Supplementary Document for the Manuscript entitled "A Functional Single Index Model"

S1 Definition and Regularity Conditions

Let $\mathcal{H}$ be the separable Hilbert space of square integrable functions on $[0,1]$, and let $<\phi_1,\phi_2>\equiv\int_0^1\phi_1(t)\phi_2(t)dt$ denote the inner product of the functions ϕ_1,ϕ_2 in $\mathcal{H}$. Then the L_2 norm $\|\cdot\|_2$ is the norm induced by the inner product. Let $\alpha(\cdot)$ be a function in $\mathcal{H}$. Let $\boldsymbol{\gamma}$ be the B-spline coefficient such that $\mathbf{B}_r(\cdot)^{\mathrm{T}}\boldsymbol{\gamma}$ converges to $\alpha(\cdot)$, uniformly on [0, 1] when the number of the B-spline inner knots goes to infinite, and $\mathbf{B}_r(\cdot)^{\mathrm{T}}\boldsymbol{\gamma}_0$ converges to $\alpha_0(\cdot)$, the true function. Here $\mathbf{B}_r=(B_{r1},\ldots,B_{rd_{\boldsymbol{\gamma}}})^{\mathrm{T}}$ is the rth order B-spline basis vector. Note that for identifiability, we assumed $\alpha_0(0)=\mathbf{B}_r(0)^{\mathrm{T}}\boldsymbol{\gamma}_0=1$, i.e., $\gamma_1=B_{r1}(0)^{-1}\{1-\sum_{k=2}^{d_{\boldsymbol{\gamma}}}B_{rk}(0)\gamma_k\}=1$. So we only have $d_{\boldsymbol{\gamma}}-1$ B-spline coefficients to estimate, and we denote $\boldsymbol{\gamma}^-=(\gamma_2,\ldots,\gamma_{d_{\boldsymbol{\gamma}}})^{\mathrm{T}}$. Further, we denote $\boldsymbol{\alpha}(\mathbf{X})=\int_0^1\mathbf{X}(t)\alpha(t)dt$ and $\boldsymbol{\alpha}_0(\mathbf{X})=\int_0^1\mathbf{X}(t)\alpha_0(t)dt$. We denote the functional sup norm as $\|\cdot\|_\infty$, and the L_p norm as $\|\cdot\|_p$.

In addition, we define $\Gamma(\boldsymbol{\beta})$ to be a second moment based linear operator such that

$$\begin{aligned}\Gamma(\boldsymbol{\beta})\phi(t) &\equiv \int_0^1 E\{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}(t)\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}(s)\}\phi(s)ds\\ &= E\{<\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X},\phi>\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}(t)\}\end{aligned}$$

while its empirical version can be written as

$$\Gamma_n(\boldsymbol{\beta})\phi(t)\equiv n^{-1}\sum_{i=1}^n<\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i,\phi>\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i(t).$$

Using these definitions, we can write

$$<\Gamma(\boldsymbol{\beta})B_{rk},B_{rl}> \;=\; E\left(\int_0^1B_{rk}(t)\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i(t)dt\int_0^1B_{rl}(t)\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i(t)dt\right),$$

$$< \Gamma_n(\boldsymbol{\beta})B_{rk}, B_{rl} > \;=\; n^{-1}\sum_{i=1}^{n}\int_0^1 B_{rk}(t)\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i(t)dt\int_0^1 B_{rl}(t)\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i(t)dt.$$

Define $\mathbf{C}(\boldsymbol{\beta})$ as a $d_{\boldsymbol{\gamma}}\times d_{\boldsymbol{\gamma}}$ matrix with its (k,l) element $< \Gamma(\boldsymbol{\beta})B_{rk}, B_{rl} >$, and $\widehat{\mathbf{C}}(\boldsymbol{\beta})$ as a $d_{\boldsymbol{\gamma}}\times d_{\boldsymbol{\gamma}}$ matrix matrix with its (k,l) element $< \Gamma_n(\boldsymbol{\beta})B_{rk}, B_{rl} >$. The locally efficient score function for $\boldsymbol{\beta},\boldsymbol{\gamma}$ can be written as

$$\begin{aligned}
&\mathbf{S}_{\mathrm{eff}\boldsymbol{\beta}}(Y,\mathbf{Z},\boldsymbol{\beta},\boldsymbol{\gamma},g^*)\\
=\;&[g^*(Y,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma})-E\{g^*(Y,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma})|\boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma}\}]\boldsymbol{\Theta}_{\beta}\{\mathbf{Z}-E(\mathbf{Z}\mid\boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma})\}\boldsymbol{\gamma},\\
&\mathbf{S}_{\mathrm{eff}\boldsymbol{\gamma}}(Y,\mathbf{Z},\boldsymbol{\beta},\boldsymbol{\gamma},g^*)\\
=\;&[g^*(Y,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma})-E\{g^*(Y,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma})|\boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma}\}]\boldsymbol{\Theta}_{\gamma}\{\mathbf{Z}-E(\mathbf{Z}\mid\boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma})\}^{\mathrm{T}}\boldsymbol{\beta}.
\end{aligned}$$

Here $g^*(Y,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma}) = f_2^{*'}(Y,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma})/f^*(Y,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma})$, $f^*(Y,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma})$ is the possibly misspecified conditional density of Y given $\boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma}$ and $f_2^{*'}(Y,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma})$ is the derivative of $f^*(Y,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma})$ with respect to $\boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma}$. When f is correctly specified, i.e. when $f^* = f$, we denote the resulting g^* as g. $\boldsymbol{\Theta}_{\boldsymbol{\beta}} = (\mathbf{I}_{J-1},\mathbf{0})$ and $\boldsymbol{\Theta}_{\boldsymbol{\gamma}} = (\mathbf{0},\mathbf{I}_{d_{\gamma}-1})$. Note that when $f^*(Y,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma}) = f(Y,\boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma})$, the true density function, the locally efficient score is the efficient score of (2). Throughout the text, $\mathbf{A}_1 = o_p(\mathbf{A}_2)$ for arbitrary vectors or matrices $\mathbf{A}_1$, $\mathbf{A}_2$ means that $\mathbf{A}_1$ has smaller order than $\mathbf{A}_2$ component wise.

We first list the regularity conditions under which we perform our theoretical analysis.

(A1) The kernel function $K(\cdot)$ is non-negative, has compact support, and satisfies $\int K(s)ds = 1$, $\int K(s)sds = 0$ and $\int K(s)s^2ds < \infty$, and $\int K^2(s)ds < \infty$.

(A2) The bandwidth h in the kernel smoothing satisfies $nh^2 \to \infty$ and $nh^8 \to 0$ when $n \to \infty$.

(A3) Assume $\alpha_0 \in \{\alpha \in C^q([0,1])$, α is one-to-one, and $\alpha_0(0) = 1\}$. The spline order $r \geq q$.

(A4) We define the knots $t_{-r+1} \leq \ldots \leq t_0 = 0$ and $t_{N+r} \geq \ldots \geq t_{N+1} = 1$. Let N be the number of interior knots and $[0,1]$ be divided into $N+1$

subintervals. In this case, $d_{\boldsymbol{\gamma}} = N + r$. Let h_p be the distance between the $(p+1)$th and pth interior knots of the order r B-spline functions and let $h_b = \max_{r \le p \le N+r} h_p$. There exists a constant C_{h_n}, $0 < c_{h_b} < \infty$, such that

$$\max_{r \le p \le N+r} h_p = O(N^{-1}) \text{ and } h_b / \min_{r \le p \le N+r} h_p < c_{h_b},$$

where N is the number of knots which satisfies $N \to \infty$ as $n \to \infty$, and $N^{-1} n (\log n)^{-1} \to \infty$ and $N n^{-1/(2q+1)} \to \infty$.

(A5) $\boldsymbol{\gamma}_0$ is a spline coefficient such that $\sup_{t \in [0,1]} |\mathbf{B}_r(t)^{\mathrm{T}} \boldsymbol{\gamma}_0 - \alpha_0(t)| = O_p(h_b^q)$. The existence of such $\boldsymbol{\gamma}_0$ has been shown in De Boor (1978).

(A6) $X_j(\cdot), j = 1, \ldots, J$ are continuous random functions in $t \in [0,1]$. For each t, $|E\{X_j(t)\}| < \infty$. $\|\boldsymbol{\beta}^{\mathrm{T}} \mathbf{X}(\cdot)\|_2$ is finite almost surely for any bounded $\boldsymbol{\beta}$. $E\{\boldsymbol{\beta}^{\mathrm{T}} \mathbf{X}_i(\cdot) \mid \boldsymbol{\beta}^{\mathrm{T}} \mathbf{Z}_i \boldsymbol{\gamma}\}$ and its first two derivatives have finite L_2 norm almost surely for any bounded $\boldsymbol{\beta}, \boldsymbol{\gamma}$. The second moment operator $\Gamma(\boldsymbol{\beta})$ is strictly positive definite.

(A7)

$$E[\{\mathbf{S}_{\mathrm{eff}\boldsymbol{\beta}}(Y_i, \mathbf{Z}_i, \boldsymbol{\beta}, \boldsymbol{\gamma}, g^*)^{\mathrm{T}}, \mathbf{S}_{\mathrm{eff}\boldsymbol{\gamma}}(Y_i, \mathbf{Z}_i, \boldsymbol{\beta}, \boldsymbol{\gamma}, g^*)^{\mathrm{T}}\}^{\mathrm{T}}]$$

is a smooth function of $(\boldsymbol{\beta}^{\mathrm{T}}, \boldsymbol{\gamma}^{\mathrm{T}})^{\mathrm{T}}$ and has unique root for $\boldsymbol{\beta}, \boldsymbol{\gamma}$.

(A8) $f^*(Y, \boldsymbol{\beta}^{\mathrm{T}} \mathbf{Z} \boldsymbol{\gamma})$ is a continuous density function and has bounded derivative with respect to the second argument. $f^*(Y, \boldsymbol{\beta}^{\mathrm{T}} \mathbf{Z} \boldsymbol{\gamma})$ is bounded away from 0 and ∞ on its support. This implies $g^*(Y, \boldsymbol{\beta}^{\mathrm{T}} \mathbf{Z} \boldsymbol{\gamma})$ is a continuous bounded function and $g^*(Y, \boldsymbol{\beta}^{\mathrm{T}} \mathbf{Z} \boldsymbol{\gamma})$ is bounded away from 0 and ∞ on its support. Let $f_{\mathbf{Z}}(\boldsymbol{\beta}_0^{\mathrm{T}} \mathbf{Z}_i \boldsymbol{\gamma}_0)$ be the density for $\boldsymbol{\beta}_0^{\mathrm{T}} \mathbf{Z} \boldsymbol{\gamma}_0$. $f_{\mathbf{Z}}$ has two bounded derivatives, and $f_{\mathbf{Z}}$ is bounded away from 0 and ∞ on its support.

S2 Proofs of the propositions in the main text

S2.1 Proof of Proposition 1

Assume the problem is not identifiable. Then there exist $\{f, \boldsymbol{\beta}, \alpha(t)\} \neq \{\widetilde{f}, \widetilde{\boldsymbol{\beta}}, \widetilde{\alpha}(t)\}$ such that

$$f\left\{Y, \sum_{j=1}^{J} \beta_j \int_0^1 \alpha(t) X_j(t) dt\right\} = \widetilde{f}\left\{Y, \sum_{j=1}^{J} \widetilde{\beta}_j \int_0^1 \widetilde{\alpha}(t) X_j(t) dt\right\} \tag{S1}$$

for all Y and all functions $X_1(\cdot), \ldots, X_J(\cdot)$. Because (S1) holds for all $X_j(\cdot), j = 1, \ldots, J$, it holds when $X_j(t) = 0$ for $j = 1, \ldots, J-1$ as well. Note that $\beta_J = \widetilde{\beta}_J = 1$, we then have

$$f\left\{Y, \int_0^1 \alpha(t) X_J(t) dt\right\} = \widetilde{f}\left\{Y, \int_0^1 \widetilde{\alpha}(t) X_J(t) dt\right\}. \tag{S2}$$

Let $X_J(t) = a(t) + \delta b(t)$ and let $\delta \to 0$, this yields

$$\begin{aligned}
& f_2'\left\{Y, \int_0^1 \alpha(t) a(t) dt\right\} \int_0^1 \alpha(t) b(t) dt \\
= & \lim_{\delta \to 0} \left(f\left[Y, \int_0^1 \alpha(t)\{a(t) + \delta b(t)\} dt\right] - f\left\{Y, \int_0^1 \alpha(t) a(t) dt\right\}\right) / \delta \\
= & \lim_{\delta \to 0} \left(\widetilde{f}\left[Y, \int_0^1 \widetilde{\alpha}(t)\{a(t) + \delta b(t)\} dt\right] - \widetilde{f}\left\{Y, \int_0^1 \widetilde{\alpha}(t) a(t) dt\right\}\right) / \delta \\
= & \widetilde{f}_2'\left\{Y, \int_0^1 \widetilde{\alpha}(t) a(t) dt\right\} \int_0^1 \widetilde{\alpha}(t) b(t) dt, \qquad \text{(S3)}
\end{aligned}$$

where $f_2'(Y, \cdot), \widetilde{f}_2'(Y, \cdot)$ denote the partial derivatives of $f, \widetilde{f}$ with respect to the second component. Set $b(t)$ to be the delta function with mass at 0. Because $\alpha(0) = \widetilde{\alpha}(0) = 1$, this yields

$$f_2'\left\{Y, \int_0^1 \alpha(t) a(t) dt\right\} = \widetilde{f}_2'\left\{Y, \int_0^1 \widetilde{\alpha}(t) a(t) dt\right\}$$

for any $a(\cdot)$, hence subsequently from (S3),

$$\int_0^1 \alpha(t) b(t) dt = \int_0^1 \widetilde{\alpha}(t) b(t) dt$$

for any $b(\cdot)$. Because $\alpha(t), \widetilde{\alpha}(t)$ are continuous, the latter directly yields $\alpha(t) = \widetilde{\alpha}(t)$. From (S2), this further indicates $f(Y, \cdot) = \widetilde{f}(Y, \cdot)$. It is now easy to see from (S1) that by taking $X_1(t), \dots, X_{j-1}(t), X_{j+1}(t), \dots, X_J(t)$ to be zero, we can obtain $\beta_j = \widetilde{\beta}_j$ for $j = 1, \dots, J-1$. This contradicts the assumption $\{f, \boldsymbol{\beta}, \alpha(t)\} \neq \{\widetilde{f}, \widetilde{\boldsymbol{\beta}}, \widetilde{\alpha}(t)\}$. Hence the model in (1) is indeed identifiable. $\square$

S2.2 Proof of Proposition 2

Recall that we can write $\boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}$ as $\int_0^1 \mathbf{B}_r(t)^{\mathrm{T}}\boldsymbol{\gamma}\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i(t)dt$. Note that

$$E([\mathbf{S}_{\mathrm{eff}\boldsymbol{\beta}}\{Y_i, \mathbf{X}_i(\cdot); \boldsymbol{\beta}_0, \alpha_0(\cdot), g^*\}^{\mathrm{T}}, \mathbf{S}_{\mathrm{eff}\boldsymbol{\gamma}}\{Y_i, \mathbf{X}_i(\cdot); \boldsymbol{\beta}_0, \alpha_0(\cdot), g^*\}^{\mathrm{T}}]^{\mathrm{T}}) = \mathbf{0},$$

where $[\mathbf{S}_{\mathrm{eff}\boldsymbol{\beta}}\{Y_i, \mathbf{X}_i(\cdot); \boldsymbol{\beta}_0, \alpha_0(\cdot), g^*\}^{\mathrm{T}}, \mathbf{S}_{\mathrm{eff}\boldsymbol{\gamma}}\{Y_i, \mathbf{X}_i(\cdot); \boldsymbol{\beta}_0, \alpha_0(\cdot), g^*\}^{\mathrm{T}}]^{\mathrm{T}}$ is the vector which replaces the occurrence of $\mathbf{B}_r(\cdot)^{\mathrm{T}}\boldsymbol{\gamma}$ in the vector $[\mathbf{S}_{\mathrm{eff}\boldsymbol{\beta}}(Y_i, \mathbf{Z}_i, \boldsymbol{\beta}_0, \boldsymbol{\gamma}, g^*)^{\mathrm{T}}, \mathbf{S}_{\mathrm{eff}\boldsymbol{\gamma}}(Y_i, \mathbf{Z}_i, \boldsymbol{\beta}_0, \boldsymbol{\gamma}, g^*)^{\mathrm{T}}]^{\mathrm{T}}$ by $\alpha_0(\cdot)$. So the uniform convergence of $\mathbf{B}_r(\cdot)^{\mathrm{T}}\boldsymbol{\gamma}_0$ to $\alpha_0(\cdot)$ leads to

$$E\{\mathbf{S}_{\mathrm{eff}\boldsymbol{\beta}}(Y_i, \mathbf{Z}_i, \boldsymbol{\beta}_0, \boldsymbol{\gamma}_0, g^*)^{\mathrm{T}}, \mathbf{S}_{\mathrm{eff}\boldsymbol{\gamma}}(Y_i, \mathbf{Z}_i, \boldsymbol{\beta}_0, \boldsymbol{\gamma}_0, g^*)^{\mathrm{T}}\}^{\mathrm{T}} = o(1).$$

On the other hand, from the estimating equation

$$n^{-1}\sum_{i=1}^{n}\{\widehat{\mathbf{S}}_{\mathrm{eff}\boldsymbol{\beta}}(Y_i, \mathbf{Z}_i, \widehat{\boldsymbol{\beta}}, \widehat{\boldsymbol{\gamma}}, g^*)^{\mathrm{T}}, \widehat{\mathbf{S}}_{\mathrm{eff}\boldsymbol{\gamma}}(Y_i, \mathbf{Z}_i, \widehat{\boldsymbol{\beta}}, \widehat{\boldsymbol{\gamma}}, g^*)^{\mathrm{T}}\}^{\mathrm{T}} = \mathbf{0},$$

using the consistency of the nonparametric kernel estimation, we have

$$n^{-1}\sum_{i=1}^{n}\{\mathbf{S}_{\mathrm{eff}\boldsymbol{\beta}}(Y_i, \mathbf{Z}_i, \widehat{\boldsymbol{\beta}}, \widehat{\boldsymbol{\gamma}}, g^*)^{\mathrm{T}}, \mathbf{S}_{\mathrm{eff}\boldsymbol{\gamma}}(Y_i, \mathbf{Z}_i, \widehat{\boldsymbol{\beta}}, \widehat{\boldsymbol{\gamma}}, g^*)^{\mathrm{T}}\}^{\mathrm{T}} = o_p(1).$$

The law of large numbers further leads to

$$\begin{aligned} & E\{\mathbf{S}_{\mathrm{eff}\boldsymbol{\beta}}(Y_i, \mathbf{Z}_i, \boldsymbol{\beta}, \boldsymbol{\gamma}, g^*)^{\mathrm{T}}, \\ & \mathbf{S}_{\mathrm{eff}\boldsymbol{\gamma}}(Y_i, \mathbf{Z}_i, \boldsymbol{\beta}, \boldsymbol{\gamma}, g^*)^{\mathrm{T}}\}^{\mathrm{T}}|_{\boldsymbol{\beta}=\widehat{\boldsymbol{\beta}}, \boldsymbol{\gamma}=\widehat{\boldsymbol{\gamma}}} = o_p(1). \end{aligned} \tag{S4}$$

The unique root property in Condition (A7) implies that the derivative of

$$E[\{\mathbf{S}_{\mathrm{eff}\boldsymbol{\beta}}(Y_i, \mathbf{Z}_i, \boldsymbol{\beta}, \boldsymbol{\gamma}, g^*)^{\mathrm{T}}, (Y_i, \mathbf{Z}_i, \boldsymbol{\beta}, \boldsymbol{\gamma}, g^*)^{\mathrm{T}}\}]^{\mathrm{T}}$$

with respect to $(\boldsymbol{\beta}^{\mathrm{T}}, \boldsymbol{\gamma}^{\mathrm{T}})^{\mathrm{T}}$ is a nonsingular matrix in the neighborhood where the function value is $\mathbf{0}$. Therefore, the left hand side of both (S4) and (S4) are invertible functions of $(\boldsymbol{\beta}^{\mathrm{T}}, \boldsymbol{\gamma}^{\mathrm{T}})^{\mathrm{T}}$. This implies $\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0 = o_p(1)$ and $\widehat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0 = o_p(1)$ by the continuous mapping theorem. $\square$

S2.3 Proof of Proposition 3

Denote the nuisance tangent space corresponding to f and the marginal density of $\mathbf{Z}$ as respectively Λ_f and Λ_z. Taking derivative of the loglikelihood function with respect to the nuisance parameters in each parametric submodel and then taking the closure of their union, we have

$$\begin{aligned}
\Lambda_z &= \{\mathbf{f}(\mathbf{Z}) : \forall \mathbf{f} \text{ such that } E(\mathbf{f}) = \mathbf{0}\} \\
\Lambda_f &= \{\mathbf{f}(Y, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma}) : \forall \mathbf{f} \text{ such that } E(\mathbf{f} \mid \mathbf{Z}) = E(\mathbf{f} \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma}) = \mathbf{0}\}.
\end{aligned}$$

We can easily verify that $\Lambda_z \perp \Lambda_f$, hence $\Lambda = \Lambda_z \oplus \Lambda_f$, where Λ is the nuisance tangent space.

