to make the constant B_0^* in (S4.19) satisfies

$$B_0^* \geq \kappa + (2\lambda - 1)/\lambda.$$

And using $(2\lambda - 1)/\lambda = 2 - 1/\lambda > 1$, for $\lambda > 2$, then (S4.21) yields

$$pr\,(V_{n2}^{*} \geq B_{\kappa,\lambda} a_n) \leq \bar{M}_0 (\log n)^{-1} n^{-\kappa}. \tag{S4.22}$$

When $\kappa \geq 2$ in (S4.22), $pr\,(V_{n2}^{*} \geq B_{\kappa,\lambda} a_n)$ is summable in n. So, applying the Borel−Cantelli lemma,

$$V_{n2}^{*} = O(a_n), \quad a.s. \tag{S4.23}$$

Thus, form (S4.10), (S4.14), (S4.15) and (S4.23), we have

$$V_{n2} = O(a_n), \quad a.s.$$

Similarly, we can also prove $V_{n1} = O(a_n), \quad a.s.$ □

Lemma 4. *Let h be a bandwidth and $c_n = 2h$. Assume that $h \to 0$ and $h^{-1}(\log n/n)^{1-2/\lambda} = o(1)$, let*

$$U_{nk}(\boldsymbol{\alpha}) = \frac{1}{n}\sum_{i=1}^{n}\int_0^{\tau} I(C_i \geq u)K_h(u-t)((u-t)/h)^k\Big\{z_i - \frac{S_{n,1}(u,\boldsymbol{\alpha})}{S_{n,0}(u,\boldsymbol{\alpha})}\Big\}d\widetilde{N}_i(u), \tag{S4.24}$$

Then, we have

$$\sup_{t\in T}\sup_{\alpha_k\in\mathcal{N}_0}(nh/\log n)^{1/2}|U_{nk}(\boldsymbol{\alpha}) - E\left(U_{nk}(\boldsymbol{\alpha})\right)| = O(1), \quad a.s. \tag{S4.25}$$

Proof. Since $K(\cdot)$ is bounded variation function, so we can write $K(\cdot) = K_1(\cdot) - K_2(\cdot)$, where $K_1(\cdot)$ and $K_2(\cdot)$ are both increasing functions. Without loss of generality, suppose that $K_1(-1) = K_2(-1) = 0$. Next up, we apply Lemma 3 by letting $c_n = 2h$. It is clear that the assumption of Lemma 3 hold here. Write

$$\begin{aligned} U_{nk}(\boldsymbol{\alpha}) = &\int_{-h}^{h}\Bigg[\frac{1}{n}\sum_{i=1}^{n}\int_0^{\tau} I(C_i \geq u)I(v < u-t < h)(\frac{u-t}{h})^k\Big\{z_i - \frac{S_{n,1}(u,\boldsymbol{\alpha})}{S_{n,0}(u,\boldsymbol{\alpha})}\Big\} \\ &d\widetilde{N}_i(u)\Bigg]dK_h(v) \\ = &\int_{-h}^{h}\Big\{G_{\alpha_k n1}(t+v, t+h) - G_{\alpha_k n2}(t+v, t+h)\Big\}dK_h(v), \end{aligned}$$

where $G_{\alpha_k n1}$ and $G_{\alpha_k n2}$ defined as (S4.1) and (S4.2), respectively. So, we have

$$\begin{aligned}&\sup_{t\in T}\sup_{\alpha_k\in\mathcal{N}_0}|U_{nk}(\boldsymbol{\alpha})-E(U_{nk}(\boldsymbol{\alpha}))|\\&\le\sup_{t\in T}\sup_{\alpha_k\in\mathcal{N}_0}\Big\{V_{\alpha_k n1}(t,2h)+V_{\alpha_k n2}(t,2h)\Big\}\int_{-h}^{h}dK_h(v)\\&\le(K_1(1)+K_2(1))h^{-1}\sup_{t\in T}\sup_{\alpha_k\in\mathcal{N}_0}\Big\{V_{\alpha_k n1}(t,2h)+V_{\alpha_k n2}(t,2h)\Big\}.\end{aligned}$$

Hence, by the consequence of Lemma 3, we can derive

$$\sup_{t\in T}\sup_{\alpha_k\in\mathcal{N}_0}|U_{nk}(\boldsymbol{\alpha})-E(U_{nk}(\boldsymbol{\alpha}))|=O((\log n/(nh))^{1/2}),\quad a.s. \tag{S4.26}$$

Thus establishing (S4.25). □

Prove Theorem1

Proof. Since $\boldsymbol{\alpha}=\boldsymbol{H}(\boldsymbol{\beta}-\boldsymbol{\beta}^*)$ and $\alpha_k(t)=\alpha_k=h^k\big\{\beta_k(t)-\beta^{(k)}(t)/k!\big\}$ defined in Section **S1**, we have

$$\mathcal{L}_n(\boldsymbol{\alpha})=$$
$$\frac{1}{n}\sum_{i=1}^{n}\int_0^{\tau}I(C_i\ge u)K_h(u-t)\Big\{\boldsymbol{\alpha}^{\mathrm{T}}\tilde{z}_i(\boldsymbol{u})+\boldsymbol{\beta}^{*\mathrm{T}}z_i(\boldsymbol{u})-\log S_{n,0}(u,\boldsymbol{\alpha})\Big\}d\widetilde{N}_i(u), \tag{S4.27}$$

and

$$
\begin{aligned}
U_{nk}(\boldsymbol{\alpha}) &= \partial \mathcal{L}_n(\boldsymbol{\alpha})/\partial \alpha_k \\
&= \frac{1}{n}\sum_{i=1}^{n}\int_0^{\tau} I(C_i \geq u)K_h(u-t)\big\{(u-t)/h\big\}^k \Big\{z_i - \frac{S_{n,1}(u,\boldsymbol{\alpha})}{S_{n,0}(u,\boldsymbol{\alpha})}\Big\} d\widetilde{N}_i(u).
\end{aligned}
$$

By the assumption of $C3$, we have $w(h) = \sup_{|t-t'|\leq h} |\alpha_k(t) - \alpha_k(t')| = O(h)$. In this, we consider α_k in the neighborhood of zero, that is $\alpha_k \in \mathcal{N}_0$. And we take $\epsilon = \epsilon_k = \max\big\{2w(h), 6l_n/(\mu_{2k}M_1)\big\}$. Now, we consider $\alpha_k \in (-\epsilon_k, \epsilon_k)$, without loss of generality, we assume $\epsilon_k < 1$. Define

$$
U_{nk}(\epsilon_k) = \frac{1}{n}\sum_{i=1}^{n}\int_0^{\tau} I(C_i \geq 0)K_h(u-t)\big\{(u-t)/h\big\}^k \Big\{z_i - \frac{S_{n,1}(\epsilon_k, u)}{S_{n,0}(\epsilon_k, u)}\Big\} d\widetilde{N}_i(u), \tag{S4.28}
$$

with, for $j = 0, 1, 2$,

$$
S_{n,j}(\epsilon_k, u) = \sum_{i=1}^{n} I(C_i \geq 0)\exp\Big(\epsilon_k z_i \big\{(u-t)/h\big\}^k + \boldsymbol{\beta}^{*\mathrm{T}} z_i(\boldsymbol{u})\Big)\, o_i(u) z_i^j.
$$

So by Lemma 3 and Lemma 4, we have (as $n \to \infty$) $a.s.$, for any $t \in T$,

$$
|U_{nk}(\pm\epsilon_k) - E\left(U_{nk}(\pm\epsilon_k)\right)| \leq l_n, \tag{S4.29}
$$

where $l_n = O((\log n/(nh))^{1/2})$.

Under $C1$ to $C5$, and by Lemma 1, we have,

$$E\left(U_{nk}(\epsilon_k)\right) = \int_0^\tau K_h(u-t)\big\{(u-t)/h\big\}^k\big\{q_1(u)-q_0(u)S_1(\epsilon_k,u)/S_0(\epsilon_k,u)\big\}du, \tag{S4.30}$$

where, for $j = 1, 2, 3$,

$$q_j(u) = E\left(p_1(u \mid z)p_2(u \mid z)\mu_0(u)\exp(\beta(u)z)z^j\right),$$

$$S_j(\epsilon_k, u) = E\left(p_1(u \mid z)p_2(u \mid z)\exp(\epsilon_k z((u-t)/h)^k + \boldsymbol{\beta}^{*\mathrm{T}} z(\boldsymbol{u}))z^j\right).$$

Let $(u - t)/h = v$, and h sufficiently small, by Taylor expansion, we have,

$$E\left(U_{nk}(\epsilon_k, u)\right) = \int K(v)v^k\Big\{q_1(t) - q_0(t)E\big(p_1(t \mid z)p_2(t \mid z)z\exp(\epsilon_k z v^k +$$

$$\beta(t)z)\big)/E\left(p_1(t \mid z)p_2(t \mid z)\exp(\epsilon_k z v^k + \beta(t)z)\right)\Big\}dv + O(h).$$