Thus, $\Lambda^\perp = \Lambda_z^\perp \cap \Lambda_f^\perp$. It is easy to see that $\Lambda_z^\perp = \{\mathbf{f}(Y, \mathbf{Z}) : E(\mathbf{f} \mid \mathbf{Z}) = 0\}$. We now show that $\Lambda_f^\perp = \{\mathbf{f}(Y, \mathbf{Z}) : E(\mathbf{f} \mid Y, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma}) = E(\mathbf{f} \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma})\}$. We show this in two aspects. First, it is easy to verify that functions having the above conditional expectation property are elements in $\Lambda_f^\perp$. To show the second aspect that elements in $\Lambda_f^\perp$ have to satisfy the conditional expectation requirement, consider any $\mathbf{f}(Y, \mathbf{Z}) \in \Lambda_f^\perp$. We choose $\mathbf{g} = E(\mathbf{f} \mid Y, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma}) - E(\mathbf{f} \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma})$. Obviously, $\mathbf{g} \in \Lambda_f$ hence $E(\mathbf{g}^{\mathrm{T}}\mathbf{f}) = 0$. We write this relation alternatively as

$$\begin{aligned}
0 = E(\mathbf{g}^{\mathrm{T}}\mathbf{f}) &= E\{\mathbf{g}^{\mathrm{T}} E(\mathbf{f} \mid Y, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma})\} \\
&= E(\mathbf{g}^{\mathrm{T}}\mathbf{g}) + E\{\mathbf{g}^{\mathrm{T}} E(\mathbf{f} \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma})\} \\
&= E(\mathbf{g}^{\mathrm{T}}\mathbf{g}) + E\{E(\mathbf{g} \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma})^{\mathrm{T}} E(\mathbf{f} \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma})\} \\
&= E(\mathbf{g}^{\mathrm{T}}\mathbf{g}) + 0.
\end{aligned}$$

This implies $\mathbf{g}$ itself should be zero. This means $\mathbf{f}$ indeed satisfies the conditional expectation requirement.

We are now ready to prove the form of $\Lambda^{\perp}$. For convenience, we denote set described in the proposition A first and we will establish $\Lambda^{\perp} = A$ through proving $\Lambda^{\perp} \subset A$ and $A \subset \Lambda^{\perp}$. To see $A \subset \Lambda^{\perp}$, we can verify that for any $\mathbf{f}(Y, \mathbf{Z}) - E(\mathbf{f} \mid Y, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma}) \in A$, we obviously have $E\{\mathbf{f}(Y, \mathbf{Z}) - E(\mathbf{f} \mid Y, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma}) \mid \mathbf{Z}\} = E(\mathbf{f} \mid \mathbf{Z}) - E(\mathbf{f} \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma}) = \mathbf{0}$, thus $A \subset \Lambda_z^{\perp}$. The form of elements in $\mathbf{A}$ also ensures that $E\{\mathbf{f}(Y, \mathbf{Z}) - E(\mathbf{f} \mid Y, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma}) \mid Y, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma}\} = \mathbf{0} = E\{\mathbf{f}(Y, \mathbf{Z}) - E(\mathbf{f} \mid Y, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma}) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma}\}$. Thus $A \subset \Lambda_f^{\perp}$. This shows $A \subset \Lambda^{\perp}$. We now show $\Lambda^{\perp} \subset A$. Assume $\mathbf{g}(Y, \mathbf{Z}) \in \Lambda^{\perp}$. Then $E(\mathbf{g} \mid \mathbf{Z}) = \mathbf{0}$ and $E(\mathbf{g} \mid Y, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma}) = E(\mathbf{g} \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma}) = E\{E(\mathbf{g} \mid \mathbf{Z}) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma}\} = \mathbf{0}$. This means we can always write $\mathbf{g}(Y, \mathbf{Z})$ as $\mathbf{f}(Y, \mathbf{Z}) - E(\mathbf{f} \mid Y, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma})$. The additional requirement of $E(\mathbf{g} \mid \mathbf{Z}) = \mathbf{0}$ further imposes $E(\mathbf{f} \mid \mathbf{Z}) = E\{E(\mathbf{f} \mid Y, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma}) \mid \mathbf{Z}\} = E(\mathbf{f} \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma})$. Thus, we indeed have $\Lambda^{\perp} \subset A$. $\square$

S3 Proofs of the theorems in the main text

S3.1 Proof of Theorem 1

By the consistency shown in Proposition 2, we expand the score function as

$$\begin{aligned}
\mathbf{0} &= n^{-1/2}\sum_{i=1}^{n}\Bigg\{g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)\} \qquad \text{(S5)}\\
&\quad - \frac{\sum_{j=1}^{J} K_h\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0) - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)\}g^*\{Y_j, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)\}}{\sum_{j=1}^{J} K_h\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0) - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)\}}\Bigg\}\\
&\quad \boldsymbol{\Theta}_{\gamma}\left\{\mathbf{Z}_i - \frac{\sum_{j=1}^{J} K_h\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0) - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)\}\mathbf{Z}_j}{\sum_{j=1}^{J} K_h\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0) - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)\}}\right\}^{\mathrm{T}}\boldsymbol{\beta}_0\\
&= \mathbf{R} + \mathbf{T}n^{1/2}\{\widehat{\boldsymbol{\gamma}}^{-}(\boldsymbol{\beta}_0) - \boldsymbol{\gamma}_0^{-}\}, \qquad \text{(S6)}
\end{aligned}$$

where

$$\begin{aligned}
\mathbf{R} &= n^{-1/2}\sum_{i=1}^{n}\left\{g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0) - \frac{\sum_{j=1}^{J} K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0 - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)g^*(Y_j, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0)}{\sum_{j=1}^{J} K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0 - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)}\right\}\\
&\quad \times\boldsymbol{\Theta}_{\gamma}\left\{\mathbf{Z}_i - \frac{\sum_{j=1}^{J} K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0 - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)\mathbf{Z}_j}{\sum_{j=1}^{J} K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0 - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)}\right\}^{\mathrm{T}}\boldsymbol{\beta}_0,
\end{aligned}$$

$$
\begin{aligned}
\mathbf{T} &= n^{-1}\sum_{i=1}^{n}\partial\Bigg[\left\{g^*(Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma})-\frac{\sum_{j=1}^{J}K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}-\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma})g^*(Y_j,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma})}{\sum_{j=1}^{J}K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}-\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma})}\right\} \\
&\quad\times\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\left\{\mathbf{Z}_i-\frac{\sum_{j=1}^{J}K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}-\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma})\mathbf{Z}_j}{\sum_{j=1}^{J}K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}-\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma})}\right\}^{\mathrm{T}}\boldsymbol{\beta}_0\Bigg]\Big/\partial(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma})\Big|_{\boldsymbol{\gamma}=\boldsymbol{\gamma}^*}\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}},
\end{aligned}
$$

and $\boldsymbol{\gamma}^*$ is the point on the line connecting $\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)$ and $\boldsymbol{\gamma}_0$.

The Asymptotic Property of R Now $\mathbf{R}$ can be written as $\mathbf{R}=\mathbf{R}_0+\mathbf{R}_1+\mathbf{R}_2+\mathbf{R}_3$, where

$$
\begin{aligned}
\mathbf{R}_0 &= n^{-1/2}\sum_{i=1}^{n}\left[g^*(Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)-E\{g^*(Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}\right]\boldsymbol{\Theta}_{\boldsymbol{\gamma}} \\
&\quad\times\left\{\mathbf{Z}_i-E(\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)\right\}^{\mathrm{T}}\boldsymbol{\beta}_0, \\
\mathbf{R}_1 &= n^{-1/2}\sum_{i=1}^{n}\left[g^*(Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)-E\{g^*(Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}\right]\boldsymbol{\Theta}_{\boldsymbol{\gamma}} \\
&\quad\times\left\{E(\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)-\frac{\sum_{j=1}^{J}K_h\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0-\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}\mathbf{Z}_j}{\sum_{j=1}^{J}K_h\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0-\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}}\right\}^{\mathrm{T}}\boldsymbol{\beta}_0, \\
\mathbf{R}_2 &= n^{-1/2}\sum_{i=1}^{n}\Bigg[E\{g^*(Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\} \\
&\quad-\frac{\sum_{j=1}^{J}K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0-\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)g^*(Y_j,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0)}{\sum_{j=1}^{J}K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0-\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)}\Bigg]\boldsymbol{\Theta}_{\boldsymbol{\gamma}} \\
&\quad\times\left\{\mathbf{Z}_i-E(\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)\right\}^{\mathrm{T}}\boldsymbol{\beta}_0, \\
\mathbf{R}_3 &= n^{-1/2}\sum_{i=1}^{n}\Bigg[E\{g^*(Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\} \\
&\quad-\frac{\sum_{j=1}^{J}K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0-\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)g^*(Y_j,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0)}{\sum_{j=1}^{J}K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0-\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)}\Bigg]\boldsymbol{\Theta}_{\boldsymbol{\gamma}} \\
&\quad\times\left\{E(\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)-\frac{\sum_{j=1}^{J}K_h\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0-\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}\mathbf{Z}_j}{\sum_{j=1}^{J}K_h\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0-\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}}\right\}^{\mathrm{T}}\boldsymbol{\beta}_0.
\end{aligned}
$$

Further, we write $\mathbf{R}_1=\mathbf{R}_{11}+\mathbf{R}_{12}$, where

$$
\begin{aligned}
\mathbf{R}_{11} &= n^{-1/2}\sum_{i=1}^{n}\left[g^*(Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)-E\{g^*(Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}\right]\boldsymbol{\Theta}_{\boldsymbol{\gamma}} \\
&\quad\times\Bigg\{\frac{E(\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)\sum_{j=1}^{J}K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0-\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)}{nf_{\mathbf{Z}}(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)}
\end{aligned}
$$

$$
\left. - \frac{\sum_{j=1}^{J} K_h\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0 - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}\mathbf{Z}_j}{nf_{\mathbf{Z}}(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)} \right\}^{\mathrm{T}} \boldsymbol{\beta}_0,
$$

$$
\begin{aligned}
\mathbf{R}_{12} &= n^{-1/2}\sum_{i=1}^{n}\left[g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0) - E\{g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}\right]\boldsymbol{\Theta}_{\boldsymbol{\gamma}} \\
&\quad \times\left\{ \frac{E(\mathbf{Z}_i \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)\sum_{j=1}^{J} K_h\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0 - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}}{nf_{\mathbf{Z}}(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)} \right. \\
&\quad \left. - \frac{\sum_{j=1}^{J} K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0 - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)\mathbf{Z}_j}{nf_{\mathbf{Z}}(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)} \right\}^{\mathrm{T}} \boldsymbol{\beta}_0 \\
&\quad \times\left\{ \frac{nf_{\mathbf{Z}}(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)}{\sum_{j=1}^{J} K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0 - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)} - 1 \right\} \\
&= \mathbf{R}_{11}\left\{ \frac{nf_{\mathbf{Z}}(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)}{\sum_{j=1}^{J} K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0 - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)} - 1 \right\},
\end{aligned}
$$

where $f_{\mathbf{Z}}(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)$ is the density for $\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma}_0$.

We first analyze $\mathbf{R}_{11}$. Using the U statistics property, we write $\mathbf{R}_{11}$ as $\mathbf{R}_{11} = \mathbf{R}_{111} + \mathbf{R}_{112} - \mathbf{R}_{113} + o_p(\mathbf{R}_{111} + \mathbf{R}_{112} - \mathbf{R}_{113})$, where

$$
\begin{aligned}
&\mathbf{R}_{11} \\
&= n^{-3/2}\sum_{i=1}^{n}\sum_{j=1}^{J}\left[g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0) - E\{g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}\right]\boldsymbol{\Theta}_{\boldsymbol{\gamma}} \\
&\quad \times\left\{ \frac{E(\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0 - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)}{nf_{\mathbf{Z}}\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)} \right. \\
&\quad \left. - \frac{K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0 - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)\mathbf{Z}_j}{nf_{\mathbf{Z}}\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)} \right\}^{\mathrm{T}} \boldsymbol{\beta}_0 \\
&= \mathbf{R}_{111} + \mathbf{R}_{112} - \mathbf{R}_{113} + o_p(\mathbf{R}_{111} + \mathbf{R}_{112} - \mathbf{R}_{113}),
\end{aligned}
$$

where

$$
\begin{aligned}
\mathbf{R}_{111} &= \left\{ n^{-1/2}\sum_{i=1}^{n} E\Bigg(\left[g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0) - E\{g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}\right]\boldsymbol{\Theta}_{\boldsymbol{\gamma}} \right. \\
&\quad \times\left\{ \frac{E(\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0 - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)}{nf_{\mathbf{Z}}\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)} \right. \\
&\quad \left. - \frac{K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0 - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)\mathbf{Z}_j}{nf_{\mathbf{Z}}\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)} \right\}^{\mathrm{T}} \boldsymbol{\beta}_0|\mathbf{O}_i\Bigg), \\
\mathbf{R}_{112} &= n^{-1/2}\sum_{j=1}^{J} E\Bigg(\left[g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0) - E\{g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}\right]\boldsymbol{\Theta}_{\boldsymbol{\gamma}}
\end{aligned}
$$

$$
\begin{aligned}
&\times\left\{\frac{E(\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0-\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)}{nf_{\mathbf{Z}}\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)}\right.\\
&\left.-\frac{K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0-\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)\mathbf{Z}_j}{nf_{\mathbf{Z}}\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)}\right\}^{\mathrm{T}}\boldsymbol{\beta}_0|\mathbf{O}_j\Bigg),\\
\mathbf{R}_{113} \;=\;& n^{1/2}E\Bigg(\left[g^*(Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)-E\{g^*(Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}\right]\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\\
&\times\left\{\frac{E(\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0-\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)}{nf_{\mathbf{Z}}\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)}\right.\\
&\left.-\frac{K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0-\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)\mathbf{Z}_j}{nf_{\mathbf{Z}}\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)}\right\}^{\mathrm{T}}\boldsymbol{\beta}_0\Bigg)\Bigg\},
\end{aligned}
$$

where $\mathbf{O}_i$ is the random variable corresponding to the ith observation. Now because the summand for $\mathbf{R}_{111}$, i.e.,

$$
\begin{aligned}
&E\Bigg(\left[g^*(Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)-E\{g^*(Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}\right]\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\\
&\times\left\{\frac{E(\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0-\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)}{nf_{\mathbf{Z}}\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)}\right.\\
&\left.-\frac{K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0-\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)\mathbf{Z}_j}{nf_{\mathbf{Z}}\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)}\right\}^{\mathrm{T}}\boldsymbol{\beta}_0|\mathbf{O}_i\Bigg)\\
=\;&\left[g^*(Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)-E\{g^*(Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}\right]\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\\
&\times\left[\frac{E(\mathbf{Z}_i\mid\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)E\{K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0-\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)|\mathbf{O}_i\}}{f_{\mathbf{Z}}(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)}\right.\\
&\left.-\frac{E\{K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0)-\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)\mathbf{Z}_j|\mathbf{O}_i\}}{f_{\mathbf{Z}}(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)}\right]^{\mathrm{T}}\boldsymbol{\beta}_0\\
=\;&\left[g^*(Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)-E\{g^*(Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}\right]\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\\
&\times\left[\frac{E(\mathbf{Z}_i\mid\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)E\{K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0-\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)|\mathbf{O}_i\}}{f_{\mathbf{Z}}(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)}\right.\\
&\left.-\frac{E[K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0-\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)E(\mathbf{Z}_j|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0)|\mathbf{O}_i]}{f_{\mathbf{Z}}(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)}\right]^{\mathrm{T}}\boldsymbol{\beta}_0\\
=\;&\left[g^*(Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)-E\{g^*(Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}\right]\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\\
&\times\Bigg[E(\mathbf{Z}_i\mid\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)^{\mathrm{T}}\boldsymbol{\beta}_0\Bigg\{1+\int\frac{K(s)s^2}{2f_{\mathbf{Z}}(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)}\\
&\times\frac{\partial^2 f_{\mathbf{Z}}\{a(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0,sh)\}}{\partial a(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0,sh)^2}dsh^2\Bigg\}-E(\mathbf{Z}_i\mid\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)^{\mathrm{T}}\boldsymbol{\beta}_0-\int\frac{s^2K(s)}{2f_{\mathbf{Z}}(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)}\\
&\times\frac{\partial^2[E\{\mathbf{Z}_i\mid\boldsymbol{\beta}_0^{\mathrm{T}}b(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0,sh)\}^{\mathrm{T}}\boldsymbol{\beta}_0 f_{\mathbf{Z}}\{b(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0,sh)\}]}{\partial b(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0,sh)^2}dsh^2\Bigg]
\end{aligned}
$$

$$
\begin{aligned}
= \;& \left[g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0) - E\{g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}\right] \boldsymbol{\Theta}_{\boldsymbol{\gamma}} \\
& \times\Bigg[E(\mathbf{Z}_i \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)^{\mathrm{T}}\boldsymbol{\beta}_0 \int \frac{\partial^2 f_{\mathbf{Z}}\{a(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0, sh)\}}{\partial a(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0, sh)^2}\frac{s^2K(s)}{2f_{\mathbf{Z}}(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)}ds \\
& - \int \frac{\partial^2[E\{\mathbf{Z}_i \mid b(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0, sh)\}^{\mathrm{T}}\boldsymbol{\beta}_0 f_{\mathbf{Z}}\{b(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0, sh)\}]}{\partial b(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0, sh)^2}\frac{s^2K(s)}{2f_{\mathbf{Z}}(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)}ds\Bigg]h^2, \qquad \text{(S7)}
\end{aligned}
$$

where $a(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0, sh), b(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0, sh)$ are points on the line connecting $\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0$ and $\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0+sh$. Note that $a(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0, sh), b(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0, sh)$ depend on the values of both $\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0$ and sh and will go to $\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0$ when $h \to 0$. Now note that

$$
\begin{aligned}
& \left[g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0) - E\{g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}\right] \boldsymbol{\Theta}_{\boldsymbol{\gamma}} \\
& \times\Bigg[E(\mathbf{Z}_i \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)^{\mathrm{T}}\boldsymbol{\beta}_0 \int \frac{\partial^2 f_{\mathbf{Z}}\{a(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0, sh)\}}{\partial a(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0, sh)^2}\frac{s^2K(s)}{2f_{\mathbf{Z}}(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)}ds \\
& - \int \frac{\partial^2[E\{\mathbf{Z}_i \mid b(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0, sh)\}^{\mathrm{T}}\boldsymbol{\beta}_0 f_{\mathbf{Z}}\{b(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0, sh)\}]}{\partial b(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0, sh)^2}\frac{s^2K(s)}{2f_{\mathbf{Z}}(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)}ds\Bigg]
\end{aligned}
$$

can be written as

$$
\int_0^1 C_{1i}(t)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\mathbf{B}_r(t)dt = \int_0^1 C_{1i}(t)\{B_{rk}(t), k = 2, \ldots, d_{\boldsymbol{\gamma}}\}^{\mathrm{T}}dt,
$$

with

$$
\begin{aligned}
& C_{1i}(t) \\
= \;& \left[g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0) - E\{g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}\right] \\
& \times\Bigg[E(\mathbf{X}_i(t)^{\mathrm{T}}\boldsymbol{\beta}_0 \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0) \int \frac{\partial^2 f_{\mathbf{Z}}\{a(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0, sh)\}}{\partial a(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0, sh)^2}\frac{s^2K(s)}{2f_{\mathbf{Z}}(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)}ds \\
& - \int \frac{\partial^2[E\{\mathbf{X}_i(t)^{\mathrm{T}}\boldsymbol{\beta}_0 \mid b(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0, sh)\} f_{\mathbf{Z}}\{b(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0, sh)\}]}{\partial b(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0, sh)^2}\frac{s^2K(s)}{2f_{\mathbf{Z}}(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)}ds\Bigg].
\end{aligned}
$$