For

$\exp(\epsilon_k z v^k+\beta(t)z) = \exp(\beta(t)z)\exp(\epsilon_k v^k z) = \exp(\beta(t)z)\big\{1+\epsilon_k v^k z+o(\epsilon_k)\big\}.$

Then

$$\begin{aligned} E\left(U_{nk}(\epsilon_k)\right) &= \int K(v)v^k\Big\{q_1(t) - q_0(t)\frac{q_1(t) + q_2(t)\epsilon_k v^k}{q_0(t) + q_1(t)\epsilon_k v^k}\Big\}dv + o(\epsilon_k) \\ &= -\int K(v)v^{2k}\sigma_1(t)\epsilon_k/\big\{1 + o(\epsilon_k) + \epsilon_k v^k q_1(t)/q_0(t)\big\}dv. \end{aligned}$$

Similarly,

$$E\left(U_{nk}(-\epsilon_k)\right) = \int K(v)v^{2k}\sigma_1(t)\epsilon_k/\left\{1 + o(\epsilon_k) - \epsilon_k v^k q_1(t)/q_0(t)\right\}dv.$$

Hence, under $C5$, we have,

$$E(U_{nk}(\epsilon_k)) \leq -3^{-1}\mu_{2k}M_1\epsilon_k, \tag{S4.31}$$

$$E(U_{nk}(-\epsilon_k)) \geq 3^{-1}\mu_{2k}M_1\epsilon_k. \tag{S4.32}$$

Therefore, combing (S4.27), (S4.29) and (S4.30), we obtain that (as $n \to \infty$) $a.s.$, for any $t \in T$,

$$U_{nk}(\epsilon_k) \leq l_n - 3^{-1}\mu_{2k}M_1\epsilon_k < 0,$$

$$U_{nk}(-\epsilon_k) \geq -l_n + 3^{-1}\mu_{2k}M_1\epsilon_k > 0.$$

Then the two above inequalities imply that $a.s.$, for any $t \in T$, there exists $\widehat{\alpha}_k(t) = \widehat{\alpha}_k \in (-\epsilon_k, \epsilon_k)$, such that $U_{nk}(\widehat{\alpha}_k(t)) = 0$, and $\widehat{\alpha}_k(t) = h^k(\widehat{\beta}_k(t) - \beta^{(k)}(t)/k!)$. Thus, we have,

$$\sup_{t\in T} |\widehat{\alpha}_k(t)| \leq \epsilon_k, \quad a.s.$$

and the above proof follows from $\epsilon_k = O((\log n/(nh))^{1/2} + h)$. Hence,

$$\sup_{t\in T} |\widehat{\beta}_k(t) - \beta^{(k)}(t)/k!| = O(h^{-k}\{\log n/(nh))^{1/2} + h\}), \quad a.s.$$

then Theorem 1 holds. □

S4.2 The detailed proof of Theorem 2

The proof of the asymptotic normality for the coefficient estimator is basically based on the functional central limit theorem of Pollard (1990). Similar to the proof of the Theorem 2.1 of Bilias et al. (1997), we will first show the asymptotic distribution of stochastic functions by the following lemma, which play a crucial role in the proof of Theorem 2.

Lemma 5. *For any nonzero vector* $\boldsymbol{a} = (a_1, \ldots, a_p)^{\mathrm{T}}$, *let*

$$u_1(s) = (h/n)^{1/2} \sum_{i=1}^{n} \int_0^s K_h(u-t)\boldsymbol{a}^{\mathrm{T}}(\boldsymbol{u}-\boldsymbol{t})dM_i(u), \tag{S4.33}$$

$$u_2(s) = (h/n)^{1/2} \sum_{i=1}^{n} \int_0^s K_h(u-t)\boldsymbol{a}^{\mathrm{T}}(\boldsymbol{u}-\boldsymbol{t})z_i dM_i(u), \tag{S4.34}$$

where

$$dM_i(u) = I(C_i \geq u)\big\{d\widetilde{N}_i(u) - \mu_0(u)\exp(\beta(u)z_i)dO_i(u)\big\}.$$

Under C1 *to* C5, *we have* $\{u_1(s), s \in T\}$ *and* $\{u_2(s), s \in T\}$ *converge in distribution to Gaussian processes* ξ_1 *and* ξ_2, *respectively, with continuous* *sample paths, mean 0 and covariance functions identified by*

$$E(\xi_1(s_1)\xi_1^{'}(s_2)) = \int_0^{s_1 \wedge s_2} hK_h^2(u-t)(\boldsymbol{a}^{\mathrm{T}}(\boldsymbol{u}-\boldsymbol{t}))^2 E(p_1(u \mid z)p_2(u \mid z) \sigma(u \mid z))du, \tag{S4.35}$$

$$E(\xi_2(s_1)\xi_2^{'}(s_2)) = \int_0^{s_1 \wedge s_2} hK_h^2(u-t)(\boldsymbol{a}^{\mathrm{T}}(\boldsymbol{u}-\boldsymbol{t}))^2 E(p_1(u \mid z)p_2(u \mid z)z^2 \sigma(u \mid z))du. \tag{S4.36}$$

Proof. Since u_1 is a special case of u_2, when we use 1 substitute for z_i in (S4.31), we only need to prove the convergence for u_2. In order to get the desired convergence, Theorem 10.7 (the functional central limit theorem) of Pollard (1990) was invoked. Therefore, conditions (i)-(v) need to be verified.

To verify (i), using the lemma A.1 of Bilias et al. (1997), it suffices to show both $\left\{ \int_0^s K_h(u-t)\boldsymbol{a}^{\mathrm{T}}(\boldsymbol{u}-\boldsymbol{t})I(C_i \geq u)z_i d\widetilde{N}_i(u), s \in T \right\}$ and $\left\{ \int_0^s K_h(u-t)\boldsymbol{a}^{\mathrm{T}}(\boldsymbol{u}-\boldsymbol{t})I(C_i \geq u)\mu_0(u)\exp(\beta(u)z_i)dO_i(u), s \in T \right\}$ are manageable. Without loss of generality, we assume $\boldsymbol{a}^{\mathrm{T}}(\boldsymbol{u}-\boldsymbol{t}) > 0$ and $z_i > 0$. Thus, for each i, $\int_0^s K_h(u-t)\boldsymbol{a}^{\mathrm{T}}(\boldsymbol{u}-\boldsymbol{t})I(C_i \geq u)z_i d\widetilde{N}_i(u)$ is nondecreasing in s. Then it has pseudodimension at most 1. By Theorem 4.8 of Pollard (1990), therefore it must be Euclidean and manageable. Similarly, $\left\{ \int_0^s K_h(u-t)\boldsymbol{a}^{\mathrm{T}}(\boldsymbol{u}-\boldsymbol{t})I(C_i \geq u)\mu_0(u)\exp(\beta(u)z_i)dO_i(u), s \in T \right\}$ are also

Euclidean and manageable. Thus (i) holds.

To verify (ii), under $C1$ to $C5$ and lemma 1,

$\lim_{n\to\infty} E\left(u_2(s_1)u_2(s_2)\right)$

$= \lim_{n\to\infty} \frac{h}{n}\sum_{i=1}^{n} E\big(\big\{\int_0^{s_1} K_h(u-t)\boldsymbol{a}^{\mathrm{T}}(\boldsymbol{u}-\boldsymbol{t})z_i dM_i(u\big\}\big\{\int_0^{s_2} K_h(u-t)\boldsymbol{a}^{\mathrm{T}}$

$(\boldsymbol{u}-\boldsymbol{t})z_i dM_i(u)\big\}\big)$

$= \int_0^{s_1\wedge s_2} hK_h^2(u-t)\big\{\boldsymbol{a}^{\mathrm{T}}(\boldsymbol{u}-\boldsymbol{t})\big\}^2 E\left(p_1(u\mid z)p_2(u\mid z)z^2\sigma(u\mid z)\right)du.$

Thus (ii) holds. By the classical multivariate central limit theorem, we obtain that the convergence of finite-dimensional distributions of u_2 to those of ξ_2 is straightforward. The latter issue is tightness.

For (iii), (iv), under $C2$ and $C3$, envelops can be chosen as $B^*/\sqrt{n}$, for some constant B^*. Thus (iii) and (iv) holds.

To test (v), for any $s_1, s_2 \in T$, define

$$\rho_n(s_1,s_2) = E\left(u_2(s_1)-u_2(s_2)\right)^2, \quad \rho(s_1,s_2) = E\left(\xi_2(s_2)-\xi_2(s_1)\right)^2.$$

Here,

$\rho_n(s_1,s_2) = E\left(u_2(s_2)-u_2(s_1)\right)^2$

$= \frac{1}{n}\sum_{i=1}^{n} E\left(h\big\{\int_{s_1}^{s_2} K_h(u-t)\boldsymbol{a}^{\mathrm{T}}(\boldsymbol{u}-\boldsymbol{t})z_i dM_i(u)\big\}^2\right)$

$= \frac{1}{n}\sum_{i=1}^{n} E\big(\big|\int_{s_1}^{s_2} hK_h^2(u-t)\big\{\boldsymbol{a}^{\mathrm{T}}(\boldsymbol{u}-\boldsymbol{t})\big\}^2 z_i^2 I(C_i\geq u)\mu_0^2(u)\exp(2\beta(u)z_i)$

$o_i(u)du\big|\big).$

Clearly, ρ_n is equicontinuous on T, and $\lim_{n\to\infty}\rho_n(s_1,s_2) = \rho(s_1,s_2), \rho$ is pseudometric on T. Thus ρ_n converges to ρ, uniformly on T. Furthermore, we set $\{s_1^n\}, \{s_2^n\}$ be any two sequences in T, it follows that if $\rho(s_1^n, s_2^n) \to 0$, then $\rho_n(s_1^n, s_2^n) \to 0$. Thus (v) holds.

Therefore, using Theorem 10.7 (the functional central limit theorem) of Pollard (1990), we can state u_2 converges in distribution to Gaussian process on T having continuous sample path. Hence, $\{u_1(s), s \in T\}$ and $\{u_2(s), s \in T\}$ converges in distribution to Gaussian processes ξ_1 and ξ_2, respectively.

□

Prove Theorem 2

Proof. Let $\gamma_n = (nh)^{-1/2}$, $\boldsymbol{\alpha} = \gamma_n^{-1}\boldsymbol{H}(\boldsymbol{\beta} - \boldsymbol{\beta}^*)$, then

$$X_n(\gamma_n\boldsymbol{\alpha}, \tau) = \frac{1}{n}\sum_{i=1}^{n}\int_0^{\tau} I(C_i \geq u)K_h(u-t)\Bigg[\gamma_n\boldsymbol{\alpha}^{\mathrm{T}}\tilde{z}_i(\boldsymbol{u}) - \log\Bigg\{\sum_{i=1}^{n} I(C_i \geq u)\exp(\gamma_n\boldsymbol{\alpha}^{\mathrm{T}} \tilde{z}_j(\boldsymbol{u}) + \boldsymbol{\beta}^{*\mathrm{T}}z_j(\boldsymbol{u}))o_j(u)/\sum_{i=1}^{n} I(C_i \geq u)\exp(\boldsymbol{\beta}^{*\mathrm{T}}z_j(\boldsymbol{u}))o_j(u)\Bigg\}\Bigg]d\widetilde{N}_i(u).$$

Let

$$I(C_i \geq u)d\widetilde{N}_i(u) = dM_i(u) + I(C_i \geq u)\mu_0(u)\exp(\beta(u)z_i)dO_i(u),$$

then

$$X_n(\gamma_n\boldsymbol{\alpha}, \tau) = A_n(\gamma_n\boldsymbol{\alpha}, \tau) + U_n(\gamma_n\boldsymbol{\alpha}, \tau), \tag{S4.37}$$

where

$A_n(\gamma_n\boldsymbol{\alpha},\tau) = \frac{1}{n}\sum_{i=1}^n \int_0^\tau K_h(u-t)\left[\gamma_n\boldsymbol{\alpha}^{\mathrm{T}}\tilde{z}_i(\boldsymbol{u}) - \log\left\{\frac{S_{n,0}(u,\gamma_n\boldsymbol{\alpha})}{\widetilde{S}_{n,0}(u)}\right\}\right] I(C_i \geq u)\mu_0(u)\exp(\beta(u)z_i)o_i(u)du,$

$U_n(\gamma_n\boldsymbol{\alpha},\tau) = \frac{1}{n}\sum_{i=0}^n \int_0^\tau K_h(u-t)\left[\gamma_n\boldsymbol{\alpha}^{\mathrm{T}}\tilde{z}_i(\boldsymbol{u}) - \log\left\{\frac{S_{n,0}(u,\gamma_n\boldsymbol{\alpha})}{\widetilde{S}_{n,0}(u)}\right\}\right] dM_i(u).$

For

$A_n(\gamma_n\boldsymbol{\alpha},\tau) = \int_0^\tau K_h(u-t)\left[\widetilde{S}^*_{n,1}(u)^{\mathrm{T}}\gamma_n\boldsymbol{\alpha} - \log\left\{\frac{S_{n,0}(u,\gamma_n\boldsymbol{\alpha})}{\widetilde{S}_{n,0}(u)}\right\}\widetilde{S}^*_{n,0}(u)\right]\mu_0(u)du,$

by Taylor expansion of $S_{n,0}(u,\gamma_n\boldsymbol{\alpha})$ at $\boldsymbol{\alpha}=0$, it follows,

$$
\begin{aligned}
&\log\left\{S_{n,0}(u,\gamma_n\boldsymbol{\alpha})/\widetilde{S}_{n,0}(u)\right\}\\
&= (\widetilde{S}_{n,1}(u)/\widetilde{S}_{n,0}(u))^{\mathrm{T}}\gamma_n\boldsymbol{\alpha} + 2^{-1}\gamma_n^2\boldsymbol{\alpha}^{\mathrm{T}}\left\{\frac{\widetilde{S}_{n,2}(u)}{\widetilde{S}_{n,0}(u)} - (\frac{\widetilde{S}_{n,1}(u)}{\widetilde{S}_{n,0}(u)})^{\otimes 2}\right\}\boldsymbol{\alpha} + o_p(\gamma_n^2)\\
&= (\widetilde{S}_1(u)/\widetilde{S}_0(u))^{\mathrm{T}}\gamma_n\boldsymbol{\alpha} + 2^{-1}\gamma_n^2\boldsymbol{\alpha}^{\mathrm{T}}\left\{\frac{\widetilde{S}_2(u)}{\widetilde{S}_0(u)} - (\frac{\widetilde{S}_1(u)}{\widetilde{S}_0(u)})^{\otimes 2}\right\}\boldsymbol{\alpha} + o_p(\gamma_n^2).
\end{aligned}
$$

Hence

$$A_n(\gamma_n\boldsymbol{\alpha},\tau) = \gamma_n A_{n,1}(\tau)^{\mathrm{T}}\boldsymbol{\alpha} - 2^{-1}\gamma_n^2\boldsymbol{\alpha}^{\mathrm{T}}F_{n,1}(\tau)\boldsymbol{\alpha} + o_p(\gamma_n^2),$$

where

$$
\begin{aligned}
A_{n,1}(\tau) &= \int_0^\tau K_h(u-t)\left\{\widetilde{S}^*_1(u) - \widetilde{S}_1(u)\widetilde{S}^*_0(u)/\widetilde{S}_0(u)\right\}\mu_0(u)du,\\
F_{n,1}(\tau) &= \int_0^\tau K_h(u-t)\left\{\widetilde{S}_2(u)/\widetilde{S}_0(u) - (\widetilde{S}_1(u)/\widetilde{S}_0(u))^{\otimes 2}\right\}\widetilde{S}^*_0(u)\mu_0(u)du.
\end{aligned}
$$

For $\mid u-t \mid< ch$, let $u=t+hv$, under $C1$ to $C5$, we have

$$F_{n,1}(\tau)=\int K(v)\Big\{\frac{\widetilde{S}_2(t+hv)}{\widetilde{S}_0(t+hv)}-(\frac{\widetilde{S}_1(t+hv)}{\widetilde{S}_0(t+hv)})^{\otimes 2}\Big\}\widetilde{S}_0^*(t+hv)\mu_0(t+hv)dv$$
$$=\sigma_1(t)\Omega_1+o_p(1),$$

where $\Omega_1=\int K(v)\boldsymbol{v}^{\otimes 2}dv$, and $\boldsymbol{v}=(1,v,\dots,v^p)^{\mathrm{T}}$.

Thus

$$A_n(\gamma_n\boldsymbol{\alpha},\tau)=\gamma_n A_{n,1}(\tau)^{\mathrm{T}}\boldsymbol{\alpha}-2^{-1}\gamma_n^2\boldsymbol{\alpha}^{\mathrm{T}}\sigma_1(t)\Omega_1\boldsymbol{\alpha}+o_p(\gamma_n^2). \qquad \text{(S4.38)}$$

Similarly, we have

$$U_n(\gamma_n\boldsymbol{\alpha},\tau)=\gamma_n\boldsymbol{\alpha}^{\mathrm{T}}U_{n,1}(\tau)-2^{-1}\gamma_n^2\boldsymbol{\alpha}^{\mathrm{T}}F_{n,2}(\tau)\boldsymbol{\alpha}+o_p(\gamma_n^2),$$

where

$$U_{n,1}(\tau)=\int_0^\tau K_h(u-t)\frac{1}{n}\sum_{i=1}^n\Big\{\tilde{z}_i(\boldsymbol{u})-\widetilde{S}_{n,1}(u)/\widetilde{S}_{n,0}(u)\Big\}dM_i(u),$$
$$F_{n,2}(\tau)=\frac{1}{n}\sum_{i=1}^n\int_0^\tau K_h(u-t)\Big\{\widetilde{S}_{n,2}(u)/\widetilde{S}_{n,0}(u)-(\widetilde{S}_{n,1}(u)/\widetilde{S}_{n,0}(u))^{\otimes 2}\Big\}dM_i(u).$$