Obviously $E\{C_{1i}(t)\} = 0$ and hence $|E\{C_{1i}(t)\}| < \infty$. From Conditions (A6) and (A8), $\|C_{1i}(\cdot)\|_2 < \infty,\ a.s.$, so by Lemma 3, we have

$$
\|\mathbf{R}_{111}\|_\infty = O_p\{h^2\sqrt{h_b \log(n)}\}, \qquad \text{(S8)}
$$

which also implies

$$
\|\mathbf{R}_{113}\|_\infty = \|E(\mathbf{R}_{111})\|_\infty = O\{h^2\sqrt{h_b \log(n)}\}. \qquad \text{(S9)}
$$

Further, the summand of $\mathbf{R}_{112}$, i.e.,

$$
\begin{aligned}
& E\Bigg(\left[g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0) - E\{g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}\right] \boldsymbol{\Theta}_{\boldsymbol{\gamma}} \\
& \times\Bigg\{ \frac{E(\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0 - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)}{nf_{\mathbf{Z}}\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)} \\
& -\frac{K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0 - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)\mathbf{Z}_j}{nf_{\mathbf{Z}}\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)}\Bigg\}^{\mathrm{T}} \boldsymbol{\beta}_0|\mathbf{O}_j\Bigg) \\
= \;& E\Bigg(E\left(\left[g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0) - E\{g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}\right]|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0, \mathbf{O}_j\right) \boldsymbol{\Theta}_{\boldsymbol{\gamma}} \\
& \times\Bigg\{ \frac{E(\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0 - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)}{nf_{\mathbf{Z}}\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)} \\
& -\frac{K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0 - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)\mathbf{Z}_j}{nf_{\mathbf{Z}}\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)}\Bigg\}^{\mathrm{T}} \boldsymbol{\beta}_0\Bigg|\mathbf{O}_j\Bigg) \\
= \;& \mathbf{0}.
\end{aligned} \tag{S10}
$$

Combining (S8), (S10), (S9), we get

$$
\mathbf{R}_{11} = O_p\{h^2\sqrt{h_b \log(n)}\}. \tag{S11}
$$

Further, to treat $\mathbf{R}_{12}$, note that

$$
\begin{aligned}
& \left\{\frac{nf_{\mathbf{Z}}(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)}{\sum_{j=1}^J K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0 - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)} - 1\right\} \\
= \;& \left\{\frac{f_{\mathbf{Z}}(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0) - n^{-1}\sum_{j=1}^J K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0 - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)}{n^{-1}\sum_{j=1}^J K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0 - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)}\right\} \\
= \;& O_p\{h^2 + (nh)^{-1/2}\}
\end{aligned} \tag{S12}
$$

hence

$$
\|\mathbf{R}_{12}\|_\infty = o_p\{h^2\sqrt{h_b \log(n)}\}. \tag{S13}
$$

Thus, combining (S11) and (S13), we have

$$
\|\mathbf{R}_1\|_\infty = O_p\{h^2\sqrt{h_b \log(n)}\}. \tag{S14}
$$

Using similar derivation as those led to (S14) while exchange the roles of $\mathbf{Z}$ and g^*, we can obtain

$$\|\mathbf{R}_2\|_\infty = O_p\{h^2\sqrt{h_b \log(n)}\}. \tag{S15}$$

For $\mathbf{R}_3$, we first note that

$$\begin{aligned}&\left|\left[E\{g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\} - \frac{\sum_{j=1}^J K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0 - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)g^*(Y_j, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0)}{\sum_{j=1}^J K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0 - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)}\right]\right|\\ &= O_p(h^2 + n^{-1/2}h^{-1/2})\end{aligned}$$

by the uniform consistency of the kernel estimator. Further, the summand in $\mathbf{R}_3$ has the form

$$\begin{aligned}&\left[E\{g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\} - \frac{\sum_{j=1}^J K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0 - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)g^*(Y_j, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0)}{\sum_{j=1}^J K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0 - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)}\right]\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\\ &\times\left\{E(\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0) - \frac{\sum_{j=1}^J K_h\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0 - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}\mathbf{Z}_j}{\sum_{j=1}^J K_h\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0 - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}}\right\}^{\mathrm{T}}\boldsymbol{\beta}_0\\ =&\int_0^1 (h^4 + n^{-1}h^{-1})C_{2i}(t)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\mathbf{B}_r(t)dt\\ =&\int_0^1 (h^4 + n^{-1}h^{-1})C_{2i}(t)\{B_{rk}(t), k = 2, \ldots, d_{\boldsymbol{\gamma}}\}^{\mathrm{T}}dt,\end{aligned}$$

where

$$\begin{aligned}&C_{2i}(t)\\ =&(h^4 + n^{-1}h^{-1})^{-1}\left[E\{g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}\right.\\ &\left.-\frac{\sum_{j=1}^J K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0 - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)g^*(Y_j, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0)}{\sum_{j=1}^J K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0 - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)}\right]\\ &\times\left\{E\{\mathbf{X}_i(t)^{\mathrm{T}}\boldsymbol{\beta}_0|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\} - \frac{\sum_{j=1}^J K_h\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0 - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}\mathbf{X}_j(t)^{\mathrm{T}}\boldsymbol{\beta}_0}{\sum_{j=1}^J K_h\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0 - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}}\right\}.\end{aligned}$$

Note that $E\{C_{2i}(t)\} \neq 0$, $|E\{C_{2i}(t)\}| < \infty$, $\|C_{2i}(\cdot)\|_2 < \infty$, $a.s.$ by Conditions (A2), (A6) and the uniform consistency of the kernel estimator. So by Lemma 3, we have

$$\mathbf{R}_3 = n^{-1/2}\sum_{i=1}^n \int_0^1 (h^4 + n^{-1}h^{-1})C_{2i}\{B_{rk}(t), k = 2, \ldots, d_{\boldsymbol{\gamma}}\}^{\mathrm{T}}dt$$

$$= \quad O_p\{(n^{1/2}h^4 + n^{-1/2}h^{-1})h_b\}. \tag{S16}$$

We now assess (S7). We write

$$\begin{aligned}
&\mathbf{R}_0 \\
= \quad & n^{-1/2}\sum_{i=1}^n [g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\} - E\{g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}]\boldsymbol{\Theta}_{\boldsymbol{\gamma}} \\
& \times \left\{\mathbf{Z}_i - E(\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)\right\}^{\mathrm{T}}\boldsymbol{\beta}_0 \\
= \quad & n^{-1/2}\sum_{i=1}^n \int_0^1 C_{3i}(t)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\mathbf{B}_r(t)dt \\
= \quad & n^{-1/2}\sum_{i=1}^n \int_0^1 C_{3i}(t)\{B_{rk}(t), k = 2, \ldots, d_{\boldsymbol{\gamma}}\}^{\mathrm{T}}dt,
\end{aligned}$$

where

$$\begin{aligned}
C_{3i}(t) \quad = \quad & [g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\} - E\{g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}][\mathbf{X}_i(t)^{\mathrm{T}}\boldsymbol{\beta}_0 \\
& -E\{\mathbf{X}_i(t)^{\mathrm{T}}\boldsymbol{\beta}_0|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}].
\end{aligned}$$

We can easily verify that $E\{C_{3i}(t)\} \neq 0$, $E\{C_{3i}(t)\} < \infty$ and $\|C_{3i}(\cdot)\|_2 < \infty$, *a.s.* by Condition (A6). Therefore, by Lemma 3 we know

$$\|\mathbf{R}_0\|_\infty = O_p(n^{1/2}h_b). \tag{S17}$$

Comparing (S14), (S15), (S16) with (S17), from Condition (A2), it is clear that $\mathbf{R}_0$ dominates $\mathbf{R}_1, \mathbf{R}_2$ and $\mathbf{R}_3$. We further analyze $\mathbf{R}_0$ by writing

$$\mathbf{R}_0 \quad = \quad \mathbf{R}_{00} + \mathbf{R}_{01} + \mathbf{R}_{02} \tag{S18}$$

where

$$\begin{aligned}
\mathbf{R}_{00} \quad = \quad & n^{-1/2}\sum_{i=1}^n [g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\} - E\{g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}]\boldsymbol{\Theta}_{\boldsymbol{\gamma}} \\
& \times \left[\mathbf{Z}_i - E\{\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\right]^{\mathrm{T}}\boldsymbol{\beta}_0
\end{aligned} \tag{S19}$$

and

$$\mathbf{R}_{01} \quad = \quad n^{-1/2}\sum_{i=1}^n \Big([g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\} - E\{g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}]$$

$$-[g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\} - E\{g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}]\Big)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}$$
$$\times \left[\mathbf{Z}_i - E\{\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}\right]^{\mathrm{T}}\boldsymbol{\beta}_0,$$

$$\begin{aligned}\mathbf{R}_{02} &= n^{-1/2}\sum_{i=1}^{n}[g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\} - E\{g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}]\boldsymbol{\Theta}_{\boldsymbol{\gamma}} \\ &\quad \times\left(\left[\mathbf{Z}_i - E\{\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}\right] - \left[\mathbf{Z}_i - E\{\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\right]\right)^{\mathrm{T}}\boldsymbol{\beta}_0.\end{aligned}$$

To analyze $\mathbf{R}_{00}$, we write it as

$$\begin{aligned}\mathbf{R}_{00} &= n^{-1/2}\sum_{i=1}^{n}\int_0^1 C_{4i}(t)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\mathbf{B}_r(t)dt \\ &= n^{-1/2}\sum_{i=1}^{n}\int_0^1 C_{4i}(t)\{B_{rk}(t), k=2,\ldots,d_{\boldsymbol{\gamma}}\}^{\mathrm{T}}dt,\end{aligned}$$

where

$$\begin{aligned}&C_{4i}(t) \\ =\ &[g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\} - E\{g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}][\mathbf{X}_i(t)^{\mathrm{T}}\boldsymbol{\beta}_0 \\ &-E\{\mathbf{X}_i(t)^{\mathrm{T}}\boldsymbol{\beta}_0|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}].\end{aligned}$$

We can check that $E\{C_{4i}(t)\} = 0$ and $\|C_{4i}(\cdot)\|_2 < \infty$ *a.s.* by Condition (A6). Thus, by Lemma 3,

$$\|\mathbf{R}_{00}\|_\infty = \sqrt{h_b \log(n)}. \tag{S20}$$

Further

$$\begin{aligned}\mathbf{R}_{02} &= n^{-1/2}\sum_{i=1}^{n}\int_0^1 h_b^q C_{5i}(t)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\mathbf{B}_r(t)dt \\ &= n^{-1/2}\sum_{i=1}^{n}\int_0^1 h_b^q C_{5i}(t)\{B_{rk}(t), k=2,\ldots,d_{\boldsymbol{\gamma}}\}^{\mathrm{T}}dt,\end{aligned}$$

where

$$C_{5i}(t) = h_b^{-q}[g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\} - E\{g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}]$$

$$\times\Bigg([\mathbf{X}_i(t)^{\mathrm{T}}\boldsymbol{\beta}_0 - E\{\mathbf{X}_i(t)^{\mathrm{T}}\boldsymbol{\beta}_0|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}] - [\mathbf{X}_i(t)^{\mathrm{T}}\boldsymbol{\beta}_0$$
$$-E\{\mathbf{X}_i(t)^{\mathrm{T}}\boldsymbol{\beta}_0|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}]\Bigg).$$

Now $E\{C_{5i}(t)\} = 0$ because

$$\begin{aligned}
& E\Bigg\{[g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\} - E\{g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}] \\
& \times\Bigg(\Big[\mathbf{X}_i(t)^{\mathrm{T}}\boldsymbol{\beta}_0 - E\{\mathbf{X}_i(t)^{\mathrm{T}}\boldsymbol{\beta}_0|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}\Big] - [\mathbf{X}_i(t)^{\mathrm{T}}\boldsymbol{\beta}_0 \\
& -E\{\mathbf{X}_i(t)^{\mathrm{T}}\boldsymbol{\beta}_0|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}]\Bigg)\Bigg\} \\
= \; & E\Bigg[E\Bigg\{[g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\} - E\{g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}]|\mathbf{X}_i\Bigg\} \\
& \times\Bigg(\Big[\mathbf{X}_i(t)^{\mathrm{T}}\boldsymbol{\beta}_0 - E\{\mathbf{X}_i(t)^{\mathrm{T}}\boldsymbol{\beta}_0|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}\Big] - [\mathbf{X}_i(t)^{\mathrm{T}}\boldsymbol{\beta}_0 \\
& -E\{\mathbf{X}_i(t)^{\mathrm{T}}\boldsymbol{\beta}_0|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}]\Bigg)\Bigg] \\
= \; & 0.
\end{aligned}$$

Further $\|C_{5i}(\cdot)\|_2 < \infty$, *a.s.* by Condition (A6). Therefore, by Lemma 3 we have

$$\begin{aligned}
& \|\mathbf{R}_{02}\|_\infty \\
= \; & \|n^{-1/2}\sum_{i=1}^n [g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\} - E\{g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}]\boldsymbol{\Theta}_{\boldsymbol{\gamma}} \\
& \times\Bigg(\Big[\mathbf{Z}_i - E\{\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}\Big] - \Big[\mathbf{Z}_i - E\{\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\Big]\Bigg)^{\mathrm{T}}\boldsymbol{\beta}_0\|_\infty \\
= \; & O_p\{h_b^q\sqrt{h_b \log(n)}\}.
\end{aligned}$$

In addition,

$$\begin{aligned}
\mathbf{R}_{01} &= n^{-1/2}\sum_{i=1}^n\int_0^1 h_b^q C_{6i}(t)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\mathbf{B}_r(t)dt \\
&= n^{-1/2}\sum_{i=1}^n\int_0^1 h_b^q C_{6i}(t)\{B_{rk}(t), k = 2, \ldots, d_{\boldsymbol{\gamma}}\}^{\mathrm{T}}dt,
\end{aligned}$$

where

$$C_{6i}(t) = h_b^{-q}\Bigg([g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\} - E\{g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}]$$

$$
\begin{aligned}
&-[g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\} - E\{g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}]\Big) \\
&\times \left[\mathbf{X}_i(t)^{\mathrm{T}}\boldsymbol{\beta}_0 - E\{\mathbf{X}_i(t)^{\mathrm{T}}\boldsymbol{\beta}_0|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}\right],
\end{aligned}
$$

with $E\{C_{6i}(t)\} \neq 0$, $|E\{C_{6i}(t)\}| < \infty$, $\|C_{6i}(\cdot)\|_2 < \infty,\ a.s.$ by Condition (A6). Therefore, by Lemma 3 we have

$$
\begin{aligned}
\|\mathbf{R}_{01}\|_\infty &= \|n^{-1/2}\sum_{i=1}^n \Big([g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\} - E\{g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}] \\
&\quad -[g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\} - E\{g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}]\Big)\boldsymbol{\Theta}_{\boldsymbol{\gamma}} \\
&\quad \times \left[\mathbf{Z}_i - E\{\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}\right]^{\mathrm{T}}\boldsymbol{\beta}_0\|_\infty \\
&= O_p(h_b^{q+1}n^{1/2}).
\end{aligned}
$$

Now by the fact that $\mathbf{R}_{01} + \mathbf{R}_{02} = O_p\{h_b^{q+1}n^{1/2} + h_b^q\sqrt{h_b\log(n)}\}$ and $\|\mathbf{R}_{00}\| = O_p\{\sqrt{h_b\log(n)}\}$ by (S20), we have $\mathbf{R}_{01} + \mathbf{R}_{02} = o_p(\mathbf{R}_{00})$ by Condition (A4). Therefore, by (S18) we have $\mathbf{R}_0 = \mathbf{R}_{00} + o_p(\mathbf{R}_{00})$, combining with the fact that $\mathbf{R}_0$ dominates $\mathbf{R}_1, \mathbf{R}_2, \mathbf{R}_3$ and $\mathbf{R} = \mathbf{R}_0 + \mathbf{R}_1 + \mathbf{R}_2 + \mathbf{R}_3$ in (S5), we have

$$
\mathbf{R} = \mathbf{R}_{00} + o_p(\mathbf{R}_{00}), \tag{S21}
$$

i.e.,

$$
\begin{aligned}
&n^{-1/2}\sum_{i=1}^n \left\{g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\} - \frac{\sum_{j=1}^J K_h\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0 - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}g^*\{Y_j, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0\}}{\sum_{j=1}^J K_h\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0 - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}}\right\} \\
&\times\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\left\{\mathbf{Z}_i - \frac{\sum_{j=1}^J K_h\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0 - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}\mathbf{Z}_j}{\sum_{j=1}^J K_h\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma}_0 - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}}\right\}^{\mathrm{T}}\boldsymbol{\beta}_0 = \mathbf{R}_{00} + o_p(\mathbf{R}_{00}),
\end{aligned}
$$

where $\mathbf{R}_{00}$ is given in (S19).