For $F_{n,2}(\tau)$, similar to Lemma 5, we have $\Big\{\int_0^s K_h(u-t)dM_i(u), s\in T\Big\}$ is manageable. Let constant $\bar{B}/\sqrt{n}$ as envelope. Thus, using Theorem 8.3

(the uniform law of large numbers) of Pollard (1990), we can derive

$$\lim_{n\to\infty}\sup_{s\in T}\|\frac{1}{n}\sum_{i=1}^{n}\int_0^s K_h(u-t)dM_i(u)-0\|=0.$$

Also, by Lemma 1, as $n\to\infty$,

$\sup_{s\in T}\|\{\widetilde{S}_{n,2}(u)/\widetilde{S}_{n,0}(u)-(\widetilde{S}_{n,1}(u)/\widetilde{S}_{n,0}(u))^{\otimes 2}\}-\{\widetilde{S}_2(u)/\widetilde{S}_0(u)-(\widetilde{S}_1(u)/\widetilde{S}_0(u))^{\otimes 2}\}\|\to 0.$

Then, by lemma 2, we have

$$F_{n,2}(\tau)=O_p(\gamma_n),$$

Therefore,

$$U_n(\gamma_n\boldsymbol{\alpha},\tau)=\gamma_n\boldsymbol{\alpha}^{\mathrm{T}}U_{n,1}(\tau)+O_p(\gamma_n^2). \tag{S4.39}$$

From (S4.37), (S4.38) and (S4.39), we obtain

$$X_n(\gamma_n\boldsymbol{\alpha},\tau)=\Big\{A_{n,1}(\tau)+U_{n,1}(\tau)\Big\}^{\mathrm{T}}\gamma_n\boldsymbol{\alpha}-2^{-1}\gamma_n^2\boldsymbol{\alpha}^{\mathrm{T}}\sigma_1(t)\Omega_1\boldsymbol{\alpha}+o_p(\gamma_n^2).$$

Using Quadratic Approximation Lemma of Fan and Gijbels (1996), we derive

$$\widehat{\boldsymbol{\alpha}}=\gamma_n^{-1}(\sigma_1(t)\Omega_1)^{-1}\Big\{A_{n,1}(\tau)+U_{n,1}(\tau)\Big\}+o_p(1). \tag{S4.40}$$

For

$$A_{n,1}(\tau)=\int_0^\tau K_h(u-t)\Big\{\widetilde{S}_1^*(u)-\widetilde{S}_1(u)\widetilde{S}_0^*(u)/\widetilde{S}_0(u)\Big\}\mu_0(u)du.$$

We apply Taylor expansion to the term:

$$\widetilde{S}_1^*(u)-\widetilde{S}_1(u)\widetilde{S}_0^*(u)/\widetilde{S}_0(u)=\widetilde{S}_1^*(u)-\widetilde{S}_1(u)-\widetilde{S}_1(u)\{\widetilde{S}_0^*(u)-\widetilde{S}_0(u)\}/\widetilde{S}_0(u).$$

Note that

$\beta(u)z \approx \beta(t)z + \beta'(t)z(u-t) + \cdots + \beta^{(p)}(t)z(u-t)^p/p! + \beta^{(p+1)}(t)z(u-t)^{p+1}/(p+1)! = \boldsymbol{\beta}^{*\mathrm{T}}z(\boldsymbol{u}) + \beta^{(p+1)}(t)z(u-t)^{p+1}/(p+1)!$.

Then

$\exp(\beta(u)z) - \exp(\boldsymbol{\beta}^{*\mathrm{T}}z(\boldsymbol{u}))$

$\approx \exp(\beta(u)z)\big\{1 - \exp(-\beta^{(p+1)}(t)z(u-t)^{p+1}/(p+1)!)\big\}$

$\approx \exp(\beta(u)z)\beta^{(p+1)}(t)z(u-t)^{p+1}/(p+1)!$.

Thus

$\widetilde{S}_1^*(u) - \widetilde{S}_1(u) = E\left(p_1(u \mid z)p_2(u \mid z)\exp(\beta(u)z)z\tilde{z}(\boldsymbol{u})\right)\beta^{(p+1)}(t)(u-t)^{p+1}/(p+1)! + o((u-t)^{p+1})$,

$\widetilde{S}_0^*(u) - \widetilde{S}_0(u) = E\left(p_1(u \mid z)p_2(u \mid z)\exp(\beta(u)z)z\right)\beta^{(p+1)}(t)(u-t)^{p+1}/(p+1)! + o((u-t)^{p+1})$,

$\widetilde{S}_0(u) = E\left(p_1(u \mid z)p_2(u \mid z)\exp(\beta(u)z)\right) + O((u-t)^{p+1})$,

$\widetilde{S}_1(u) = E\left(p_1(u \mid z)p_2(u \mid z)\exp(\beta(u)z)\tilde{z}(\boldsymbol{u})\right) + O((u-t)^{p+1})$.

Therefore, we have

$A_{n,1}(\tau) = \int_0^\tau K_h(u-t)\Bigg[\Big\{E(p_1(t \mid z)p_2(t \mid z)\exp(\beta(u)z)z\tilde{z}(\boldsymbol{u})) - \widetilde{S}_1^*(u) S_1^*(u,\beta(u))/S_0^*(u,\beta(u))\Big\}\beta^{(p+1)}(t)(u-t)^{p+1}/(p+1)! + o((u-t)^{p+1})\Bigg]du.$

Let $u = t + hv$, we derive

$$A_{n,1}(\tau) = \int K(v)\boldsymbol{v}v^{p+1}dv\sigma_1(t)h^{p+1}\beta^{(p+1)}(t)/(p+1)! + o(h^{p+1}). \quad \text{(S4.41)}$$

From (S4.40) and (S4.41), we obtain (let $\boldsymbol{b} = \int K(v)v^{p+1}\boldsymbol{v}dv$)

$$\widehat{\boldsymbol{\alpha}} = \gamma_n^{-1}\Omega_1^{-1}\boldsymbol{b}h^{p+1}\beta^{(p+1)}(t)/(p+1)! + \gamma_n^{-1}\sigma_1^{-1}(t)\Omega_1^{-1}U_{n,1}(\tau) + o_p(1).$$

Hence,

$$\begin{aligned}&(nh)^{1/2}\Big\{\boldsymbol{H}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*) - \Omega_1^{-1}\boldsymbol{b}h^{p+1}\beta^{(p+1)}(t)/(p+1)!\Big\}\\ &= \gamma_n^{-1}\sigma_1^{-1}(t)\Omega_1^{-1}U_{n,1}(\tau) + o_p(1).\end{aligned} \tag{S4.42}$$

Therefore, (S4.40) can be reduced to prove the multivariate normality of $(nh)^{1/2}U_{n,1}(\tau)$. That is equivalent to prove the normality of $\boldsymbol{a}^{\mathrm{T}}(nh)^{1/2}U_{n,1}(\tau)$, for any nonzero vector $\boldsymbol{a} = (a_1, \ldots, a_p)^{\mathrm{T}}$. Write $\widetilde{U}_n(s) = \boldsymbol{a}^{\mathrm{T}}(nh)^{1/2}U_{n,1}(s)$ is empirical process, we will show that it converges to Gaussian process $\tilde{\xi}$. In fact,

$$\widetilde{U}_n(s) = \widetilde{U}_{n1}(s) + \widetilde{U}_{n2}(s),$$

where

$$\widetilde{U}_{n1}(s) = \sqrt{\frac{h}{n}}\sum_{i=1}^n \int_0^s K_h(u-t)\boldsymbol{a}^{\mathrm{T}}(\boldsymbol{u}-\boldsymbol{t})\Big\{z_i - \frac{S_1(u,\boldsymbol{\beta}^*)}{S_0(u,\boldsymbol{\beta}^*)}\Big\}dM_i(u),$$
$$\widetilde{U}_{n2}(s) = \sqrt{\frac{h}{n}}\sum_{i=1}^n \int_0^s K_h(u-t)\boldsymbol{a}^{\mathrm{T}}(\boldsymbol{u}-\boldsymbol{t})\Big\{\frac{S_1(u,\boldsymbol{\beta}^*)}{S_0(u,\boldsymbol{\beta}^*)} - \frac{S_{n,1}(u,\boldsymbol{\beta}^*)}{S_{n,0}(u,\boldsymbol{\beta}^*)}\Big\}dM_i(u).$$

For $\widetilde{U}_{n2}(s)$, by Lemma 5 and the Strong Representation Theorem of Pollard (1990), we can construct a new probability space, and have

$$\sup_{s\in T}\|u_1(s) - \xi_1(s)\| \to 0, \quad as \quad n\to\infty,$$

and by Lemma 1, we have

$\sup_{s\in T} \|S_1(u,\boldsymbol{\beta}^*)/S_0(u,\boldsymbol{\beta}^*) - S_{n,1}(u,\boldsymbol{\beta}^*)/S_{n,0}(u,\boldsymbol{\beta}^*)\| \to 0, \quad as \quad n\to\infty.$