The Asymptotic Property of T Now consider the term

$$
\begin{aligned}
\mathbf{T} &= \Bigg(n^{-1}\sum_{i=1}^n \partial\Bigg[\Bigg\{g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}) \\
&\quad - \frac{\sum_{j=1}^J K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma} - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma})g^*(Y_j, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma})}{\sum_{j=1}^J K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma} - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma})}\Bigg\}
\end{aligned}
$$

$$
\times \boldsymbol{\Theta}_{\boldsymbol{\gamma}} \left\{ \mathbf{Z}_i - \frac{\sum_{j=1}^J K_h(\boldsymbol{\beta}_0^{\mathrm{T}} \mathbf{Z}_j \boldsymbol{\gamma} - \boldsymbol{\beta}_0^{\mathrm{T}} \mathbf{Z}_i \boldsymbol{\gamma}) \mathbf{Z}_j}{\sum_{j=1}^J K_h(\boldsymbol{\beta}_0^{\mathrm{T}} \mathbf{Z}_j \boldsymbol{\gamma} - \boldsymbol{\beta}_0^{\mathrm{T}} \mathbf{Z}_i \boldsymbol{\gamma})} \right\}^{\mathrm{T}} \boldsymbol{\beta}_0 \Bigg] \\
/\partial(\boldsymbol{\beta}_0^{\mathrm{T}} \mathbf{Z}_i \boldsymbol{\gamma}) \Big|_{\boldsymbol{\gamma}=\boldsymbol{\gamma}^*} \boldsymbol{\beta}_0^{\mathrm{T}} \mathbf{Z}_i \boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}} \Bigg) \qquad \text{(S22)}
$$

in (S5). We decompose $\mathbf{T}$ as

$$
\mathbf{T} = \mathbf{T}_0 + \mathbf{T}_1 + \mathbf{T}_2,
$$

where

$$
\begin{aligned}
\mathbf{T}_0 &= \Bigg\{ n^{-1} \sum_{i=1}^n \partial \Bigg(\left[g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{\alpha}_0(\mathbf{X}_i)) - E\{g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{\alpha}_0(\mathbf{X}_i)) | \boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{\alpha}_0(\mathbf{X}_i)\} \right] \boldsymbol{\Theta}_{\boldsymbol{\gamma}} \\
&\quad \times \left[\mathbf{Z}_i - E\{\mathbf{Z}_i | \boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{\alpha}_0(\mathbf{X}_i)\} \right]^{\mathrm{T}} \boldsymbol{\beta}_0 \Bigg) / \partial\{\boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{\alpha}_0(\mathbf{X}_i)\} \boldsymbol{\beta}_0^{\mathrm{T}} \mathbf{Z}_i \boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}} \Bigg\}, \\
\mathbf{T}_1 &= \Bigg\{ n^{-1} \sum_{i=1}^n \partial \Bigg(\left[g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}} \mathbf{Z}_i \boldsymbol{\gamma}) - E\{g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}} \mathbf{Z}_i \boldsymbol{\gamma}) | \boldsymbol{\beta}_0^{\mathrm{T}} \mathbf{Z}_i \boldsymbol{\gamma}\} \right] \boldsymbol{\Theta}_{\boldsymbol{\gamma}} \\
&\quad \times \left[\mathbf{Z}_i - E\{\mathbf{Z}_i | \boldsymbol{\beta}_0^{\mathrm{T}} \mathbf{Z}_i \boldsymbol{\gamma}\} \right]^{\mathrm{T}} \boldsymbol{\beta}_0 \Bigg) / \partial(\boldsymbol{\beta}_0^{\mathrm{T}} \mathbf{Z}_i \boldsymbol{\gamma}) \Big|_{\boldsymbol{\gamma}=\boldsymbol{\gamma}^*} \boldsymbol{\beta}_0^{\mathrm{T}} \mathbf{Z}_i \boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}} \Bigg\} \\
&\quad - \Bigg\{ n^{-1} \sum_{i=1}^n \partial \Bigg(\left[g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{\alpha}_0(\mathbf{X}_i)) - E\{g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{\alpha}_0(\mathbf{X}_i)) | \boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{\alpha}_0(\mathbf{X}_i)\} \right] \boldsymbol{\Theta}_{\boldsymbol{\gamma}} \\
&\quad \times \left[\mathbf{Z}_i - E\{\mathbf{Z}_i | \boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{\alpha}_0(\mathbf{X}_i)\} \right]^{\mathrm{T}} \boldsymbol{\beta}_0 \Bigg) / \partial\{\boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{\alpha}_0(\mathbf{X}_i)\} \boldsymbol{\beta}_0^{\mathrm{T}} \mathbf{Z}_i \boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}} \Bigg\}, \\
\mathbf{T}_2 &= \Bigg(n^{-1} \sum_{i=1}^n \partial \Bigg[\left\{ g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}} \mathbf{Z}_i \boldsymbol{\gamma}) - \frac{\sum_{j=1}^J K_h(\boldsymbol{\beta}_0^{\mathrm{T}} \mathbf{Z}_j \boldsymbol{\gamma} - \boldsymbol{\beta}_0^{\mathrm{T}} \mathbf{Z}_i \boldsymbol{\gamma}) g^*(Y_j, \boldsymbol{\beta}_0^{\mathrm{T}} \mathbf{Z}_j \boldsymbol{\gamma})}{\sum_{j=1}^J K_h(\boldsymbol{\beta}_0^{\mathrm{T}} \mathbf{Z}_j \boldsymbol{\gamma} - \boldsymbol{\beta}_0^{\mathrm{T}} \mathbf{Z}_i \boldsymbol{\gamma})} \right\} \\
&\quad \times \boldsymbol{\Theta}_{\boldsymbol{\gamma}} \left\{ \mathbf{Z}_i - \frac{\sum_{j=1}^J K_h(\boldsymbol{\beta}_0^{\mathrm{T}} \mathbf{Z}_j \boldsymbol{\gamma} - \boldsymbol{\beta}_0^{\mathrm{T}} \mathbf{Z}_i \boldsymbol{\gamma}) \mathbf{Z}_j}{\sum_{j=1}^J K_h(\boldsymbol{\beta}_0^{\mathrm{T}} \mathbf{Z}_j \boldsymbol{\gamma} - \boldsymbol{\beta}_0^{\mathrm{T}} \mathbf{Z}_i \boldsymbol{\gamma})} \right\}^{\mathrm{T}} \boldsymbol{\beta}_0 \Bigg] / \partial(\boldsymbol{\beta}_0^{\mathrm{T}} \mathbf{Z}_i \boldsymbol{\gamma}) \Big|_{\boldsymbol{\gamma}=\boldsymbol{\gamma}^*} \boldsymbol{\beta}_0^{\mathrm{T}} \mathbf{Z}_i \boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}} \Bigg) \\
&\quad - \Bigg\{ n^{-1} \sum_{i=1}^n \partial \Bigg(\left[g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}} \mathbf{Z}_i \boldsymbol{\gamma}) - E\{g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}} \mathbf{Z}_i \boldsymbol{\gamma}) | \boldsymbol{\beta}_0^{\mathrm{T}} \mathbf{Z}_i \boldsymbol{\gamma}\} \right] \boldsymbol{\Theta}_{\boldsymbol{\gamma}} \\
&\quad \times \left[\mathbf{Z}_i - E\{\mathbf{Z}_i | \boldsymbol{\beta}_0^{\mathrm{T}} \mathbf{Z}_i \boldsymbol{\gamma}\} \right]^{\mathrm{T}} \boldsymbol{\beta}_0 \Bigg) / \partial(\boldsymbol{\beta}_0^{\mathrm{T}} \mathbf{Z}_i \boldsymbol{\gamma}) \Big|_{\boldsymbol{\gamma}=\boldsymbol{\gamma}^*} \boldsymbol{\beta}_0^{\mathrm{T}} \mathbf{Z}_i \boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}} \Bigg\}.
\end{aligned}
$$

Now

$$
\|\mathbf{T}_0\|_\infty = \max_{k=1,\ldots,d_{\boldsymbol{\gamma}}} \sum_{l=1}^{d_{\boldsymbol{\gamma}}} \Bigg| n^{-1} \sum_{i=1}^n \Bigg[\int_0^1 \widetilde{C}_{7i}(t) \{B_{rk}(t)\} dt
$$

$$
\begin{aligned}
& \times \int_0^1 \boldsymbol{\beta}_0^{\mathrm{T}} \mathbf{X}_i(s)\{B_{rl}(s) - B_{rl}(0)/B_{r1}(0)B_{r1}(s)\}ds\Big]\Big| \\
= & \max_{k=1,\ldots,d_\gamma} \sum_{l=1}^{d_\gamma} \Big| n^{-1} \sum_{i=1}^{n} \Big[\int_0^1 \widetilde{C}_{7i}(t)\{B_{rk}(t)\}dt \\
& \times \boldsymbol{\beta}_0^{\mathrm{T}} \mathbf{X}_i(\xi_1)\{B_{rl}(\xi_1) - B_{rl}(0)/B_{r1}(0)B_{r1}(\xi_1)\}\Big]\Big| \\
= & \max_{k=1,\ldots,d_\gamma} \Big| n^{-1} \sum_{i=1}^{n} \Big[\int_0^1 \widetilde{C}_{7i}(t)\{B_{rk}(t)\}dt \\
& \times \boldsymbol{\beta}_0^{\mathrm{T}} \mathbf{X}_i(\xi_1)\Big]\Big| \sum_{l=1}^{d_\gamma} \Big|\{B_{rl}(\xi_1) - B_{rl}(0)/B_{r1}(0)B_{r1}(\xi_1)\}\Big| \\
= & \max_{k=1,\ldots,d_\gamma} \Big| n^{-1} \sum_{i=1}^{n} \int_0^1 \widetilde{C}_{7i}(t)\{B_{rk}(t)\}dt \\
& \times \boldsymbol{\beta}_0^{\mathrm{T}} \mathbf{X}_i(\xi_1)\Big| M \\
= & \max_{k=1,\ldots,d_\gamma} \Big| n^{-1} \sum_{i=1}^{n} \int_0^1 a_n C_{7i}(t)\{B_{rk}(t)\}dt\Big|,
\end{aligned}
$$

where ξ_1 is a point in (0, 1), and

$$
\begin{aligned}
M &= \sum_{l=1}^{d_\gamma} \Big|\{B_{rl}(\xi_1) - B_{rl}(0)/B_{r1}(0)B_{r1}(\xi_1)\}\Big| \\
&\le \sum_{l=1}^{d_\gamma} B_{rl}(\xi_1) + \sum_{l=1}^{d_\gamma} B_{rl}(0)/B_{r1}(0)B_{r1}(\xi_1) \\
&< \infty,
\end{aligned}
$$

which holds by (3.4) on page 141 in DeVore and Lorentz (1993), and a_n is the sequence such that $\boldsymbol{\beta}_0^{\mathrm{T}} \mathbf{X}_i(\xi_1) = O_{a.s.}(a_n)$. Note that a_n does not need to be bounded. Here

$$
\begin{aligned}
\widetilde{C}_{7i}(t) = & \partial\Big(\big[g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)) - E\{g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i))|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\big] \\
& \times \big[\mathbf{X}_i(t)^{\mathrm{T}}\boldsymbol{\beta}_0 - E\{\mathbf{X}_i(t)^{\mathrm{T}}\boldsymbol{\beta}_0|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\big]^{\mathrm{T}}\boldsymbol{\beta}_0\Big)/\partial\{\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}
\end{aligned}
$$

and

$$
C_{7i}(t) = a_n^{-1}\partial\Big(\big[g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)) - E\{g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i))|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\big]
$$

$$\times \left[\mathbf{X}_i(t)^{\mathrm{T}}\boldsymbol{\beta}_0 - E\{\mathbf{X}_i(t)^{\mathrm{T}}\boldsymbol{\beta}_0|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\right]^{\mathrm{T}}\boldsymbol{\beta}_0\Big)/\partial\{\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}$$
$$\times M\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i(\xi_1).$$

Now $E\{C_{7i}(t)\} \neq 0$, $|E\{C_{7i}(t)\}| < \infty$, and $\|C_{7i}(\cdot)\|_2 < \infty$. By Lemma 3, we have $\|\mathbf{T}_0\|_\infty = O_p(h_b a_n)$. Similarly,

$$\|\mathbf{T}_1\|_\infty = \max_{k=1,\ldots,d_{\boldsymbol{\gamma}}} |n^{-1}\sum_{i=1}^n \int_0^1 h_b^q a_n C_{8i}(t)\{B_{rk}(t)\}dt|$$

where

$$\begin{aligned} C_{8i}(t) &= h_b^{-q}a_n^{-1}\Bigg\{\partial\Big(\left[g^*(Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}) - E\{g^*(Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma})|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}\}\right] \\ &\quad\times\left[\mathbf{X}_i(t)^{\mathrm{T}}\boldsymbol{\beta}_0 - E\{\mathbf{X}_i(t)^{\mathrm{T}}\boldsymbol{\beta}_0|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}\}\right]\Big)/\partial(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma})\Big|_{\boldsymbol{\gamma}=\boldsymbol{\gamma}^*} M\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i(\xi_1) \\ &\quad-\partial\Big(\left[g^*(Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)) - E\{g^*(Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i))|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\right] \\ &\quad\times\left[\mathbf{X}_i(t)^{\mathrm{T}}\boldsymbol{\beta}_0 - E\{\mathbf{X}_i(t)^{\mathrm{T}}\boldsymbol{\beta}_0|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\right]\Big)/\partial\{\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}M\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i(\xi_1)\Bigg\}, \end{aligned}$$

with $E\{C_{8i}(t)\} \neq 0$, $E\{C_{8i}(t)\} < \infty$, and $\|C_{8i}(\cdot)\|_2 < \infty$. By Lemma 3, we have $\|\mathbf{T}_1\|_\infty = O_p(h_b^{q+1}a_n)$. Next,

$$\|\mathbf{T}_2\|_\infty = \max_{k=1,\ldots,d_{\boldsymbol{\gamma}}} |n^{-1}\sum_{i=1}^n \int_0^1 a_n(h^2 + n^{-1/2}h^{-1/2})C_{9i}(t)\{B_{rk}(t)\}dt|,$$

where

$$\begin{aligned} C_{9i}(t) &= a_n^{-1}(h^2+n^{-1/2}h^{-1/2})^{-1}\Bigg(\partial\Bigg[\Bigg\{g^*(Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}) \\ &\quad-\frac{\sum_{j=1}^J K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma} - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma})g^*(Y_j,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma})}{\sum_{j=1}^J K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma} - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma})}\Bigg\} \\ &\quad\times\Bigg\{\mathbf{X}_i(t)^{\mathrm{T}}\boldsymbol{\beta}_0 - \frac{\sum_{j=1}^J K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma} - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma})\mathbf{X}_j(t)^{\mathrm{T}}\boldsymbol{\beta}_0}{\sum_{j=1}^J K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\boldsymbol{\gamma} - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma})}\Bigg\}\Bigg]/\partial(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma})\Bigg|_{\boldsymbol{\gamma}=\boldsymbol{\gamma}^*} \\ &\quad\times M\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i(\xi_1) - \sum_{i=1}^n\partial\Big(\left[g^*(Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}) - E\{g^*(Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma})|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}\}\right] \\ &\quad\times\left[\mathbf{X}_i(t)^{\mathrm{T}}\boldsymbol{\beta}_0 - E\{\mathbf{X}_i(t)^{\mathrm{T}}\boldsymbol{\beta}_0|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}\}\right]\Big)/\partial(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma})\Big|_{\boldsymbol{\gamma}=\boldsymbol{\gamma}^*} M\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i(\xi_1)\Bigg), \end{aligned}$$

with $E\{C_{9i}(t)\} \neq 0$, $E\{C_{9i}(t)\} < \infty$, and $\|C_{9i}(\cdot)\|_2 < \infty$. By Lemma 3, we have $\|\mathbf{T}_2\|_\infty = O_p\{h_b(h^2+n^{-1/2}h^{-1/2})a_n\}$. Combining the orders of $\mathbf{T}_0, \mathbf{T}_1$ and $\mathbf{T}_2$, we see that $\mathbf{T}_0$ clearly dominates $\mathbf{T}_1, \mathbf{T}_2$. So we can write $\mathbf{T} = \mathbf{T}_0 + o_p(\mathbf{T}_0)$.

We further analyze $\mathbf{T}_0$. Note that

$$\mathbf{T}_0 = \mathbf{T}_{00} + \mathbf{T}_{01} + \mathbf{T}_{02},$$

where

$$\begin{aligned}
\mathbf{T}_{00} &= \Bigg(n^{-1} \sum_{i=1}^n \frac{\partial \left[g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{\alpha}_0(\mathbf{X}_i)) - E\{g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{\alpha}_0(\mathbf{X}_i)) | \boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{\alpha}_0(\mathbf{X}_i)\} \right]}{\partial \boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{\alpha}_0(\mathbf{X}_i)} \boldsymbol{\Theta}_{\boldsymbol{\gamma}} \\
&\quad \times \left[\mathbf{Z}_i - E\{\mathbf{Z}_i | \boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{\alpha}_0(\mathbf{X}_i)\} \right]^{\mathrm{T}} \boldsymbol{\beta}_0 \boldsymbol{\beta}_0^{\mathrm{T}} \left[\mathbf{Z}_i - E\{\mathbf{Z}_i | \boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{\alpha}_0(\mathbf{X}_i)\} \right] \boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}} \Bigg) \\
&= \Bigg(n^{-1} \sum_{i=1}^n \boldsymbol{\Theta}_{\boldsymbol{\gamma}} \left[\mathbf{Z}_i - E\{\mathbf{Z}_i | \boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{\alpha}_0(\mathbf{X}_i)\} \right]^{\mathrm{T}} \boldsymbol{\beta}_0 \\
&\quad \times \frac{\partial \left[g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{\alpha}_0(\mathbf{X}_i)) - E\{g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{\alpha}_0(\mathbf{X}_i)) | \boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{\alpha}_0(\mathbf{X}_i)\} \right]}{\partial \boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{\alpha}_0(\mathbf{X}_i)} \boldsymbol{\beta}_0^{\mathrm{T}} \\
&\quad \left[\mathbf{Z}_i - E\{\mathbf{Z}_i | \boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{\alpha}_0(\mathbf{X}_i)\} \right] \boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}} \Bigg), \\
\mathbf{T}_{01} &= \Bigg\{ n^{-1} \sum_{i=1}^n \partial \Bigg(\left[g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{\alpha}_0(\mathbf{X}_i)) - E\{g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{\alpha}_0(\mathbf{X}_i)) | \boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{\alpha}_0(\mathbf{X}_i)\} \right] \\
&\quad / \partial \{\boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{\alpha}_0(\mathbf{X}_i)\} \boldsymbol{\Theta}_{\boldsymbol{\gamma}} \left[\mathbf{Z}_i - E\{\mathbf{Z}_i | \boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{\alpha}_0(\mathbf{X}_i)\} \right]^{\mathrm{T}} \boldsymbol{\beta}_0 \Bigg) \boldsymbol{\beta}_0^{\mathrm{T}} \\
&\quad \times [E\{\mathbf{Z}_i | \boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{\alpha}_0(\mathbf{X}_i)\}] \boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}} \Bigg\},
\end{aligned}$$

and

$$\begin{aligned}
\mathbf{T}_{02} &= \Bigg\{ n^{-1} \sum_{i=1}^n \left[g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{\alpha}_0(\mathbf{X}_i)) - E\{g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{\alpha}_0(\mathbf{X}_i)) | \boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{\alpha}_0(\mathbf{X}_i)\} \right] \\
&\quad \times \partial \Bigg(\boldsymbol{\Theta}_{\boldsymbol{\gamma}} \left[\mathbf{Z}_i - E\{\mathbf{Z}_i | \boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{\alpha}_0(\mathbf{X}_i)\} \right]^{\mathrm{T}} \boldsymbol{\beta}_0 \Bigg) / \partial \{\boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{\alpha}_0(\mathbf{X}_i)\} \boldsymbol{\beta}_0^{\mathrm{T}} \mathbf{Z}_i \boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}} \Bigg\}.
\end{aligned}$$

Now

$$\|\mathbf{T}_{00}\|_\infty = \max_{k=1,\ldots,d_{\boldsymbol{\gamma}}} |n^{-1} \sum_{i=1}^n \int_0^1 a_n C_{10i}(t) \{B_{rk}(t)\} dt|,$$

where

$$C_{10i}(t) = a_n^{-1} \left[\mathbf{X}_i(t)^{\mathrm{T}} \boldsymbol{\beta}_0 - E\{\mathbf{X}_i(t)^{\mathrm{T}} \boldsymbol{\beta}_0 | \boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{\alpha}_0(\mathbf{X}_i)\} \right]$$

$$
\begin{aligned}
&\times \frac{\partial \left[g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)) - E\{g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i))|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\right]}{\partial \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)} M \\
&\times \boldsymbol{\beta}_0^{\mathrm{T}}[\mathbf{X}_i(\xi_1) - E\{\mathbf{X}_i(\xi_1)|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}].
\end{aligned}
$$

Now $E\{C_{10i}(t)\} \neq 0$, $E\{C_{10i}(t)\} < \infty$ and $\|C_{10i}(\cdot)\|_2 < \infty$. By Lemma 3, we have $\|\mathbf{T}_{00}\|_\infty = O_p(h_b a_n)$. Similarly

$$
\|\mathbf{T}_{01} + \mathbf{T}_{02}\|_\infty = \max_{k=1,\ldots,d_{\boldsymbol{\gamma}}} |n^{-1}\sum_{i=1}^n \int_0^1 a_n C_{11i}(t)\{B_{rk}(t)\}dt|,
$$

where

$$
\begin{aligned}
C_{11i}(t) = {} & a_n^{-1}\partial\bigg(\left[g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)) - E\{g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i))|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\right] \\
& /\partial\{\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\left[\mathbf{X}_i(t)^{\mathrm{T}}\boldsymbol{\beta}_0 - E\{\mathbf{X}_i(t)^{\mathrm{T}}\boldsymbol{\beta}_0|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\right]\bigg) \\
& \times M\boldsymbol{\beta}_0^{\mathrm{T}}[E\{\mathbf{X}_i(\xi_1)|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}] + a_n^{-1}[g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)) \\
& - E\{g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i))|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}]\partial\bigg([\mathbf{X}_i(t)^{\mathrm{T}}\boldsymbol{\beta}_0 \\
& - E\{\mathbf{X}_i(t)^{\mathrm{T}}\boldsymbol{\beta}_0|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}]\bigg)/\partial\{\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}M\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i(\xi_1).
\end{aligned}
$$

Now $E\{C_{11i}(t)\} = 0$ and $\|C_{11i}(\cdot)\|_2 < \infty$. By Lemma 3, we have

$$
\|\mathbf{T}_{01} + \mathbf{T}_{02}\|_\infty = O_p\{\sqrt{h_b n^{-1}\log(n)}a_n\} = o_p(\mathbf{T}_{00}) \tag{S23}
$$

by Condition (A4). Combining the results that $\mathbf{T} = \mathbf{T}_0 + o_p(\mathbf{T}_0)$ and $\mathbf{T}_0 = \mathbf{T}_{00} + o_p(\mathbf{T}_{00})$, we have

$$
\mathbf{T} = \mathbf{T}_{00} + o_p(\mathbf{T}_{00}). \tag{S24}
$$

The Asymptotic Property of $\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)$ Combining the above results with (S5), (S21) and (S24), we obtain

$$
\begin{aligned}
& n^{1/2}\{\widehat{\boldsymbol{\gamma}}^-(\boldsymbol{\beta}_0) - \boldsymbol{\gamma}_0^-\} \\
= {} & -\mathbf{T}^{-1}\mathbf{R} \\
= {} & -\bigg\{\mathbf{T}_{00} + o_p(\mathbf{T}_{00})\bigg\}^{-1}\{\mathbf{R}_{00} + o_p(\mathbf{R}_{00})\}
\end{aligned}
$$

$$
\begin{aligned}
&= -E\Bigg(\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\left[\mathbf{Z}_i - E\{\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\right]^{\mathrm{T}}\boldsymbol{\beta}_0 \\
&\quad \times\frac{\partial\left[g^*(Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)) - E\{g^*(Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i))|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\right]}{\partial\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)} \\
&\quad \times\boldsymbol{\beta}_0^{\mathrm{T}}\left[\mathbf{Z}_i - E\{\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\right]\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\Bigg)^{-1}\{\mathbf{R}_{00} + o_p(\mathbf{R}_{00})\} \\
&\quad -\Big\{\mathbf{T}_{00} + o_p(\mathbf{T}_{00})\Big\}^{-1}\{\mathbf{R}_{00} + o_p(\mathbf{R}_{00})\} \\
&\quad +E\Bigg(\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\left[\mathbf{Z}_i - E\{\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\right]^{\mathrm{T}}\boldsymbol{\beta}_0 \\
&\quad \times\frac{\partial\left[g^*(Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)) - E\{g^*(Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i))|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\right]}{\partial\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)} \\
&\quad \times\boldsymbol{\beta}_0^{\mathrm{T}}\left[\mathbf{Z}_i - E\{\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\right]\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\Bigg)^{-1}\{\mathbf{R}_{00} + o_p(\mathbf{R}_{00})\} \\
&= -\{\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\mathbf{C}_{\mathbf{Q}}(\boldsymbol{\beta}_0)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\}^{-1}n^{-1/2}\sum_{i=1}^{n}[g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\} \\
&\quad -E\{g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}]\boldsymbol{\Theta}_{\boldsymbol{\gamma}} \\
&\quad \times\left\{\mathbf{Z}_i - E\{\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\right\}^{\mathrm{T}}\boldsymbol{\beta}_0 + o_p(\mathbf{R}_{00}) \\
&\quad -[\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\widehat{\mathbf{C}}_{\mathbf{Q}}(\boldsymbol{\beta}_0)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}} + o_p\{\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\widehat{\mathbf{C}}_{\mathbf{Q}}(\boldsymbol{\beta}_0)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\}]^{-1}\Bigg(n^{-1/2}\sum_{i=1}^{n}[g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\} \\
&\quad -E\{g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}]\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\left\{\mathbf{Z}_i - E\{\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\right\}^{\mathrm{T}}\boldsymbol{\beta}_0 \\
&\quad +o_p(\mathbf{R}_{00})\Bigg) + \{\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\mathbf{C}_{\mathbf{Q}}(\boldsymbol{\beta}_0)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\}^{-1}\Bigg(n^{-1/2}\sum_{i=1}^{n}[g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\} \\
&\quad -E\{g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}] \\
&\quad \times\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\left\{\mathbf{Z}_i - E\{\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\right\}^{\mathrm{T}}\boldsymbol{\beta}_0 + o_p(\mathbf{R}_{00})\Bigg) \\
&= -\Bigg(\{\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\mathbf{C}_{\mathbf{Q}}(\boldsymbol{\beta}_0)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\}^{-1} + o_p[\{\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\mathbf{C}_{\mathbf{Q}}(\boldsymbol{\beta}_0)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\}^{-1}]\Bigg) \\
&\quad \times\Bigg(n^{-1/2}\sum_{i=1}^{n}[g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\} - E\{g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}] \\
&\quad \boldsymbol{\Theta}_{\boldsymbol{\gamma}}\left\{\mathbf{Z}_i - E\{\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\right\}^{\mathrm{T}}\boldsymbol{\beta}_0 + o_p(\mathbf{R}_{00})\Bigg), \qquad \text{(S25)}
\end{aligned}
$$

where $\widehat{\mathbf{C}}_{\mathbf{Q}}(\boldsymbol{\beta}_0)$, $\mathbf{C}_{\mathbf{Q}}(\boldsymbol{\beta}_0)$ are defined in Corollary 1 for $\mathbf{v} = \boldsymbol{\beta}_0$ and the corresponding

$$
\begin{aligned}
&\mathbf{Q}_i(t) \\
\equiv\quad & \left(\frac{\partial\left[g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)) - E\{g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i))|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\right]}{\partial\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)}\right)^{1/2} \\
&\times[\mathbf{X}_i(t) - E\{\mathbf{X}_i(t) \mid \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}].
\end{aligned}
$$

The last equality holds because for arbitrary $d_{\boldsymbol{\gamma}-1}$-dimensional unit vector $\mathbf{u} \equiv (u_1, \ldots, u_{\boldsymbol{\gamma}-1})^{\mathrm{T}}$, we have $\mathbf{u}^{\mathrm{T}}\boldsymbol{\Theta}_{\boldsymbol{\gamma}} = \{-\sum_{k=1}^{\gamma-1} u_k B_{rk+1}(0)/B_{r1}(0), u_1, \ldots, u_{\boldsymbol{\gamma}-1}\}^{\mathrm{T}}$. Now since $B_{rk+1}(0)$ is positive for at most r k's, $\sum_{k=1}^{\gamma-1} u_k B_{rk+1}(0)/B_{r1}(0)$ is finite, which implies $\|\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\mathbf{u}\|_2^2 = 1 - u_1^2 + \{\sum_{k=1}^{\gamma-1} u_k B_{rk+1}(0)/B_{r1}(0)\}^2$ and $\|\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\|_2 = O_p(1)$. Therefore,

$$
D_7\|\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\mathbf{u}\|_2 h_b \leq \mathbf{u}^{\mathrm{T}}\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\widehat{\mathbf{C}}_{\mathbf{Q}}(\boldsymbol{\beta}_0)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\mathbf{u} \leq D_8\|\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\mathbf{u}\|_2 h_b \tag{S26}
$$

for bounded positive constants D_7, D_8 by Corollary 1, which implies the maximum and minimum eigenvalue of $\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\widehat{\mathbf{C}}_{\mathbf{Q}}(\boldsymbol{\beta}_0)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}$ is bounded between $D_7\|\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\mathbf{u}\|_2 h_b$ and $D_8\|\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\mathbf{u}\|_2 h_b$. Since $\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\widehat{\mathbf{C}}_{\mathbf{Q}}(\boldsymbol{\beta}_0)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}$ is a symmetric matrix, $\|\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\widehat{\mathbf{C}}_{\mathbf{Q}}(\boldsymbol{\beta}_0)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\|_2$ is its maximum eigenvalue, which is of order $O_p(h_b)$. Now note that $\|\{\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\widehat{\mathbf{C}}_{\mathbf{Q}}(\boldsymbol{\beta}_0)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\}^{-1}\|_2$ is the inverse of the minimum eigenvalue of $\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\widehat{\mathbf{C}}_{\mathbf{Q}}(\boldsymbol{\beta}_0)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}$. Therefore, $\|\{\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\widehat{\mathbf{C}}_{\mathbf{Q}}(\boldsymbol{\beta}_0)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\}^{-1}\|_2 = O_p(h_b^{-1})$. In summary,

$$
\begin{aligned}
\|\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\widehat{\mathbf{C}}_{\mathbf{Q}}(\boldsymbol{\beta}_0)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\|_2 &= O_p(h_b), \\
\|\{\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\widehat{\mathbf{C}}_{\mathbf{Q}}(\boldsymbol{\beta}_0)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\}^{-1}\|_2 &= O_p(h_b^{-1})
\end{aligned} \tag{S27}
$$

Similar to those arguments that led to (S27), we have

$$
\begin{aligned}
\|\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\mathbf{C}_{\mathbf{Q}}(\boldsymbol{\beta}_0)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\|_2 &= O_p(h_b), \\
\|\{\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\mathbf{C}_{\mathbf{Q}}(\boldsymbol{\beta}_0)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\}^{-1}\|_2 &= O_p(h_b^{-1}).
\end{aligned} \tag{S28}
$$

Now because as shown in Corollary 1, $\|\widehat{\mathbf{C}}_{\mathbf{Q}}(\boldsymbol{\beta}_0) - \mathbf{C}_{\mathbf{Q}}(\boldsymbol{\beta}_0)\|_2 \leq o_p\{h_b n^{(h_b-1)/2}\}$ and $\|\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\|_2 = O_p(1)$, we have

$$
\|\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\widehat{\mathbf{C}}_{\mathbf{Q}}(\boldsymbol{\beta}_0)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}} + o_p\{\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\widehat{\mathbf{C}}_{\mathbf{Q}}(\boldsymbol{\beta}_0)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\} - \boldsymbol{\Theta}_{\boldsymbol{\gamma}}\mathbf{C}_{\mathbf{Q}}(\boldsymbol{\beta}_0)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\|_2
$$

$$
\begin{aligned}
&\leq \|\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\widehat{\mathbf{C}}_{\mathbf{Q}}(\boldsymbol{\beta}_0)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}} - \boldsymbol{\Theta}_{\boldsymbol{\gamma}}\mathbf{C}_{\mathbf{Q}}(\boldsymbol{\beta}_0)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\|_2 + o_p\{\|\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\widehat{\mathbf{C}}_{\mathbf{Q}}(\boldsymbol{\beta}_0)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\|_2\} \\
&\leq o_p\{h_b n^{(h_b-1)/2}\} + o_p(h_b).
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
&\|[\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\widehat{\mathbf{C}}_{\mathbf{Q}}(\boldsymbol{\beta}_0)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}} + o_p\{\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\widehat{\mathbf{C}}_{\mathbf{Q}}(\boldsymbol{\beta}_0)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\}]^{-1} - \{\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\mathbf{C}_{\mathbf{Q}}(\boldsymbol{\beta}_0)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\}^{-1}\|_2 \\
= &\|\{\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\mathbf{C}_{\mathbf{Q}}(\boldsymbol{\beta}_0)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\}^{-2}[\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\widehat{\mathbf{C}}_{\mathbf{Q}}(\boldsymbol{\beta}_0)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}} + o_p\{\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\widehat{\mathbf{C}}_{\mathbf{Q}}(\boldsymbol{\beta}_0)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\} \\
&-\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\ \mathbf{C}_{\mathbf{Q}}(\boldsymbol{\beta}_0)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}]\{1+o_p(1)\}\|_2 \\
\leq &\ o_p\{h_b^{-2}h_b n^{(h_b-1)/2}\} + o_p(h_b^{-2}h_b) \\
= &\ o_p\{h_b^{-1}n^{(h_b-1)/2}\} + o_p(h_b^{-1}) \\
= &\ o_p[\{\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\mathbf{C}_{\mathbf{Q}}(\boldsymbol{\beta}_0)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\}^{-1}]. \qquad \text{(S29)}
\end{aligned}
$$

As a result, (S25) holds and we can write

$$
\begin{aligned}
&n^{1/2}\{\widehat{\boldsymbol{\gamma}}^{-}(\boldsymbol{\beta}_0) - \boldsymbol{\gamma}_0^{-}\} \\
= &-(\{\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\mathbf{C}_{\mathbf{Q}}(\boldsymbol{\beta}_0)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\}^{-1} + o_p[\{\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\mathbf{C}_{\mathbf{Q}}(\boldsymbol{\beta}_0)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\}^{-1}])\Bigg(n^{-1/2}\sum_{i=1}^n[g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\} \\
&E\{g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}]\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\left[\mathbf{Z}_i - E\{\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\right]^{\mathrm{T}}\boldsymbol{\beta}_0 + o_p(\mathbf{R}_{00})\Bigg) \\
= &-\Bigg(\Bigg\{E\Bigg(\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\left[\mathbf{Z}_i - E\{\mathbf{Z}_i|\boldsymbol{\beta}_0\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\right]^{\mathrm{T}}\boldsymbol{\beta}_0 \\
&\times\frac{\partial\left(g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\} - E[g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)]\right]\right)}{\partial\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)} \\
&\times\boldsymbol{\beta}_0^{\mathrm{T}}[\mathbf{Z}_i - E\{\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}]\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\Bigg)\Bigg\}^{-1} + o_p[\{\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\mathbf{C}_{\mathbf{Q}}(\boldsymbol{\beta}_0)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\}^{-1}]\Bigg) \\
&\times\Bigg(n^{-1/2}\sum_{i=1}^n[g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\} - E\{g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}]\boldsymbol{\Theta}_{\boldsymbol{\gamma}} \\
&\times\left[\mathbf{Z}_i - E\{\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\right]^{\mathrm{T}}\boldsymbol{\beta}_0 + o_p(\mathbf{R}_{00})\Bigg) \\
= &\ \mathbf{L} + o_p(\mathbf{L}),
\end{aligned}
$$

where recall that we have defined in the statement that

$$
\mathbf{L} = -\{\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\mathbf{C}_{\mathbf{Q}}(\boldsymbol{\beta}_0)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\}^{-1}\Bigg(n^{-1/2}\sum_{i=1}^n[g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}
$$

$$
\begin{aligned}
& -E\{g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}]\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\left[\mathbf{Z}_i - E\{\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\right]^{\mathrm{T}}\boldsymbol{\beta}_0\Bigg) \\
= & -\Bigg(\Bigg\{E\Bigg(\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\left[\mathbf{Z}_i - E\{\mathbf{Z}_i|\boldsymbol{\beta}_0\boldsymbol{\alpha}_0(\mathbf{X}_i\}\right]^{\mathrm{T}}\boldsymbol{\beta}_0 \qquad \text{(S30)} \\
& \times\frac{\partial\left(g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\} - E[g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)]\right)}{\partial\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)} \\
& \times\boldsymbol{\beta}_0^{\mathrm{T}}[\mathbf{Z}_i - E\{\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}]\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\Bigg)\Bigg\}^{-1}\Bigg)\Bigg(n^{-1/2}\sum_{i=1}^{n}[g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\} \\
& -E\{g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}]\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\left[\mathbf{Z}_i - E\{\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\right]^{\mathrm{T}}\boldsymbol{\beta}_0\Bigg).
\end{aligned}
$$

Now for arbitrary $d_{\boldsymbol{\gamma}-1}$-dimensional vector $\mathbf{a}$ with $\|\mathbf{a}\|_2 < \infty$ by using the same argument as those lead to (S23), we have

$$
\begin{aligned}
[\mathbf{a}^{\mathrm{T}}n^{1/2}\{\widehat{\boldsymbol{\gamma}}^{-}(\boldsymbol{\beta}_0) - \boldsymbol{\gamma}_0^{-}\}]^2 &= \mathbf{a}^{\mathrm{T}}\mathbf{L}\mathbf{L}^{\mathrm{T}}\mathbf{a} + o_p(\mathbf{a}^{\mathrm{T}}\mathbf{L}\mathbf{L}^{\mathrm{T}}\mathbf{a}) \\
&= \mathbf{L}_1 + o_p(\mathbf{L}_1),
\end{aligned}
$$

where

$$
\begin{aligned}
\mathbf{L}_1 = & -\|\mathbf{a}^{\mathrm{T}}\{\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\mathbf{C}_{\mathbf{Q}}(\boldsymbol{\beta}_0)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\}^{-1}\|_2^2\frac{\mathbf{a}^{\mathrm{T}}\{\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\mathbf{C}_{\mathbf{Q}}(\boldsymbol{\beta}_0)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\}^{-1}}{\|\mathbf{a}^{\mathrm{T}}\{\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\mathbf{C}_{\mathbf{Q}}(\boldsymbol{\beta}_0)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\}^{-1}\|_2} \\
& \times n^{-1}\sum_{i=1}^{n}\Bigg([g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\} - E\{g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}]\boldsymbol{\Theta}_{\boldsymbol{\gamma}}[\mathbf{Z}_i \\
& -E\{\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}]^{\mathrm{T}}\boldsymbol{\beta}_0\Bigg)^{\otimes 2}\frac{\{\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\mathbf{C}_{\mathbf{Q}}(\boldsymbol{\beta}_0)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\}^{-1}\mathbf{a}}{\|\mathbf{a}^{\mathrm{T}}\{\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\mathbf{C}_{\mathbf{Q}}(\boldsymbol{\beta}_0)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\}^{-1}\|_2} \\
= & -\|\mathbf{a}^{\mathrm{T}}\{\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\mathbf{C}_{\mathbf{Q}}(\boldsymbol{\beta}_0)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\}^{-1}\|_2^2\frac{\mathbf{a}^{\mathrm{T}}\{\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\mathbf{C}_{\mathbf{Q}}(\boldsymbol{\beta}_0)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\}^{-1}}{\|\mathbf{a}^{\mathrm{T}}\{\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\mathbf{C}_{\mathbf{Q}}(\boldsymbol{\beta}_0)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\}^{-1}\|_2} \\
& n^{-1}\sum_{i=1}^{n}\Bigg([g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\} - E\{g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}]\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\{\mathbf{Z}_i \\
& -E\{\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\}^{\mathrm{T}}\boldsymbol{\beta}_0\Bigg)^{\otimes 2}\frac{\{\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\mathbf{C}_{\mathbf{Q}}(\boldsymbol{\beta}_0)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\}^{-1}\mathbf{a}}{\|\mathbf{a}^{\mathrm{T}}\{\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\mathbf{C}_{\mathbf{Q}}(\boldsymbol{\beta}_0)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\}^{-1}\|_2} \\
\leq & \; O_p(h_b^{-2}h_b) \\
= & \; O_p(h_b^{-1}).
\end{aligned}
$$