Then, by Lemma 2, we can show that, almost surely, as $n \to \infty$,

$$\sqrt{\frac{h}{n}}\sum_{i=1}^{n}\int_0^s K_h(u-t)\boldsymbol{a}^{\mathrm{T}}(\boldsymbol{u}-\boldsymbol{t})\Big\{\frac{S_1(u,\boldsymbol{\beta}^*)}{S_0(u,\boldsymbol{\beta}^*)} - \frac{S_{n,1}(u,\boldsymbol{\beta}^*)}{S_{n,0}(u,\boldsymbol{\beta}^*)}\Big\}dM_i(u) \to 0.$$

which holds in original probability space since the statement is now in probability. Thus the convergence of $\widetilde{U}_n(s)$ reduces to that of $\widetilde{U}_{n1}(s)$. Here,

$$\begin{aligned}
&\lim_{n\to\infty} E\left(\widetilde{U}_{n1}(s_1)\widetilde{U}_{n1}(s_2)\right)\\
&= \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n} E\Big(h\Big[\int_0^{s_1} K_h(u-t)\boldsymbol{a}^{\mathrm{T}}(\boldsymbol{u}-\boldsymbol{t})\big\{z_i - \frac{S_1(u,\boldsymbol{\beta}^*)}{S_0(u,\boldsymbol{\beta}^*)}\big\}dM_i(u)\Big]\\
&\Big[\int_0^{s_2} K_h(u-t)\boldsymbol{a}^{\mathrm{T}}(\boldsymbol{u}-\boldsymbol{t})\big\{z_i - S_1(u,\boldsymbol{\beta}^*)/S_0(u,\boldsymbol{\beta}^*)\big\}dM_i(u)\Big]\Big)\\
&= \int_0^{s_1\wedge s_2} hK_h^2(u-t)\big\{\boldsymbol{a}^{\mathrm{T}}(\boldsymbol{u}-\boldsymbol{t})\big\}^2 E(p_1(u\mid z)p_2(u\mid z)\Big\{z - \frac{S_1(u,\boldsymbol{\beta}^*)}{S_0(u,\boldsymbol{\beta}^*)}\Big\}^2\\
&\mu_0^2(u)\exp(2\beta(u)z))du\\
&= E(\tilde{\xi}(s_1)\tilde{\xi}(s_2)).
\end{aligned}$$

Then, the convergence of finite-dimensional distributions of $\widetilde{U}_{n1}(s)$ to those of $\tilde{\xi}$ is clearly true by the classical multivariate central limit theorem, since $\widetilde{U}_{n1}$ is a sum of independent random variables. It remains to show tightness for $\widetilde{U}_{n1}$, or equivalently, tightness for

$$\widetilde{U}_{n1}(s) = \sqrt{\frac{h}{n}} \sum_{i=1}^{n} \int_0^s K_h(u-t)\boldsymbol{a}^{\mathrm{T}}(\boldsymbol{u}-\boldsymbol{t})\Big\{z_i - \frac{S_1(u,\boldsymbol{\beta}^*)}{S_0(u,\boldsymbol{\beta}^*)}\Big\}dM_i(u).$$

By Lemma 5, $\big\{(h/n)^{1/2}\sum_{i=1}^n \int_0^s K_h(u-t)\boldsymbol{a}^{\mathrm{T}}(\boldsymbol{u}-\boldsymbol{t})z_i dM_i(u), s \in T\big\}$ is tightness. And analogous to the proof of Lemma 5, we can check that $\big\{(h/n)^{1/2}\sum_{i=1}^n \int_0^s K_h(u-t)\boldsymbol{a}^{\mathrm{T}}(\boldsymbol{u}-\boldsymbol{t})S_1(u,\boldsymbol{\beta}^*)/S_0(u,\boldsymbol{\beta}^*)dM_i(u), s \in T\big\}$ is tightness, too. Therefore, $\widetilde{U}_{n1}(s)$ converges to $\tilde{\xi}$. Hence, $\boldsymbol{a}^{\mathrm{T}}(nh)^{1/2}U_{n,1}(\tau)$ is normal. Then, $(nh)^{1/2}U_{n,1}(\tau)$ is multivariate normal, and asymptotically covariance is as follows:

$\Sigma_2(t) = \int K^2(v)\boldsymbol{v}^{\otimes 2}dvE(p_1(t \mid z)p_2(t \mid z)\mu_0^2(t)\exp(2\beta(t)z)(z-q_1(t)/q_0(t))^2)$

$= \sigma_2(t)\Omega_2$,

where $\Omega_2 = \int K^2(v)\boldsymbol{v}^{\otimes 2}dv$.

Therefore,

$\sqrt{nh}\Big\{\boldsymbol{H}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^*)-\Omega_1^{-1}\boldsymbol{b}h^{p+1}\beta^{(p+1)}(t)/(p+1)!\Big\} \to N\left(0, \sigma_1^{-2}(t)\sigma_2(t)\Omega_1^{-1}\Omega_2\Omega_1^{-1}\right)$,

as $n \to \infty$, $h \to 0$, $nh \to \infty$. □

S4.3 Proof the consistency of covariance

Proof. As defined in main document, $\widehat{\Sigma}(t) = \widehat{\Sigma}_1^{-1}(t)\widehat{\Sigma}_2(t)\widehat{\Sigma}_1^{-1}(t)$.

Here, we will show that $\widehat{\Sigma}_1(t)$ and $\widehat{\Sigma}_2(t)$ are consistent, respectively.

First of all, we give a conclusion by the following demonstration. Under $C2$

to $C4$, there exists a neighborhood $\mathcal{B}$ of $\boldsymbol{\beta}^*$, such that functions $S_j(u,\boldsymbol{\beta}), j=0,1,2,$ are continuous in $\boldsymbol{\beta}\in\mathcal{B}$, uniformly in $u\in T$. And $S_0(u,\boldsymbol{\beta})$ is bounded away of from zero on $(u,\boldsymbol{\beta})\in T\times\mathcal{B}$. Furthermore, by Lemma 1, we can derive, for each $j=0,1,2,$

$$\sup_{\mathcal{B}\times T}\|S_{n,j}(u,\boldsymbol{\beta})-S_j(u,\boldsymbol{\beta})\|\to 0,\quad as\quad n\to\infty. \tag{S4.43}$$

Further, account for $\widehat{\Sigma}_1(t)$. We will prove $\widehat{\Sigma}_1(t)$ converges to $\Sigma_1(t)=\sigma_1(t)\Omega_1$. Let

$$v_1(u,\beta(u))=S_2^*(u,\beta(u))/S_0^*(u,\beta(u))-S_1^{*2}(u,\beta(u))/S_0^{*2}(u,\beta(u)), \tag{S4.44}$$

and from the defined $q_j(t)$, we have $S_j^*(t,\beta(t))=q_j(t)/\mu_0(t)$, $j=0,1,2.$ Then, we obtain

$$\begin{aligned}\Sigma_1(t)&=\big\{q_2(t)-q_1^2(t)/q_0(t)\big\}\Omega_1\\&=\mu_0(t)\big\{S_2^*(t,\beta(t))-S_1^{*2}(t,\beta(t))/S_0^*(t,\beta(t))\big\}\Omega_1\\&=\int K_h(u-t)(\boldsymbol{u}-\boldsymbol{t})^{\otimes 2}\mu_0(u)S_0^*(u,\beta(u))v_1(u,\beta(u))du+o(1).\end{aligned} \tag{S4.45}$$

Using triangle inequality, we have

$\|\widehat{\Sigma}_1(t)-\Sigma_1(t)\|$

$\leq\|\frac{1}{n}\sum_{i=1}^n\int_0^\tau I(C_i\geq u)K_h(u-t)(\boldsymbol{u}-\boldsymbol{t})^{\otimes 2}\big\{V_1(u,\widehat{\boldsymbol{\beta}})-v_1(u,\beta(u))\big\}d\widetilde{N}_i(u)\|$

$+\|\frac{1}{n}\sum_{i=1}^{n}\int_0^{\tau} I(C_i \geq u)K_h(u-t)(\boldsymbol{u}-\boldsymbol{t})^{\otimes 2}v_1(u,\beta(u))\big\{d\widetilde{N}_i(u)-\mu_0(u)\exp(\beta(u)$

$z_i)o_i(u)du\big\}\|$

$+\|\frac{1}{n}\sum_{i=1}^{n}\int_0^{\tau} I(C_i \geq u)K_h(u-t)(\boldsymbol{u}-\boldsymbol{t})^{\otimes 2}v_1(u,\beta(u))\mu_0(u)\exp(\beta(u)z_i)o_i(u)du$

$-\int_0^{\tau} K_h(u-t)(\boldsymbol{u}-\boldsymbol{t})^{\otimes 2}\mu_0(u)S_0^*(u,\beta(u))v_1(u,\beta(u))du\|$

$+\|\int_0^{\tau} K_h(u-t)(\boldsymbol{u}-\boldsymbol{t})^{\otimes 2}\mu_0(u)S_0^*(u,\beta(u))v_1(u,\beta(u))du-\mu_0(t)\big\{S_2^*(t,\beta(t))-$