The third line holds by noting that $\|\{\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\mathbf{C}_{\mathbf{Q}}(\boldsymbol{\beta}_0)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\}^{-1}\| = O_p(h_b^{-1})$ by (S28) and

$$n^{-1}\sum_{i=1}^{n}\Bigg([g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\} - E\{g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}]\boldsymbol{\Theta}_{\boldsymbol{\gamma}}$$
$$\times\left\{\mathbf{Z}_i - E\{\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\right\}^{\mathrm{T}}\boldsymbol{\beta}_0\Bigg)^{\otimes 2}$$

is in the form of $\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\widehat{\mathbf{C}}_{\mathbf{Q}}(\boldsymbol{\beta}_0)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}$, where $\widehat{\mathbf{C}}_{\mathbf{Q}}(\boldsymbol{\beta}_0)$ is defined in Corollary 1 with $\mathbf{v} = \boldsymbol{\beta}_0$ and

$$\begin{aligned}
&\mathbf{Q}_i(t)\\
\equiv\; &[g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\} - E\{g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}]\\
&\times[\mathbf{X}_i(t) - E\{\mathbf{X}_i(t) \mid \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}].
\end{aligned}$$

Using the arguments as those led to (S26), we have

$$\begin{aligned}
D_{10}h_b \;\le\; & \frac{\mathbf{a}^{\mathrm{T}}\{\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\mathbf{C}_{\mathbf{Q}}(\boldsymbol{\beta}_0)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\}^{-1}}{\|\mathbf{a}^{\mathrm{T}}\{\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\mathbf{C}_{\mathbf{Q}}(\boldsymbol{\beta}_0)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\}^{-1}\|_2} n^{-1}\sum_{i=1}^{n}\Bigg([g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\\
& -E\{g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}]\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\left\{\mathbf{Z}_i - E\{\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\right\}^{\mathrm{T}}\boldsymbol{\beta}_0\Bigg)^{\otimes 2}\\
& \times\frac{\{\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\mathbf{C}_{\mathbf{Q}}(\boldsymbol{\beta}_0)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\}^{-1}\mathbf{a}}{\|\mathbf{a}^{\mathrm{T}}\{\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\mathbf{C}_{\mathbf{Q}}(\boldsymbol{\beta}_0)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\}^{-1}\|_2} \le D_{11}h_b,
\end{aligned} \tag{S31}$$

where D_{10} and D_{11} are positive constants. Therefore $\mathbf{a}^{\mathrm{T}}\{\widehat{\boldsymbol{\gamma}}^{-}(\boldsymbol{\beta}_0) - \boldsymbol{\gamma}_0^{-}\} = O_p(n^{-1/2}h_b^{-1/2})$. This proves the result. □

S3.2 Proof of Theorem 2

The proof of Theorem 2 is divided into two parts. In Part I, we show that $\sqrt{n}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})$ convergences to a random quantity which exists under the approximated model. In Part II, we show that this quantity converges to a Gaussian vector defined on the measurable space under the true model.

Part I

By the consistency shown in Proposition 2, we expand the score function as

$$\mathbf{0} \;=\; n^{-1/2}\sum_{i=1}^{n}\Big[g^*\{Y_i,\widehat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\widehat{\boldsymbol{\beta}})\}$$

$$
\begin{aligned}
& -\frac{\sum_{j=1}^{J} K_h\{\widehat{\boldsymbol\beta}^{\mathrm T}\mathbf Z_j\widehat{\boldsymbol\gamma}(\widehat{\boldsymbol\beta})-\widehat{\boldsymbol\beta}^{\mathrm T}\mathbf Z_i\widehat{\boldsymbol\gamma}(\widehat{\boldsymbol\beta})\}g^*\{Y_j,\widehat{\boldsymbol\beta}^{\mathrm T}\mathbf Z_j\widehat{\boldsymbol\gamma}(\widehat{\boldsymbol\beta})\}}{\sum_{j=1}^{J} K_h\{\widehat{\boldsymbol\beta}^{\mathrm T}\mathbf Z_j\widehat{\boldsymbol\gamma}(\widehat{\boldsymbol\beta})-\widehat{\boldsymbol\beta}^{\mathrm T}\mathbf Z_i\widehat{\boldsymbol\gamma}(\widehat{\boldsymbol\beta})\}}\Bigg]\boldsymbol\Theta_{\boldsymbol\beta} \\
& \times\left[\mathbf Z_i-\frac{\sum_{j=1}^{J} K_h\{\widehat{\boldsymbol\beta}^{\mathrm T}\mathbf Z_j\widehat{\boldsymbol\gamma}(\widehat{\boldsymbol\beta})-\widehat{\boldsymbol\beta}^{\mathrm T}\mathbf Z_i\widehat{\boldsymbol\gamma}(\widehat{\boldsymbol\beta})\}\mathbf Z_j}{\sum_{j=1}^{J} K_h\{\widehat{\boldsymbol\beta}^{\mathrm T}\mathbf Z_j\widehat{\boldsymbol\gamma}(\widehat{\boldsymbol\beta})-\widehat{\boldsymbol\beta}^{\mathrm T}\mathbf Z_i\widehat{\boldsymbol\gamma}(\widehat{\boldsymbol\beta})\}}\right]\widehat{\boldsymbol\gamma}(\widehat{\boldsymbol\beta}) \\
= \; & \mathbf G+\mathbf H n^{1/2}(\widehat{\boldsymbol\beta}-\boldsymbol\beta_0), \qquad \text{(S32)}
\end{aligned}
$$

where

$$
\begin{aligned}
\mathbf G \;=\; & n^{-1/2}\sum_{i=1}^{n}\Bigg[g^*\{Y_i,\boldsymbol\beta_0^{\mathrm T}\mathbf Z_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)\} \\
& -\frac{\sum_{j=1}^{J} K_h\{\boldsymbol\beta_0^{\mathrm T}\mathbf Z_j\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)-\boldsymbol\beta_0^{\mathrm T}\mathbf Z_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)\}g^*\{Y_j,\boldsymbol\beta_0^{\mathrm T}\mathbf Z_j\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)\}}{\sum_{j=1}^{J} K_h\{\boldsymbol\beta_0^{\mathrm T}\mathbf Z_j\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)-\boldsymbol\beta_0^{\mathrm T}\mathbf Z_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)\}}\Bigg]\boldsymbol\Theta_{\boldsymbol\beta} \\
& \times\left[\mathbf Z_i-\frac{\sum_{j=1}^{J} K_h\{\boldsymbol\beta_0^{\mathrm T}\mathbf Z_j\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)-\boldsymbol\beta_0^{\mathrm T}\mathbf Z_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)\}\mathbf Z_j}{\sum_{j=1}^{J} K_h\{\boldsymbol\beta_0^{\mathrm T}\mathbf Z_j\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)-\boldsymbol\beta_0^{\mathrm T}\mathbf Z_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)\}}\right]\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)
\end{aligned}
$$

and

$$
\begin{aligned}
\mathbf H \;=\; & n^{-1}\sum_{i=1}^{n}\partial\Bigg[g^*\{Y_i,\boldsymbol\beta^{\mathrm T}\mathbf Z_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta)\} \\
& -\frac{\sum_{j=1}^{J} K_h\{\boldsymbol\beta^{\mathrm T}\mathbf Z_j\widehat{\boldsymbol\gamma}(\boldsymbol\beta)-\boldsymbol\beta^{\mathrm T}\mathbf Z_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta)\}g^*\{Y_j,\boldsymbol\beta^{\mathrm T}\mathbf Z_j\widehat{\boldsymbol\gamma}(\boldsymbol\beta)\}}{\sum_{j=1}^{J} K_h\{\boldsymbol\beta^{\mathrm T}\mathbf Z_j\widehat{\boldsymbol\gamma}(\boldsymbol\beta)-\boldsymbol\beta^{\mathrm T}\mathbf Z_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta)\}}\Bigg]\boldsymbol\Theta_{\boldsymbol\beta} \\
& \times\left[\mathbf Z_i-\frac{\sum_{j=1}^{J} K_h\{\boldsymbol\beta^{\mathrm T}\mathbf Z_j\widehat{\boldsymbol\gamma}(\boldsymbol\beta)-\boldsymbol\beta^{\mathrm T}\mathbf Z_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta)\}\mathbf Z_j}{\sum_{j=1}^{J} K_h\{\boldsymbol\beta^{\mathrm T}\mathbf Z_j\widehat{\boldsymbol\gamma}(\boldsymbol\beta)-\boldsymbol\beta^{\mathrm T}\mathbf Z_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta)\}}\right]\widehat{\boldsymbol\gamma}(\boldsymbol\beta)/\partial\boldsymbol\beta^{\mathrm T}|_{\boldsymbol\beta=\boldsymbol\beta^*},
\end{aligned}
$$

where $\boldsymbol\beta^*$ is the point on the line connecting $\widehat{\boldsymbol\beta}$ and $\boldsymbol\beta_0$. In the following, we study the asymptotic properties for $\mathbf G$ and $\mathbf H$ separately.

The Asymptotic Property for G Now we can decompose $\mathbf G$ as

$$
\mathbf G=\mathbf G_0+\mathbf G_1+\mathbf G_2+\mathbf G_3,
$$

where

$$
\begin{aligned}
\mathbf G_0 \;=\; & n^{-1/2}\sum_{i=1}^{n}\left(g^*\{Y_i,\boldsymbol\beta_0^{\mathrm T}\mathbf Z_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)\}-E[g^*\{Y_i,\boldsymbol\beta_0^{\mathrm T}\mathbf Z_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)\}|\boldsymbol\beta_0^{\mathrm T}\mathbf Z_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)]\right)\boldsymbol\Theta_{\boldsymbol\beta} \\
& \times\left[\mathbf Z_i-E\{\mathbf Z_i|\boldsymbol\beta_0^{\mathrm T}\mathbf Z_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)\}\right]\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0),
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{G}_1 &= n^{-1/2}\sum_{i=1}^{n}\left(g^*\{Y_i,\boldsymbol\beta_0^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)\} - E[g^*\{Y_i,\boldsymbol\beta_0^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)\}|\boldsymbol\beta_0^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)]\right)\boldsymbol\Theta_{\boldsymbol\beta}\\
&\quad\times\left[E\{\mathbf{Z}_i|\boldsymbol\beta_0^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)\} - \frac{\sum_{j=1}^{J}K_h\{\boldsymbol\beta_0^{\mathrm T}\mathbf{Z}_j\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)-\boldsymbol\beta_0^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)\}\mathbf{Z}_j}{\sum_{j=1}^{J}K_h\{\boldsymbol\beta_0^{\mathrm T}\mathbf{Z}_j\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)-\boldsymbol\beta_0^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)\}}\right]\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0),\\
\mathbf{G}_2 &= n^{-1/2}\sum_{i=1}^{n}\Bigg(E[g^*\{Y_i,\boldsymbol\beta_0^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)\}|\boldsymbol\beta_0^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)]\\
&\quad-\frac{\sum_{j=1}^{J}K_h\{\boldsymbol\beta_0^{\mathrm T}\mathbf{Z}_j\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)-\boldsymbol\beta_0^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)\}g^*\{Y_j,\boldsymbol\beta_0^{\mathrm T}\mathbf{Z}_j\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)\}}{\sum_{j=1}^{J}K_h\{\boldsymbol\beta_0^{\mathrm T}\mathbf{Z}_j\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)-\boldsymbol\beta_0^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)\}}\Bigg)\boldsymbol\Theta_{\boldsymbol\beta}\\
&\quad\times\left[\mathbf{Z}_i - E\{\mathbf{Z}_i|\boldsymbol\beta_0^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)\}\right]\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0),\\
\mathbf{G}_3 &= n^{-1/2}\sum_{i=1}^{n}\Bigg(E[g^*\{Y_i,\boldsymbol\beta_0^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)\}|\boldsymbol\beta_0^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)]\\
&\quad-\frac{\sum_{j=1}^{J}K_h\{\boldsymbol\beta_0^{\mathrm T}\mathbf{Z}_j\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)-\boldsymbol\beta_0^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)\}g^*\{Y_j,\boldsymbol\beta_0^{\mathrm T}\mathbf{Z}_j\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)\}}{\sum_{j=1}^{J}K_h\{\boldsymbol\beta_0^{\mathrm T}\mathbf{Z}_j\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)-\boldsymbol\beta_0^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)\}}\Bigg)\boldsymbol\Theta_{\boldsymbol\beta}\\
&\quad\times\left[E\{\mathbf{Z}_i|\boldsymbol\beta_0^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)\} - \frac{\sum_{j=1}^{J}K_h\{\boldsymbol\beta_0^{\mathrm T}\mathbf{Z}_j\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)-\boldsymbol\beta_0^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)\}\mathbf{Z}_j}{\sum_{j=1}^{J}K_h\{\boldsymbol\beta_0^{\mathrm T}\mathbf{Z}_j\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)-\boldsymbol\beta_0^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)\}}\right]\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0).
\end{aligned}
$$

We can further write $\mathbf{G}_1 = \mathbf{G}_{11} + \mathbf{G}_{12}$, where

$$
\begin{aligned}
\mathbf{G}_{11} &= n^{-1/2}\sum_{i=1}^{n}\left(g^*\{Y_i,\boldsymbol\beta_0^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)\} - E[g^*\{Y_i,\boldsymbol\beta_0^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)\}|\boldsymbol\beta_0^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)]\right)\boldsymbol\Theta_{\boldsymbol\beta}\\
&\quad\times\Bigg\{\frac{E\{\mathbf{Z}_i|\boldsymbol\beta_0^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)\}\sum_{j=1}^{J}K_h\{\boldsymbol\beta_0^{\mathrm T}\mathbf{Z}_j\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)-\boldsymbol\beta_0^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)\}}{nf_{\mathbf{Z}}\{\boldsymbol\beta_0^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)\}}\\
&\quad-\frac{\sum_{j=1}^{J}K_h\{\boldsymbol\beta_0^{\mathrm T}\mathbf{Z}_j\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)-\boldsymbol\beta_0^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)\}\mathbf{Z}_j}{nf_{\mathbf{Z}}\{\boldsymbol\beta_0^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)\}}\Bigg\}\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0),\\
\mathbf{G}_{12} &= n^{-1/2}\sum_{i=1}^{n}\left(g^*\{Y_i,\boldsymbol\beta_0^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)\} - E[g^*\{Y_i,\boldsymbol\beta_0^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)\}|\boldsymbol\beta_0^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)]\right)\boldsymbol\Theta_{\boldsymbol\beta}\\
&\quad\times\Bigg\{\frac{E\{\mathbf{Z}_i\mid\boldsymbol\beta_0^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)\}\sum_{j=1}^{J}K_h\{\boldsymbol\beta_0^{\mathrm T}\mathbf{Z}_j\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)-\boldsymbol\beta_0^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)\}}{nf_{\mathbf{Z}}\{\boldsymbol\beta_0^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)\}}\\
&\quad-\frac{\sum_{j=1}^{J}K_h\{\boldsymbol\beta_0^{\mathrm T}\mathbf{Z}_j\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)-\boldsymbol\beta_0^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)\}\mathbf{Z}_j}{nf_{\mathbf{Z}}\{\boldsymbol\beta_0^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)\}}\Bigg\}\\
&\quad\times\left\{\frac{nf_{\mathbf{Z}}\{\boldsymbol\beta_0^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)\}}{\sum_{j=1}^{J}K_h\{\boldsymbol\beta_0^{\mathrm T}\mathbf{Z}_j\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)-\boldsymbol\beta_0^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)\}}-1\right\}\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)\\
&= \mathbf{G}_{11}\left\{\frac{nf_{\mathbf{Z}}\{\boldsymbol\beta_0^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)\}}{\sum_{j=1}^{J}K_h\{\boldsymbol\beta_0^{\mathrm T}\mathbf{Z}_j\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)-\boldsymbol\beta_0^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta_0)\}}-1\right\}.
\end{aligned}
$$

Note that $\mathbf{G}_{11}, \mathbf{G}_{12}$ are of finite dimension, so the sup norm and the L_2 norm are equivalent. Using the U-statistics techniques similar to those led to (S11) and (S13), we get

$$\begin{aligned}
\mathbf{G}_{11} &= n^{-3/2}\sum_{i=1}^{n}\sum_{j=1}^{J}\left(g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)\} - E[g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)]\right)\boldsymbol{\Theta}_{\boldsymbol{\beta}} \\
&\quad\times\left\{\frac{E\{\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)\}K_h\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0) - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)\}}{nf_{\mathbf{Z}}\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)\}}\right. \\
&\quad\left.-\frac{K_h\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0) - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)\}\mathbf{Z}_j}{nf_{\mathbf{Z}}\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)\}}\right\}\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0) \\
&= \mathbf{G}_{111} + \mathbf{G}_{112} - \mathbf{G}_{113} + o_p(\mathbf{G}_{111} + \mathbf{G}_{112} - \mathbf{G}_{113}),
\end{aligned}$$

where

$$\begin{aligned}
\mathbf{G}_{111} &= n^{-1/2}\sum_{i=1}^{n}E\Bigg[\left(g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)\} - E[g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)]\right)\boldsymbol{\Theta}_{\boldsymbol{\beta}} \\
&\quad\times\left\{\frac{E\{\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)\}K_h\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0) - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)\}}{nf_{\mathbf{Z}}\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)\}}\right. \\
&\quad\left.-\frac{K_h\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0) - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)\}\mathbf{Z}_j}{nf_{\mathbf{Z}}\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)\}}\right\}\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)|\mathbf{O}_i\Bigg], \\
\mathbf{G}_{112} &= n^{-1/2}\sum_{j=1}^{J}E\Bigg[\left(g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)\} - E[g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)]\right)\boldsymbol{\Theta}_{\boldsymbol{\beta}} \\
&\quad\times\left\{\frac{E\{\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)\}K_h\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0) - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)\}}{nf_{\mathbf{Z}}\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)\}}\right. \\
&\quad\left.-\frac{K_h\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0) - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)\}\mathbf{Z}_j}{nf_{\mathbf{Z}}\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)\}}\right\}\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)|\mathbf{O}_j\Bigg], \\
\mathbf{G}_{113} &= n^{1/2}E\Bigg[\left(g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)\} - E[g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)]\right)\boldsymbol{\Theta}_{\boldsymbol{\beta}} \\
&\quad\times\left\{\frac{E\{\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)\}K_h\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0) - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)\}}{nf_{\mathbf{Z}}\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)\}}\right. \\
&\quad\left.-\frac{K_h\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0) - \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)\}\mathbf{Z}_j}{nf_{\mathbf{Z}}\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)\}}\right\}\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)\Bigg].
\end{aligned}$$