$S_1^{*2}(t,\beta(t))/S_0^*(t,\beta(t))\big\}\Omega_1\|.$

For the first term of the right-hand side, under $C1$ to $C5$, by the consequence of Theorem 1, we can derive

$$\sup_{\mathcal{B}\times T}\|S_j(u,\widehat{\boldsymbol{\beta}})-S_j^*(u,\beta(u))\|\to 0, \quad as \quad n\to\infty. \tag{S4.46}$$

Hence, from (S4.43) and (S4.46), we have

$$\sup_{\mathcal{B}\times T}\|V_1(u,\widehat{\boldsymbol{\beta}})-v_1(u,\beta(u))\|\to 0, \quad as \quad n\to\infty. \tag{S4.47}$$

By consequence of Lenglart inequality,

$$pr\left(\left\{\frac{1}{n}\sum_{i=1}^{n}\int_0^{\tau} I(C_i \geq u)K_h(u-t)d\widetilde{N}_i(u) > C\right\}\right) \leq$$
$$\frac{\delta}{C}+pr\left(\left\{\int_0^{\tau}\frac{1}{n}\sum_{i=1}^{n} I(C_i \geq u)K_h(u-t)\mu_0(u)\exp(\beta(u)z_i)o_i(u)du > \delta\right\}\right), \tag{S4.48}$$

when $\delta > \int_0^\tau K_h(u-t)\mu_0(u)S_0^*(u,\beta(u))du = \mu_0(t)S_0^*(t,\beta(t))$, the latter probability tends to zero as $n \to \infty, h \to 0$, and $nh \to \infty$. Thus the first term converges to zero.

For the second term of the right-hand side,

$$\frac{1}{n}\sum_{i=1}^{n}\int_0^\tau K_h(u-t)v_1(u,\beta(u))dM_i(u)$$

is empirical process. by Lemma 5 and $v_1(u,\beta(u))$ is non-negative function, analogous to the proof of Theorem 2, using the Theorem 8.3 (the uniform law of large numbers) of Pollard (1990), we can demonstrate the second term converges to zero.

For the third term of the right-hand side, under $C1$ to $C4$, functions $v_1(u,\beta(u))$ are bounded. So from (S4.43), it is easy to prove the third term tend to zero.

For the fourth term of the right-hand side, from (S4.45), obviously, it converges to zero.

Therefore,

$$\|\widehat{\Sigma}_1(t) - \Sigma_1(t)\| \to 0, \quad as \quad n \to \infty. \tag{S4.49}$$

Next, we will prove $\widehat{\Sigma}_2(t)$ converges to $\Sigma_2(t) = \sigma_2(t)\Omega_2$ by the following demonstration. Let

$$v_2(u,\beta(u)) = \left\{z - S_1^*(u,\beta(u))/S_0^*(u,\beta(u))\right\}^2. \tag{S4.50}$$

For

$$\begin{aligned}\Sigma_2(t) =& E\left(p_1(t \mid z)p_2(t \mid z)\mu_0^2(t)\exp(2\beta(t)z)\left\{z - q_1(t)/q_0(t)\right\}^2\right)\Omega_2 \\ =& \int_0^\tau hK_h^2(u-t)(\boldsymbol{u}-\boldsymbol{t})^{\otimes 2}\mu_0^2(u)E\big(p_1(u \mid z)p_2(u \mid z)\exp(2\beta(u)z) \\ & v_2(u,\beta(u))\big)du + o(1).\end{aligned} \tag{S4.51}$$

Then, using triangle inequality, we have

$$\begin{aligned}&\|\widehat{\Sigma}_2(t) - \Sigma_2(t)\| \leq \\ &\|\frac{1}{n}\sum_{i=1}^n \int_0^\tau hK_h^2(u-t)(\boldsymbol{u}-\boldsymbol{t})^{\otimes 2}I(C_i \geq u)V_2(u,\widehat{\boldsymbol{\beta}})\widehat{\mu}_0^2(u,\widehat{\beta}(u))\exp(2\widehat{\beta}(t) \\ &z_i)o_i(u)du - \int_0^\tau hK_h^2(u-t)(\boldsymbol{u}-\boldsymbol{t})^{\otimes 2}\widehat{\mu}_0^2(u,\widehat{\beta}(u))E(p_1(u \mid z)p_2(u \mid z)v_2(\\ &u,\beta(u))\exp(2\beta(t)z))du\| \\ &+\|\int_0^\tau hK_h^2(u-t)(\boldsymbol{u}-\boldsymbol{t})^{\otimes 2}\left\{\widehat{\mu}_0^2(u,\widehat{\beta}(u)) - \mu_0^2(u)\right\}E(p_1(u \mid z)p_2(u \mid z) \\ &v_2(u,\beta(u))\exp(2\beta(t)z))du\| \\ &+\|\int_0^\tau hK_h^2(u-t)(\boldsymbol{u}-\boldsymbol{t})^{\otimes 2}\mu_0^2(u)E(p_1(u \mid z)p_2(u \mid z)v_2(u,\beta(u))\exp(2 \\ &\beta(t)z))du - E(p_1(t \mid z)p_2(t \mid z)\mu_0^2(t)\exp(2\beta(t)z)\left\{z - q_1(t)/q_0(t)\right\}^2)\Omega_2\|.\end{aligned}$$

For the first term of the right-hand side, let $V_{21}(u,\widehat{\boldsymbol{\beta}}) = S_{n,1}(u,\widehat{\boldsymbol{\beta}})/S_{n,0}(u,\widehat{\boldsymbol{\beta}})$, and $v_{21} = S_1^*(u,\beta(u))/S_0^*(u,\beta(u))$. We have

$$\frac{1}{n}\sum_{i=1}^{n}\int_0^\tau hK_h^2(u-t)(\boldsymbol{u}-\boldsymbol{t})^{\otimes 2}I(C_i \geq u)V_2(u,\widehat{\boldsymbol{\beta}})\widehat{\mu}_0^2(u,\widehat{\beta}(u))\exp(2\widehat{\beta}(t)z_i)$$
$$o_i(u)du = \int_0^\tau hK_h^2(u-t)(\boldsymbol{u}-\boldsymbol{t})^{\otimes 2}\widehat{\mu}_0^2(u,\widehat{\beta}(u))\big\{S_{n,2}^*(u,2\widehat{\beta}(t)) - 2S_{n,1}^*(u,$$
$$2\widehat{\beta}(t))V_{21}(u,\widehat{\boldsymbol{\beta}}) + S_{n,0}^*(u,2\widehat{\beta}(t))V_{21}^2(u,\widehat{\boldsymbol{\beta}})\big\}du. \tag{S4.52}$$

From (S4.43), we can derive, for $j = 0,1,2$,

$$\sup_{\mathcal{B}\times T}\|S_{n,j}(u,2\widehat{\beta}(t)) - S_j(u,2\beta(t))\| \to 0, \quad as \quad n\to\infty. \tag{S4.53}$$

Analogous to the proof of $V_1(u,\widehat{\boldsymbol{\beta}})$, we can obtain

$$\sup_{\mathcal{B}\times T}\|V_{21}(u,\widehat{\boldsymbol{\beta}}) - v_{21}(u,\beta(u))\| \to 0, \quad as \quad n\to\infty, \tag{S4.54}$$

and

$$\sup_{\mathcal{B}\times T}\|V_{21}^2(u,\widehat{\boldsymbol{\beta}}) - v_{21}^2(u,\beta(u))\| \to 0, \quad as \quad n\to\infty. \tag{S4.55}$$

Hence, from (S4.52), (S4.53), (S4.54) and (S4.55), the convergence of the first term can be demonstrated.

For the second term of the right-hand side, let

$$\widehat{\mu}_0^2(u,\widehat{\beta}(u))-\mu_0^2(u)=\big\{\widehat{\mu}_0(u,\widehat{\beta}(u))+\mu_0(u)\big\}\big\{\widehat{\mu}_0(u,\widehat{\beta}(u))-\mu_0(u)\big\}, \quad \text{(S4.56)}$$

and

$$\begin{aligned}\widehat{\mu}_0(u,\widehat{\beta}(u))-\mu_0(u) &= \big\{\widehat{\mu}_0(u,\widehat{\beta}(u))-\widehat{\mu}_0(u,\beta(u))\big\}+\big\{\widehat{\mu}_0(u,\beta(u))-\mu_0(u)\big\}\\ &= H_n(u,\beta^*)\big\{\widehat{\beta}(u)-\beta(u)\big\}+\frac{1}{n}\sum_{i=1}^n dM_i(u)/S_{n,0}^*(u,\beta(u)), \quad \text{(S4.57)}\end{aligned}$$

where $H_n(u,\beta^*)=-S_{n,1}^*(u,\beta^*)\sum_{i=1}^n I(C_i\geq u)N_i(u)o_i(u)/n(S_{n,0}^*(u,\beta^*))^2$, and β^* is between $\beta(u)$ and $\widehat{\beta}(u)$.

Under $C1$ to $C5$, by Lemma 1, $H_n(u,\beta^*)$ is bounded, and $\widehat{\beta}(u)$ uniform converges to $\beta(u)$. Hence, we can derive

$$\sup_{u\in T}\|H_n(u,\beta^*)(\widehat{\beta}(u)-\beta(u))\|\to 0, \quad as \quad n\to\infty. \quad \text{(S4.58)}$$

For $\frac{1}{n}\sum_{i=1}^n dM_i(u)/S_{n,0}^*(u,\beta(u))$ is empirical process, analogous to the proof of Theorem 2, using Theorem 8.3 (the uniform law of large numbers) of Pollard (1990), we can obtain

$$\sup_{u\in T}\|\frac{1}{n}\sum_{i=1}^n dM_i(u)/S_{n,0}^*(u,\beta(u))\|\to 0, \quad as \quad n\to\infty. \quad \text{(S4.59)}$$

Under $C1$ to $C5$, $\widehat{\mu}_0(u, \widehat{\beta}(u)) + \mu_0(u)$ is bounded, in conjunction with (S4.56), (S4.57), (S4.58) and (S4.59), we obtain

$$\sup_{u \in T} \|\widehat{\mu}_0^2(u, \widehat{\beta}(u)) - \mu_0^2(u)\| \to 0, \quad as \quad n \to \infty. \tag{S4.60}$$

Therefore, the second term converges to zero.

For the third term of the right-hand side, from (S4.51), obviously, converges to zero.

Hence,

$$\|\widehat{\Sigma}_2(t) - \Sigma_2(t)\| \to 0, \quad as \quad n \to \infty. \tag{S4.61}$$

Therefore, from (S4.49) and (S4.61), we have $\widehat{\Sigma}(t)$ is consistent. □

S4.4 Proof of the asymptotic normality of $\widehat{\mu}_0(t, \widehat{\beta}(t))$

Proof. Let

$$\begin{aligned}
&\sqrt{nh}\Big\{\widehat{\mu}_0(t, \widehat{\beta}(t)) - \mu_0(t)\Big\} \\
&= \sqrt{nh}\Big\{\widehat{\mu}_0(t, \widehat{\beta}(t)) - \widehat{\mu}_0(t, \beta(t))\Big\} + \sqrt{nh}\Big\{\widehat{\mu}_0(t, \beta(t)) - \mu_0(t)\Big\} \\
&= H_n(t, \beta^*)\sqrt{nh}\Big\{\widehat{\beta}(t) - \beta(t)\Big\} + \sqrt{nh}\sum_{i=1}^{n} dM_i(t)/nS_{n,0}^*(t, \beta(t)).
\end{aligned} \tag{S4.62}$$

where $H_n(t, \beta^*) = -S_{n,1}^*(t, \beta^*)\sum_{i=1}^{n} I(C_i \geq t)N_i(t)o_i(t)/nS_{n,0}^{*2}(t, \beta^*)$, and β^* is between $\widehat{\beta}(t)$ and $\beta(t)$.

For the second term of right-hand side of (S4.62), let

$$\begin{aligned}&\sqrt{nh}\sum_{i=1}^{n}\frac{dM_i(t)}{nS_{n,0}^{*}(t,\beta(t))}\\&=\sqrt{nh}\sum_{i=1}^{n}\frac{dM_i(t)}{nS_0^{*}(t,\beta(t))}+\sqrt{\frac{h}{n}}\sum_{i=1}^{n}\Big\{S_{n,0}^{*-1}(t,\beta(t))-S_0^{*-1}(t,\beta(t))\Big\}dM_i(t).\end{aligned} \tag{S4.63}$$

Define

$$U_3(s)=\sqrt{\frac{h}{n}}\sum_{i=1}^{n}dM_i(s). \tag{S4.64}$$

Clearly, $E[U_3(s)]=0$.

Analogous to the proof of Lemma 5, using the functional central limit theorem of Pollard (1990), we shall argument U_3 converges to Gaussian process ξ_3. Now we test the conditions (i)-(v) in Pollard (1990).

Under $C1$ to $C5$, by Lemma 1, we have, for any $s_1, s_2 \in T$,

$$\lim_{n\to\infty} E\left(U_3(s_1)U_3(s_2)\right)$$
$$= \lim_{n\to\infty} E\left(\frac{h}{n}\sum_{i=1}^{n}\Big\{dM_i(s_1)dM_i(s_2)\Big\}\right)$$
$$= hE\Big(\big\{I(C \geq s_1)\mu_0(s_1)\exp(\beta(s_1)z)o(s_1)\big\}\big\{I(C \geq s_2)\mu_0(s_2)\exp(\beta(s_2)z)$$
$$o(s_2)\big\}\Big) + o_p(1)$$
$$= \begin{cases} hE\left(p_1(s \mid z)p_2(s \mid z)\mu_0^2(s)\exp(2\beta(s)z)\right) + o_p(1), & s_1 = s_2 = s, \\ 0, & s_1 \neq s_2. \end{cases}$$

Then, by the classical multivariate central limit theorem for independent random vectors, the finite-dimensional distributions of U_3 converge to those of gaussian process ξ_3, which converge to zero, as h gets to zero. Thus condition (ii) holds. Next, checking the tightness. Under $C1$ to $C5$, we know $\big\{I(C_i \geq t)N_i(t)o_i(t), t \in T\big\}$ has finite points, and $\exp(\beta(t)z_i)o_i(t)$ are bounded variation functions. Thus, $\big\{dM_i(t), t \in T\big\}$ is manageable, and the envelops can be chosen as constant $\bar{B}/\sqrt{n}$, then (i)(iii)(iv) holds. To verify (v), for any $s_1, s_2 \in T$, define

$$\rho_n(s_1, s_2) = E\left(U_3(s_1) - U_3(s_2)\right)^2, \quad \rho(s_1, s_2) = E\left(\xi_3(s_1) - \xi_3(s_2)\right)^2.$$

Further,

$$\rho_n(s_1, s_2) = E\left(\sqrt{\frac{h}{n}}\sum_{i=1}^{n}\Big\{dM_i(s_1) - dM_i(s_2)\Big\}\right)^2$$
$$= \frac{h}{n}\sum_{i=1}^{n} E\big(I(C_i \geq s_1)\mu_0^2(s_1)\exp(2\beta(s_1)z_i)o_i(s_1) + I(C_i \geq s_2)\mu_0^2(s_2)\exp$$
$$(2\beta(s_2)z_i)o_i(s_2)\big).$$

Clearly, $\{\rho_n\}$ is equicontinuous on T, and $\lim_{n\to\infty}\rho_n(s_1, s_2) = \rho(s_1, s_2), \rho$ is pseudometric on T. Thus ρ_n converges, uniformly on T, to ρ. And, let $\{s_1^{(n)}\}, \{s_2^{(n)}\}$ be any two sequence in T, it follows that if $\rho(s_1^{(n)}, s_2^{(n)}) \to 0$, then $\rho_n(s_1^{(n)}, s_2^{(n)}) \to 0$, then (v) holds. Therefore, U_3 converges in distribution to Gaussian process on T, and covariance matrix is diagonal matrix, and the matrix is zero, when h gets to zero. That is, U_3 converges in distribution zero, as $n \to \infty$, $h \to 0$ and $nh \to \infty$. Moreover, using the Strong Representation Theorem of Pollard (1990), we have a new probability space and

$$\sup_{s\in T}\|U_3(s) - 0\| \to 0, \quad as \quad n \to \infty. \tag{S4.65}$$

By Lemma 1 and $C1$ to $C5$, we can obtain

$$\sup_{s\in T}\|S_{n,0}^{*-1}(t, \beta(t)) - S_0^{*-1}(t, \beta(t))\| \to 0, \quad as \quad n \to \infty. \tag{S4.66}$$

Then, by Lemma 2 combined with (S4.65) and (S4.66), we can derive, in probability,

$$(h/n)^{1/2}\sum_{i=1}^{n}\Big\{S_{n,0}^{*-1}(t,\beta(t))-S_{0}^{*-1}(t,\beta(t))\Big\}dM_i(t)\to 0, \quad as \quad n\to\infty. \tag{S4.67}$$

which holds in the original probability space. And in analogy with the prove of $U_3(s)$, we can check that, in distribution,

$$(nh)^{1/2}\sum_{i=1}^{n}dM_i(t)/nS_0^{*}(t,\beta(t))\to 0, \quad as \quad n\to\infty. \tag{S4.68}$$

Therefore, from (S4.63), (S4.67) and (S4.68), we can obtain, in probability,

$$(nh)^{1/2}\sum_{i=1}^{n}dM_i(t)/nS_{n,0}^{*}(t,\beta(t))\to 0, \quad as \quad n\to\infty. \tag{S4.69}$$

For the first term of right-hand side of (S4.62). Under $C1$ to $C5$ and Lemma 1, we obtain

$$H_n(t,\beta^*)\to -q_1(t)/q_0(t), \quad n\to\infty. \tag{S4.70}$$

By the assumption of $nh^5 = o(1)$, we can derive

$$(nh)^{1/2}\big\{\widehat{\beta}(t) - \beta(t)\big\} \to N(0, \nu_0\sigma_1^{-2}(t)\sigma_2(t)), \quad as \quad n \to \infty. \tag{S4.71}$$