Similar to $\mathbf{R}_{111}$, using the same derivation in (S7), we can write the summand of $\mathbf{G}_{111}$ as

$$E\Bigg[\left(g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)\} - E[g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)]\right)\boldsymbol{\Theta}_{\boldsymbol{\beta}}$$

$$
\begin{aligned}
& \times\Bigg\{ \frac{E\{\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)\}K_h\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)-\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)\}}{nf_{\mathbf{Z}}\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)\}} \\
& -\frac{K_h\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)-\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)\}\mathbf{Z}_j}{nf_{\mathbf{Z}}\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)\}}\Bigg\}\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)|\mathbf{O}_i\Bigg] \\
= & \left(g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)\}-E[g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)]\right)\boldsymbol{\Theta}_{\boldsymbol{\beta}} \\
& \times\Bigg(E\{\mathbf{Z}_i\mid\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)\}\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)\int\frac{\partial^2 f_{\mathbf{Z}}[c\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0),sh\}]}{\partial c\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0),sh\}^2}\frac{s^2K(s)}{2f_{\mathbf{Z}}\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)\}}ds \\
& -\int\frac{\partial^2[E\{\mathbf{Z}_i\mid d\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0),sh\}\}\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)f_{\mathbf{Z}}\{d\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0),sh\}\}]}{\partial d\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0),sh\}^2} \\
& \times\frac{s^2K(s)}{2f_{\mathbf{Z}}\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)\}}ds\Bigg)h^2,
\end{aligned}
$$

which is a mean $\mathbf{0}$ finite dimensional vector. Here $c\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0),sh\}$ and $d\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0),sh\}$ are points on the line connecting $\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)$ and $\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)+sh$. Therefore, we have $\|h^{-2}\mathbf{G}_{111}\|=O_p(1)$, and in turn $\mathbf{G}_{111}$ and $\mathbf{G}_{113}$ have the order of $O_p(h^2)$. Further, similar to $\mathbf{R}_{112}$, we have $\mathbf{G}_{112}=\mathbf{0}$. As a result, $\mathbf{G}_{11}=\mathbf{G}_{111}+\mathbf{G}_{112}-\mathbf{G}_{113}+o_p(\mathbf{G}_{111}+\mathbf{G}_{112}-\mathbf{G}_{113})=O_p(h^2)$. Now similar to (S12), since

$$
\left\{\frac{nf_{\mathbf{Z}}\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)\}}{\sum_{j=1}^{J}K_h\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_j\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)-\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)\}}-1\right\}=O_p\{h^2+(nh)^{-1/2}\},
$$

we obtain $\mathbf{G}_{12}=o_p(\mathbf{G}_{11})$. Thus, we have

$$
\|\mathbf{G}_1\|_2=\|\mathbf{G}_{11}+\mathbf{G}_{12}\|_2=O_p(h^2). \tag{S33}
$$

Using similar derivation as those led to (S33) while exchanging the roles of $\mathbf{Z}$ and g^*, we obtain $\|\mathbf{G}_2\|_2=O_p(h^2)$. Using the same reasoning as those led to (S16), and noting that $\mathbf{G}_3$ is a finite dimensional vector, we have that the order of $\mathbf{G}_3$ is the same as $n^{1/2}$ times the square of the order for kernel estimation errors, i.e., $\|\mathbf{G}_3\|_2=O_p(n^{1/2}h^4+n^{-1/2}h^{-1})=o_p(1)$ under Condition (A2).

Now we can further write $\mathbf{G}_0=\mathbf{G}_{00}+\mathbf{G}_{01}+\mathbf{G}_{02}$, where

$$
\begin{aligned}
\mathbf{G}_{00} = & n^{-1/2}\sum_{i=1}^{n}\left(g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}-E[g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)]\right)\boldsymbol{\Theta}_{\boldsymbol{\beta}} \\
& \times\left[\boldsymbol{\alpha}_0(\mathbf{X}_i)-E\{\boldsymbol{\alpha}_0(\mathbf{X}_i)|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\right],
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{G}_{01} &= n^{-1/2}\sum_{i=1}^{n}\left(g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)\} - E[g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)]\right)\boldsymbol{\Theta}_{\boldsymbol{\beta}}\\
&\quad\times\left[\mathbf{Z}_i - E\{\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)\}\right]\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0) - n^{-1/2}\sum_{i=1}^{n}[g^*(Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)\\
&\quad-E\{g^*(Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}]\boldsymbol{\Theta}_{\boldsymbol{\beta}}\left\{\mathbf{Z}_i - E(\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)\right\}\boldsymbol{\gamma}_0,\\
\mathbf{G}_{02} &= n^{-1/2}\sum_{i=1}^{n}\left[g^*(Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0) - E\{g^*(Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}\right]\boldsymbol{\Theta}_{\boldsymbol{\beta}}\{\mathbf{Z}_i\\
&\quad-E(\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)\}\boldsymbol{\gamma}_0 - n^{-1/2}\sum_{i=1}^{n}(g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\\
&\quad-E[g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)])\boldsymbol{\Theta}_{\boldsymbol{\beta}}\left[\boldsymbol{\alpha}_0(\mathbf{X}_i) - E\{\boldsymbol{\alpha}_0(\mathbf{X}_i)|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\right]\\
&= n^{-1/2}\sum_{i=1}^{n}\left(g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\} - E[g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)]\right)\boldsymbol{\Theta}_{\boldsymbol{\beta}}\\
&\quad\times\frac{\partial\left[\boldsymbol{\alpha}(\mathbf{X}_i) - E\{\boldsymbol{\alpha}(\mathbf{X}_i)|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}(\mathbf{X}_i)\}\right]}{\partial\boldsymbol{\alpha}(\mathbf{X}_i)}|_{\boldsymbol{\alpha}=\boldsymbol{\alpha}^*(\mathbf{X}_i,\mathbf{Z}_i)\}}\{\mathbf{Z}_i\boldsymbol{\gamma}_0 - \boldsymbol{\alpha}_0(\mathbf{X}_i)\},
\end{aligned}
$$

where $\boldsymbol{\alpha}^*(\mathbf{X}_i,\mathbf{Z}_i)$ is the point on the line connecting $\boldsymbol{\alpha}_0(\mathbf{X}_i)$ and $\mathbf{Z}_i\boldsymbol{\gamma}_0$. Clearly we have $E(\mathbf{G}_{02}) = 0$. Further from $\|\mathbf{B}_r(\cdot)^{\mathrm{T}}\boldsymbol{\gamma}_0 - \alpha_0(\cdot)\|_\infty = O_p(h_b^q)$, we have

$$
\begin{aligned}
&\|\mathbf{Z}_i\boldsymbol{\gamma}_0 - \boldsymbol{\alpha}_0(\mathbf{X}_i)\|_2\\
=\ &\|\int_0^1 \mathbf{B}_r(t)^{\mathrm{T}}\boldsymbol{\gamma}_0\mathbf{X}_i(t)dt - \int_0^1 \alpha_0(t)\mathbf{X}_i(t)dt\|_2\\
\le\ &\left\{\int_0^1 |\mathbf{B}_r(t)^{\mathrm{T}}\boldsymbol{\gamma}_0 - \alpha_0(t)|^2 dt\right\}^{1/2}\|\mathbf{X}_i(\cdot)\|_2\\
=\ &O_p(h_b^q),
\end{aligned}
$$

so $\|h_b^{-q}\mathbf{G}_{02}\|_2 = O_p(1)$. Therefore, $\|\mathbf{G}_{02}\|_2 = O_p(h_b^q)$.

Next

$$
\begin{aligned}
&\mathbf{G}_{01}\\
=\ &\left(n^{-1}\sum_{i=1}^{n}\frac{\partial\left[g^*(Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0) - E\{g^*(Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}\right]\boldsymbol{\Theta}_{\boldsymbol{\beta}}\left\{\mathbf{Z}_i - E(\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)\right\}\boldsymbol{\gamma}_0}{\partial\boldsymbol{\gamma}_0^{-\mathrm{T}}}\right.\\
&\left.n^{1/2}\{\widehat{\boldsymbol{\gamma}}^-(\boldsymbol{\beta}_0) - \boldsymbol{\gamma}_0^-\}\right)\{1+o_p(1)\}\\
=\ &\mathbf{G}_{011}^{\mathrm{T}}n^{1/2}\{\widehat{\boldsymbol{\gamma}}^-(\boldsymbol{\beta}_0) - \boldsymbol{\gamma}_0^-\}\{1+o_p(1)\},
\end{aligned}
$$

where

$$
\begin{aligned}
& \mathbf{G}_{011}^{\mathrm{T}} \\
= \quad & n^{-1} \sum_{i=1}^{n} \frac{\partial \left[g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}} \mathbf{Z}_i \boldsymbol{\gamma}_0) - E\{g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}} \mathbf{Z}_i \boldsymbol{\gamma}_0) | \boldsymbol{\beta}_0^{\mathrm{T}} \mathbf{Z}_i \boldsymbol{\gamma}_0\} \right]}{\partial \boldsymbol{\beta}_0^{\mathrm{T}} \mathbf{Z}_i \boldsymbol{\gamma}_0} \boldsymbol{\Theta}_{\boldsymbol{\beta}} \left\{ \mathbf{Z}_i - E(\mathbf{Z}_i | \boldsymbol{\beta}_0^{\mathrm{T}} \mathbf{Z}_i \boldsymbol{\gamma}_0) \right\} \boldsymbol{\gamma}_0 \\
& \times \boldsymbol{\beta}_0^{\mathrm{T}} \{\mathbf{Z}_i - E(\mathbf{Z}_i | \boldsymbol{\beta}_0^{\mathrm{T}} \mathbf{Z}_i \boldsymbol{\gamma}_0)\} \boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}.
\end{aligned}
$$

The last equality holds by using the same arguments as those lead to (S23) Now

$$
\mathbf{G}_{011}^{\mathrm{T}} = n^{-1} \sum_{i=1}^{n} \int_0^1 \mathbf{C}_{12i}(t) \mathbf{B}_r^{\mathrm{T}}(t) \boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}} dt
$$

is a $(J-1) \times (d_{\boldsymbol{\gamma}} - 1)$ matrix where

$$
\begin{aligned}
\mathbf{C}_{12i}(t) = & \frac{\partial \left[g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}} \mathbf{Z}_i \boldsymbol{\gamma}_0) - E\{g^*(Y_i, \boldsymbol{\beta}_0^{\mathrm{T}} \mathbf{Z}_i \boldsymbol{\gamma}_0) | \boldsymbol{\beta}_0^{\mathrm{T}} \mathbf{Z}_i \boldsymbol{\gamma}_0\} \right]}{\partial \boldsymbol{\beta}_0^{\mathrm{T}} \mathbf{Z}_i \boldsymbol{\gamma}_0} \boldsymbol{\Theta}_{\boldsymbol{\beta}} \left\{ \mathbf{Z}_i - E(\mathbf{Z}_i | \boldsymbol{\beta}_0^{\mathrm{T}} \mathbf{Z}_i \boldsymbol{\gamma}_0) \right\} \boldsymbol{\gamma}_0 \\
& \times [\boldsymbol{\beta}_0^{\mathrm{T}} \mathbf{X}_i(t) - E\{\boldsymbol{\beta}_0^{\mathrm{T}} \mathbf{X}_i(t) | \boldsymbol{\beta}_0^{\mathrm{T}} \mathbf{Z}_i \boldsymbol{\gamma}_0\}].
\end{aligned}
$$

Because J is finite, $\|\mathbf{G}_{011}\|_\infty$ and $\|\mathbf{G}_{011l}\|_\infty$ have the same order as their elements did, where $\mathbf{G}_{011l}$ is the lth column of $\mathbf{G}_{011}$. The (k,l) element of $\mathbf{G}_{011}$ is

$$
G_{011kl} = n^{-1} \sum_{i=1}^{n} \int_0^1 \{B_{rk}(t)\} C_{12li} dt,
$$

where C_{12li} is the lth element of $\mathbf{C}_{12i}$. Now $E\{C_{12li}(t)\} \neq 0$, $|E\{C_{12li}(t)\}| < \infty$, $\|C_{12li}(\cdot)\|_2 < \infty$, *a.s.* by Condition (A6). Therefore, by Lemma 3 we have $|G_{011kl}| = O_p(h_b)$, and thus $\mathbf{G}_{011l}$ and $\mathbf{G}_{011}$ satisfy

$$
\|\mathbf{G}_{011l}\|_\infty = O_p(h_b), \ \|\mathbf{G}_{011}\|_\infty = O_p(h_b). \tag{S34}
$$

Since $\|\mathbf{a}\|_2 \leq N^{1/2} \|\mathbf{a}\|_\infty$ for arbitrary vector $\mathbf{a}$ with length N, and $\|\mathbf{b}\|_2 \leq N^{1/2} \|\mathbf{b}\|_\infty$ for arbitrary matrix $\mathbf{b}$ with N rows, we obtain

$$
\begin{aligned}
\|\mathbf{G}_{011l}\|_2 \leq N^{1/2} O_p(h_b) &= O_p(h_b^{1/2}), \\
\|\mathbf{G}_{011}\|_2 \leq N^{1/2} O_p(h_b) &= O_p(h_b^{1/2}).
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
& \mathbf{G}_{011}^{\mathrm{T}} n^{1/2}\{\widehat{\boldsymbol{\gamma}}^{-}(\boldsymbol{\beta}_0)-\boldsymbol{\gamma}_0^{-}\} \\
= \; & O_p[\max_{1\le l\le J} \mathbf{G}_{011l}^{\mathrm{T}} n^{1/2}\{\widehat{\boldsymbol{\gamma}}^{-}(\boldsymbol{\beta}_0)-\boldsymbol{\gamma}_0^{-}\}] \\
= \; & h_b^{1/2} O_p[\max_{1\le l\le J} h_b^{-1/2}\mathbf{G}_{011l}^{\mathrm{T}} n^{1/2}\{\widehat{\boldsymbol{\gamma}}^{-}(\boldsymbol{\beta}_0)-\boldsymbol{\gamma}_0^{-}\}] \\
= \; & O_p(1)
\end{aligned}
$$

by Theorem 1 that $|\mathbf{a}^{\mathrm{T}}\{\widehat{\boldsymbol{\gamma}}^{-}(\boldsymbol{\beta}_0)-\boldsymbol{\gamma}_0^{-}\}| = O_p(n^{-1/2}h_b^{-1/2})$ for arbitrary $\mathbf{a}$ with $\|\mathbf{a}\|_2<\infty$.

Further,

$$
\begin{aligned}
\mathbf{G}_{011}^{\mathrm{T}} n^{1/2}\{\widehat{\boldsymbol{\gamma}}^{-}(\boldsymbol{\beta}_0)-\boldsymbol{\gamma}_0^{-}\} \;=\; & \mathbf{G}_{013}^{\mathrm{T}} n^{1/2}\{\widehat{\boldsymbol{\gamma}}^{-}(\boldsymbol{\beta}_0)-\boldsymbol{\gamma}_0^{-}\} + (\mathbf{G}_{011}-\mathbf{G}_{012})^{\mathrm{T}} n^{1/2}\{\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0^{-})-\boldsymbol{\gamma}_0^{-}\} \\
& +(\mathbf{G}_{012}-\mathbf{G}_{013})^{\mathrm{T}} n^{1/2}\{\widehat{\boldsymbol{\gamma}}^{-}(\boldsymbol{\beta}_0)-\boldsymbol{\gamma}_0^{-}\}
\end{aligned}
$$

where

$$
\begin{aligned}
& \mathbf{G}_{012}^{\mathrm{T}} \\
= \; & E\Bigg(\frac{\partial\left[g^*(Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)-E\{g^*(Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}\right]}{\partial\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0}\boldsymbol{\Theta}_{\boldsymbol{\beta}}\left\{\mathbf{Z}_i-E(\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)\right\}\boldsymbol{\gamma}_0 \\
& \times\boldsymbol{\beta}_0^{\mathrm{T}}\{\mathbf{Z}_i-E(\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)\}\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\Bigg)
\end{aligned}
$$

and

$$
\begin{aligned}
\mathbf{G}_{013}^{\mathrm{T}} \;=\; & E\Bigg\{\frac{\partial\left(g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}-E[g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)]\right)}{\partial\{\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}}\boldsymbol{\Theta}_{\boldsymbol{\beta}} \\
& \times\left[\boldsymbol{\alpha}_0(\mathbf{X}_i)-E\{\boldsymbol{\alpha}_0(\mathbf{X}_i)|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\right]\boldsymbol{\beta}_0^{\mathrm{T}}[\mathbf{Z}_i-E\{\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}]\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\Bigg\}.
\end{aligned}
$$

Therefore,

$$
\mathbf{G}_{011}^{\mathrm{T}}-\mathbf{G}_{012}^{\mathrm{T}} = n^{-1}\sum_{i=1}^{n}\int_0^1 \mathbf{C}_{13i}(t)\mathbf{B}_r^{\mathrm{T}}(t)\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}dt,
$$

where

$$
\begin{aligned}
& \mathbf{C}_{13i} \\
= \; & \frac{\partial\left[g^*(Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)-E\{g^*(Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}\right]}{\partial\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0}\boldsymbol{\Theta}_{\boldsymbol{\beta}}
\end{aligned}
$$

$$\times \left\{\mathbf{Z}_i - E(\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)\right\}\boldsymbol{\gamma}_0[\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i(t) - E\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i(t)|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}]$$
$$-E\Bigg(\frac{\partial\left[g^*(Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0) - E\{g^*(Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}\right]}{\partial\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0}\boldsymbol{\Theta}_{\boldsymbol{\beta}}$$
$$\times \left\{\mathbf{Z}_i - E(\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0)\right\}\boldsymbol{\gamma}_0[\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i(t) - E\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i(t)|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\boldsymbol{\gamma}_0\}]\Bigg),$$

with its lth element $C_{13l}(t)$ satisfies $E\{C_{13li}(t)\} = 0$, $|E\{C_{13li}(t)\}| < \infty$, $\|C_{13li}(\cdot)\|_2 < \infty$, *a.s.* by Condition (A6). Therefore, using the similar argument as those led to (S34) and Lemma 3 we have

$$\|\mathbf{G}_{011} - \mathbf{G}_{012}\|_\infty = O_p\{\sqrt{h_b n^{-1}\log(n)}\} = o_p(h_b).$$

Further, $\|\mathbf{G}_{012} - \mathbf{G}_{013}\|_\infty = O_p(h_b^q)$ by Condition (A5). As a result, we have

$$\|\mathbf{G}_{012} - \mathbf{G}_{013}\|_\infty = o_p(\|\mathbf{G}_{011}\|_\infty) = o_p(h_b) \tag{S35}$$

and

$$\|\mathbf{G}_{013}\|_2 \le \sqrt{N}\|\mathbf{G}_{013}\|_\infty = \sqrt{N}O_p(\|\mathbf{G}_{011}\|_\infty) = O_p(h_b^{1/2}) \tag{S36}$$

by (S34) and we can write

$$\mathbf{G}_{01} \;=\; \mathbf{G}_{013}^{\mathrm{T}} n^{1/2}\{\widehat{\boldsymbol{\gamma}}^-(\boldsymbol{\beta}_0) - \boldsymbol{\gamma}_0^-\} + o_p(1).$$