Therefore, from (S4.70) and (S4.71), we have

$$H_n(t, \beta^*)(nh)^{1/2}\big\{\widehat{\beta}(t)-\beta(t)\big\} \to N(\nu_0 q_0^{-2}(t)q_1(t)\sigma_1^{-2}(t)\sigma_2(t)), \quad as \quad n \to \infty. \tag{S4.72}$$

Hence, form (S4.62), (S4.69) and (S4.72), using Slutsky's theorem, we obtain

$$(nh)^{1/2}\Big\{\widehat{\mu}_0(t, \widehat{\beta}(t)) - \mu_0(t)\Big\} \to N(0, \Sigma_3(t)), \quad n \to \infty, \tag{S4.73}$$

where $\Sigma_3(t) = \nu_0 q_0^{-2}(t)q_1(t)\sigma_1^{-2}(t)\sigma_2(t)$. □

S5 Additional Simulations

ere, we show the simulation results about the local kernel estimators $\widehat{\beta}(t)$ with corresponding setting that $\beta(t)$=0·5$\big\{$Beta(t/12,4,4)+Beta(t/12,5,5)$\big\}$ and $\beta(t) = \sin(\pi t/6)$, respectively, under sample sizes equal to 500. As the following figures show.

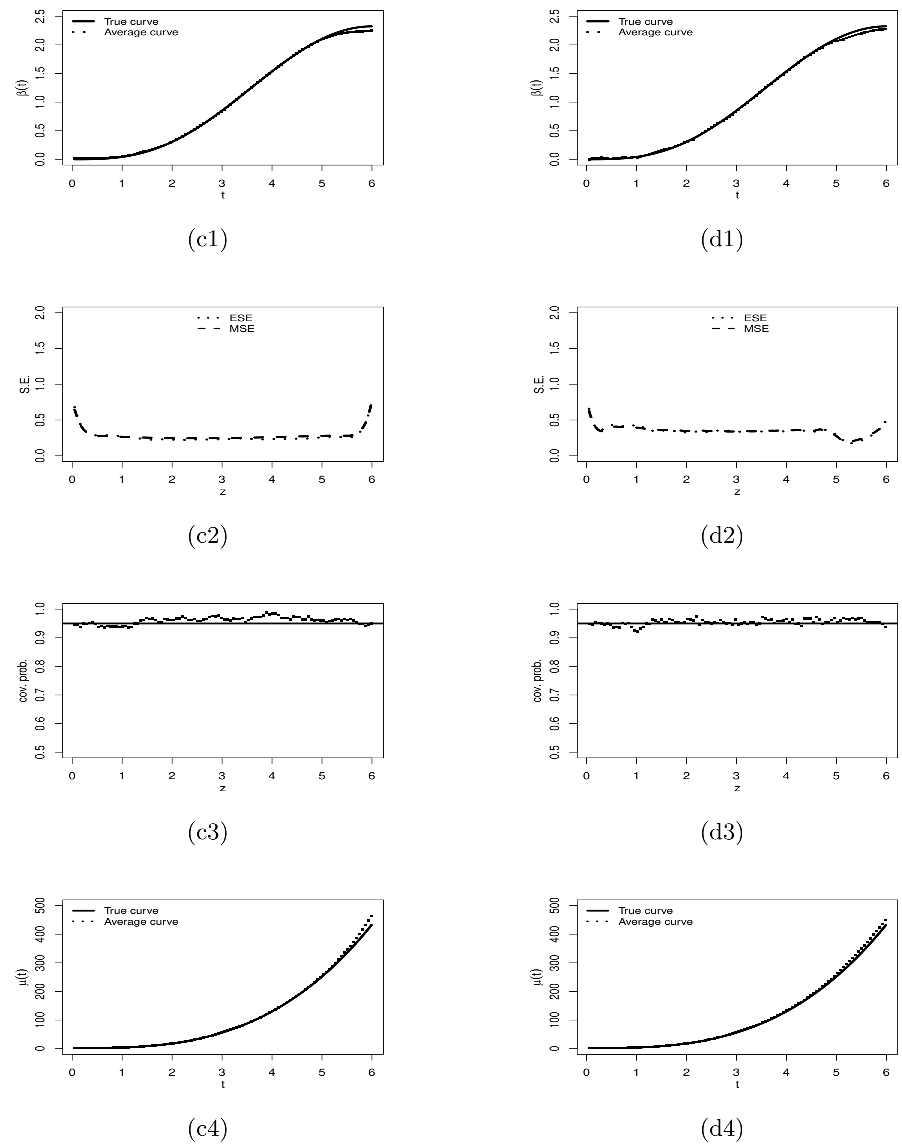
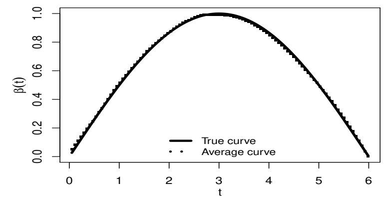
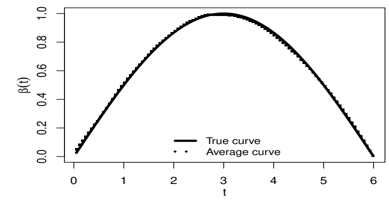


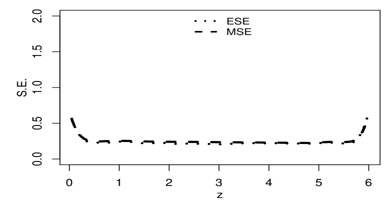
Figure 1: (c1) and (d1): The true and the average of the local kernel estimator, with h=0·5 and h_{cv}, respectively. (c2) and (d2): Comparison of empirical standard errors (ESE) and the estimated standard errors (MSE) of $\widehat{\beta}(t)$, with h=0·5 and h_{cv}, respectively. (c3) and (d3): Empirical coverage probabilities of the 95% confidence intervals for $\widehat{\beta}(t)$, with h=0·5 and h_{cv}, respectively. (c4) and (d4): Comparison of the true baseline curve and the average of the Breslow-type estimator, with h=0·5 and h_{cv}, respectively.



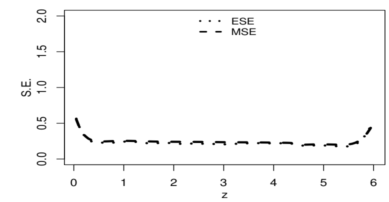
(c1)



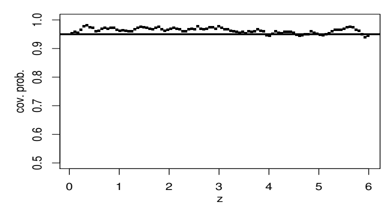
(d1)



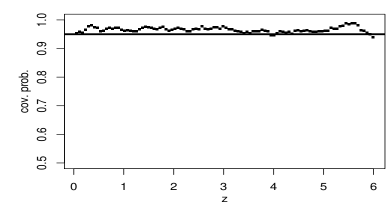
(c2)



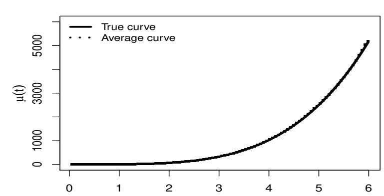
(d2)



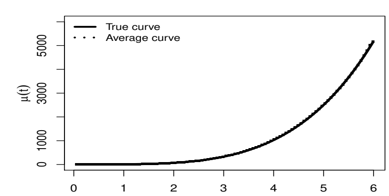
(c3)



(d3)



(c4)



(d4)

Figure 2: (c1) and (d1): The true and the average of the local kernel estimator, with h=0·5 and h_{cv}, respectively. (c2) and (d2): Comparison of empirical standard errors (ESE) and the estimated standard errors (MSE) of $\widehat{\beta}(t)$, with h=0·5 and h_{cv}, respectively. (c3) and (d3): Empirical coverage probabilities of the 95% confidence intervals for $\widehat{\beta}(t)$, with h=0·5 and h_{cv}, respectively. (c4) and (d4): Comparison of the true baseline curve and the average of the Breslow-type estimator, with h=0·5 and h_{cv}, respectively.

Bibliography

Bilias, Y., M. Gu, and Z. Ying (1997). Towards a general asymptotic theory for cox model with staggered entry. *The Annals of Statistics 25*(2), 662–682.

Fan, J. C. C. and R. Gijbels (1996). *Local polynomial modelling and its applications*. Chapman and Hall.

Hardle, W., P. Janssen, and R. Serfling (1988). Strong uniform consistency rates for estimators of conditional functionals. *The Annals of Statistics 16*(4), 1428–1449.

Pollard, D. (1990). *Empirical processes: Theory and Applications*, Volume 2. Paul Shaman, Jessica Utts and Jose L. Gonzalez.

Zhao, P.-L. (1994). Asymptotics of kernel estimators based on local maximum likelihood. *Journaltitle of Nonparametric Statistics 4*(1), 79–90.