Combining with the result that $\|\mathbf{G}_{02}\| = O_p(h_b^q)$ and the fact that $\mathbf{G}_0 = \mathbf{G}_{00} + \mathbf{G}_{01} + \mathbf{G}_{02}$, we have

$$\begin{aligned}\mathbf{G}_0 &= \mathbf{G}_{00} + \mathbf{G}_{013}^{\mathrm{T}} n^{1/2}\{\widehat{\boldsymbol{\gamma}}^-(\boldsymbol{\beta}_0) - \boldsymbol{\gamma}_0^-\} + o_p(1)\\ &= O_p(1).\end{aligned}$$

Further, $\mathbf{G}_1 = O_p(h^2)$, $\mathbf{G}_2 = O_p(h^2)$ and $\mathbf{G}_3 = o_p(1)$, $\mathbf{G} = \mathbf{G}_0 + \mathbf{G}_1 + \mathbf{G}_2 + \mathbf{G}_3$, so we have

$$\begin{aligned}\mathbf{G} &= \mathbf{G}_{00} + \mathbf{G}_{013}^{\mathrm{T}} n^{1/2}\{\widehat{\boldsymbol{\gamma}}^-(\boldsymbol{\beta}_0) - \boldsymbol{\gamma}_0^-\} + o_p(1)\\ &= O_p(1). \qquad \text{(S37)}\end{aligned}$$

The Asymptotic Property for H Note that $\mathbf{H}$ is a finite dimensional vector, and

$$
\begin{aligned}
& \|\mathbf{B}_r(\cdot)\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0) - \alpha_0(\cdot)\|_\infty \\
\leq \;& \|\mathbf{B}_r(\cdot)\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0) - \mathbf{B}_r(\cdot)\boldsymbol{\gamma}_0\|_\infty \\
& + \|\mathbf{B}_r(\cdot)\boldsymbol{\gamma}_0 - \alpha_0(\cdot)\|_\infty \\
= \;& O_p(n^{-1/2}h^{-1/2} + h_b^q) \\
= \;& o_p(1).
\end{aligned}
$$

Now by the fact that the kernel estimators are uniformly consistent, we have

$$
\mathbf{H} = \mathbf{H}_0 + \mathbf{H}_1 + o_p(1),
$$

where

$$
\begin{aligned}
\mathbf{H}_0 \;=\;& E\{\partial\left(g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\} - E[g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)]\right)\boldsymbol{\Theta}_{\boldsymbol{\beta}} \\
& \times\left[\boldsymbol{\alpha}_0(\mathbf{X}_i) - E\{\boldsymbol{\alpha}_0(\mathbf{X}_i)|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\right]/\partial\boldsymbol{\beta}_0^{\mathrm{T}}\}\boldsymbol{\Theta}_{\boldsymbol{\beta}}^{\mathrm{T}}, \qquad \text{(S38)}
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{H}_1 \;=\;& n^{-1}\sum_{i=1}^{n}[\partial\{\left(g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\gamma}(\boldsymbol{\beta}_0)\} - E[g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\gamma}(\boldsymbol{\beta}_0)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\gamma}(\boldsymbol{\beta}_0)]\right)\boldsymbol{\Theta}_{\boldsymbol{\beta}} \\
& \times\left[\mathbf{Z}_i - E\{\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{Z}_i\widehat{\gamma}(\boldsymbol{\beta}_0)\}\right]\widehat{\gamma}(\boldsymbol{\beta}_0)\}/\partial\widehat{\gamma}^{\mathrm{T}}(\boldsymbol{\beta}_0)]\partial\widehat{\gamma}(\boldsymbol{\beta})/\partial\boldsymbol{\beta}^{\mathrm{T}}|_{\boldsymbol{\beta}=\boldsymbol{\beta}^*} \\
=\;& \mathbf{G}_{011}^{\mathrm{T}}\partial\widehat{\boldsymbol{\gamma}}^-(\boldsymbol{\beta})/\partial\boldsymbol{\beta}^{\mathrm{T}}|_{\boldsymbol{\beta}=\boldsymbol{\beta}^*}\{1+o_p(1)\} \\
=\;& \mathbf{G}_{013}^{\mathrm{T}}\partial\widehat{\boldsymbol{\gamma}}^-(\boldsymbol{\beta})/\partial\boldsymbol{\beta}^{\mathrm{T}}|_{\boldsymbol{\beta}=\boldsymbol{\beta}^*}\{1+o_p(1)\} \\
=\;& \mathbf{G}_{013}^{\mathrm{T}}\mathbf{P}\{1+o_p(1)\} \\
=\;& \mathbf{G}_{013}^{\mathrm{T}}\mathbf{P} + o_p(1),
\end{aligned} \qquad \text{(S39)}
$$

where $\mathbf{P}$ is defined in (S53). The second equality holds because $\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)$ is consistent to $\boldsymbol{\gamma}_0$. The third equality holds by (S35). The fourth equality holds by (S52). The last equality holds because $\|\mathbf{G}_{013}\|_2 = O_p(h_b^{1/2})$ by (S36), and $\|\mathbf{P}\|_2 = O_p(h_b^{-1/2})$ by (S54). Combining (S37), (S38), (S39) and (S32), we have

$$
\sqrt{n}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0) \qquad \text{(S40)}
$$

$$
\begin{aligned}
&= -\mathbf{H}^{-1}\mathbf{G} \\
&= -\{\mathbf{H}_0 + \mathbf{H}_1 + o_p(1)\}^{-1}\{\mathbf{G}_{00} + \mathbf{G}_{013}^{\mathrm{T}} n^{1/2}\{\widehat{\boldsymbol{\gamma}}^{-}(\boldsymbol{\beta}_0) - \boldsymbol{\gamma}_0^{-}\} + o_p(1)\} \\
&= -\{\mathbf{H}_0 + \mathbf{G}_{013}^{\mathrm{T}}\mathbf{P} + o_p(1)\}^{-1}\{\mathbf{G}_{00} + \mathbf{G}_{013}^{\mathrm{T}}\mathbf{L} + o_p(1)\} \\
&= O_p(1)
\end{aligned}
$$

where $\mathbf{L}$ is defined in Theorem 1.

Part II

Now note that we can write $\mathbf{P} = (\mathbf{P}_1, \ldots, \mathbf{P}_{J-1})$, where $\mathbf{P}_l$ is the lth column of $\mathbf{P}$, $l = 1, \ldots, J-1$, which can be written as .

$$
\begin{aligned}
\mathbf{P}_l = \; & \mathrm{argmin}_{\widetilde{\mathbf{P}}_l} E\Bigg\{\Bigg(\int_0^1 X_{il}(t)\alpha_0(t)dt + \Bigg[\int_0^1 \mathbf{B}_r(t)^{\mathrm{T}}\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\widetilde{\mathbf{P}}_l\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i(t)dt \\
& -E\left\{\int_0^1 \mathbf{B}_r(t)^{\mathrm{T}}\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\widetilde{\mathbf{P}}_l\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i(t)dt|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\right\}\Bigg]\Bigg)^2 \\
& \times\partial[g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\} - E\{g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}]/\partial\{\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\Bigg\}.
\end{aligned}
$$

Hence,

$$
\begin{aligned}
& \mathbf{B}_r(\cdot)^{\mathrm{T}}\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\mathbf{P}_l \\
= \; & \mathrm{argmin}_{\mathbf{B}_r(\cdot)^{\mathrm{T}}\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\mathbf{P}_l} E\Bigg\{\Bigg(\int_0^1 X_{il}(t)\alpha_0(t)dt + \Bigg[\int_0^1 \mathbf{B}_r(t)^{\mathrm{T}}\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\widetilde{\mathbf{P}}_l\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i(t)dt \\
& -E\left\{\int_0^1 \mathbf{B}_r(t)^{\mathrm{T}}\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\widetilde{\mathbf{P}}_l\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i(t)dt|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\right\}\Bigg]\Bigg)^2 \\
& \times\partial[g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\} - E\{g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}]/\partial\{\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\Bigg\}.
\end{aligned}
$$

Now let $\delta_l \in C^q[0,1]$, $\delta_l(0) = 1$ and

$$
\begin{aligned}
& \delta_l(\cdot) \\
= \; & \mathrm{argmin}_{\widetilde{\delta}_l} E\Bigg\{\Bigg(\int_0^1 X_{il}(t)\alpha_0(t)dt + \Bigg[\int_0^1 \boldsymbol{\beta}_0^{\mathrm{T}}\widetilde{\delta}_l(t)\mathbf{X}_{ic}(t)dt\Bigg]\Bigg)^2 \\
& \times\Delta g_c^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\Bigg\}
\end{aligned}
\qquad \text{(S41)}
$$

where recalling that $\Delta g_c^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}$ is defined before Theorem 2, and $\mathbf{X}_{ic}(s) = \mathbf{X}_i(s) - E\{\mathbf{X}_i(s)|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}$. We further define $\boldsymbol{\alpha}_{c0}(\mathbf{X}_i) = \boldsymbol{\alpha}_0(\mathbf{X}_i) -$

$E\{\boldsymbol{\alpha}_0(\mathbf{X}_i)|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}$. As the number of spline knots goes to infinity, $\|\mathbf{B}_r(\cdot)^{\mathrm{T}}\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\mathbf{P}_l - \delta_l(\cdot)\|_\infty = O_p(h_b^q)$. Hence

$$\begin{aligned} & \left\| \mathbf{G}_{013}^{\mathrm{T}}\mathbf{P}_l - E\left[\Delta g_c^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\boldsymbol{\Theta}_{\boldsymbol{\beta}}\boldsymbol{\alpha}_{c0}(\mathbf{X}_i)\int_0^1 \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_{ic}(t)\delta_l(t)dt\right] \right\|_2 \\ = \;& o_p(1). \end{aligned} \tag{S42}$$

Further, the Gâteaux derivative of the target function in (S41) at $\widetilde{\delta}_l = \delta_l$ in the direction $\mathbf{w}_0^*(t)$ satisfies

$$\begin{aligned} \mathbf{0} &= \partial E\left(\int_0^1 X_{il}(t)\alpha_0(t)dt + \left[\int_0^1 \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_{ic}(t)\{\delta_l(t) + \mathbf{v}^{\mathrm{T}}\mathbf{w}_0^*(t)\}dt\right]\right)^2 \\ &\quad \times \Delta g_c^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}/\partial\mathbf{v}|_{\mathbf{v}=\mathbf{0}} \\ &= 2E\left[\int_0^1 X_{il}(t)\alpha_0(t)dt\Delta g_c^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\int_0^1 \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_{ic}(t)\mathbf{w}_0^*(t)dt\right] \\ &\quad +2\int_0^1 \delta_l(t)\boldsymbol{\Theta}_{\boldsymbol{\beta}}E[\boldsymbol{\alpha}_{c0}(\mathbf{X}_i)\Delta g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\mathbf{X}_{ic}^{\mathrm{T}}(t)]\boldsymbol{\beta}_0 dt \\ &= 2E\left[\int_0^1 X_{icl}(t)\alpha_0(t)dt\Delta g_c^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\int_0^1 \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_{ic}(t)\mathbf{w}_0^*(t)dt\right] \\ &\quad +2\int_0^1 \delta_l(t)\boldsymbol{\Theta}_{\boldsymbol{\beta}}E[\boldsymbol{\alpha}_{c0}(\mathbf{X}_i)\Delta g^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\mathbf{X}_{ic}^{\mathrm{T}}(t)]\boldsymbol{\beta}_0 dt, \end{aligned}$$

where $\mathbf{w}_0^*(t)$ is defined in (7), X_{icl} is the lth element of $\mathbf{X}_{ic}$. The second equality holds by the definition of $\mathbf{w}_0^*(t)$ in (7). The last equality holds because

$$\begin{aligned} & E\left[\int_0^1 E\{X_{il}(t)|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0\}\alpha_0(t)dt\Delta g_c^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\int_0^1 \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_{ic}(t)\mathbf{w}_0^*(t)dt\right] \\ = \;& \mathbf{0}. \end{aligned}$$

This implies

$$\begin{aligned} & E\left[\boldsymbol{\Theta}_{\boldsymbol{\beta}}\boldsymbol{\alpha}_{c0}(\mathbf{X}_i)\Delta g_c^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\int_0^1 \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_{ic}(t)\delta_l(t)dt\right] \\ = \;& -E\left[\int_0^1 X_{icl}(t)\alpha_0(t)dt\Delta g_c^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\int_0^1 \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_{ic}(t)\mathbf{w}_0^*(t)dt\right]. \end{aligned}$$

This holds for each $l = 1, \ldots, J-1$. Combine this result with (S42), we have

$$\left\| \mathbf{G}_{013}^{\mathrm{T}}\mathbf{P} + E\left[\Delta g_c^*\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\int_0^1 \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_{ic}(t)\mathbf{w}_0^*(t)dt\boldsymbol{\alpha}_{c0}(\mathbf{X}_i)^{\mathrm{T}}\boldsymbol{\Theta}_{\boldsymbol{\beta}}^{\mathrm{T}}\right] \right\|_2$$

$$= \quad o_p(1).$$

Next, by (S29) and (S30), where $\widehat{\mathbf{C}}_{\mathbf{Q}}(\boldsymbol{\beta}_0)$ and $\mathbf{C}_{\mathbf{Q}}(\boldsymbol{\beta}_0)$ in (S29) and (S30) are defined in (S25), we can write

$$n^{-1/2}\mathbf{L} \quad = \quad \widehat{\mathbf{L}}\{1+o_p(1)\}, \tag{S43}$$

where

$$\begin{aligned}\widehat{\mathbf{L}} \quad = \quad & \mathrm{argmin}_{\widetilde{\mathbf{L}}} n^{-1}\sum_{i=1}^{n}\Bigg([g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\} - E\{g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}] \\ & \times\Delta g_c^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}^{-1} + \left[\int_0^1 \mathbf{B}_r(t)^{\mathrm{T}}\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\widetilde{\mathbf{L}}\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_{ic}(t)dt\right]\Bigg)^2\Delta g_c^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}.\end{aligned}$$

Further, by above equality we have

$$\begin{aligned}& \mathbf{B}_r(\cdot)^{\mathrm{T}}\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\widehat{\mathbf{L}} \\ = \quad & \mathrm{argmin}_{\mathbf{B}_r(\cdot)^{\mathrm{T}}\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\widetilde{\mathbf{L}}} n^{-1}\sum_{i=1}^{n}\Bigg([g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\} - E\{g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}] \\ & \Delta g_c^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}^{-1} + \left[\int_0^1 \mathbf{B}_r(t)^{\mathrm{T}}\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\widetilde{\mathbf{L}}\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_{ic}(t)dt\right]\Bigg)^2\Delta g_c^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}.\end{aligned}$$

Let

$$\begin{aligned}& \widehat{\eta}(\cdot) \\ = \quad & \mathrm{argmin}_{\widetilde{\eta}(\cdot)} n^{-1}\sum_{i=1}^{n}\Bigg\{\Bigg([g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\} \\ & -E\{g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}]\Delta g_c^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}^{-1} \\ & +\left[\int_0^1 \widetilde{\eta}(t)\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_{ic}(t)dt\right]\Bigg)^2\Delta g_c^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\Bigg\}.\end{aligned} \tag{S44}$$

As the number of spline knots goes to infinity, $\|\mathbf{B}_r(\cdot)^{\mathrm{T}}\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\widehat{\mathbf{L}}(\cdot) - \widehat{\eta}(\cdot)\|_\infty = O_p(h_b^q)$. Combining with (S43) we have

$$\begin{aligned}& \Bigg\|\mathbf{G}_{013}^{\mathrm{T}}\mathbf{L} - \int_0^1 E\Bigg[\boldsymbol{\Theta}_{\boldsymbol{\beta}}\boldsymbol{\alpha}_c(\mathbf{X}_i) \\ & \Delta g_c^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_{ic}(t)\Bigg]n^{1/2}\widehat{\eta}(t)dt\Bigg\|_\infty \\ = \quad & o_p(1).\end{aligned} \tag{S45}$$

Now taking the Gâteaux derivative of the target function in (S44) at $\widehat{\eta}$ in the direction of $\mathbf{w}_0^*(t)$, we have

$$
\begin{aligned}
\mathbf{0} &= \partial n^{-1}\sum_{i=1}^{n}\Bigg[\Bigg\{(g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\} - E[g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)])\\
&\quad \Delta g_c^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}^{-1} + \int_0^1\{\widehat{\eta}(t)+\mathbf{v}^{\mathrm{T}}\mathbf{w}_0^*(t)\}\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_{ic}(t)dt\Bigg\}^2\\
&\quad \times\Delta g_c^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\Bigg]/\partial\mathbf{v}|_{\mathbf{v}=\mathbf{0}}\\
&= 2n^{-1}\sum_{i=1}^{n}\Bigg(\Bigg[(g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\} - E[g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)])\\
&\quad \times\Delta g_c^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}^{-1} + \int_0^1\widehat{\eta}(t)\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_{ic}(t)dt\Bigg]\\
&\quad \times\Delta g_c^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\int_0^1\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_{ic}(t)\mathbf{w}_0^*(t)dt\Bigg)\\
&= 2n^{-1}\sum_{i=1}^{n}\Bigg\{(g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\} - E[g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)])\\
&\quad \times\int_0^1\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_{ic}(t)\mathbf{w}_0^*(t)dt\Bigg\} + 2\int_0^1\widehat{\eta}(t)E\Bigg[\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_{ic}(t)dt\\
&\quad \times\Delta g_c^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\int_0^1\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_{ic}(t)\mathbf{w}_0^*(t)dt\Bigg] + o_p(n^{-1/2})\\
&= 2n^{-1}\sum_{i=1}^{n}\Bigg\{(g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\} - E[g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)])\\
&\quad \times\int_0^1\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_{ic}(t)\mathbf{w}_0^*(t)dt\Bigg\} + 2\Bigg(\int_0^1\widehat{\eta}(t)\boldsymbol{\Theta}_{\boldsymbol{\beta}}E[\boldsymbol{\alpha}_{c0}(\mathbf{X}_i)\\
&\quad \times\Delta g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\mathbf{X}_{ic}^{\mathrm{T}}(t)]\boldsymbol{\beta}_0 dt\Bigg) + o_p(n^{-1/2}).
\end{aligned}
$$

The last equality holds by the definition of $\mathbf{w}_0^*(t)$. This implies

$$
\begin{aligned}
& n^{-1}\sum_{i=1}^{n}\Bigg([g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\} - E\{g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}]\\
&\quad \times\int_0^1\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_{ic}(t)\mathbf{w}_0^*(t)dt\Bigg)\\
&= -\int_0^1\widehat{\eta}(t)\boldsymbol{\Theta}_{\boldsymbol{\beta}}E[\boldsymbol{\alpha}_{c0}(\mathbf{X}_i)\Delta g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\mathbf{X}_{ic}^{\mathrm{T}}(t)]\boldsymbol{\beta}_0 dt\\
&\quad + o_p(n^{-1/2})
\end{aligned}
$$

Combining with (S45), we have

$$\begin{aligned}
&\Bigg\|\mathbf{G}_{013}^{\mathrm{T}}\mathbf{L}+n^{-1/2}\sum_{i=1}^{n}\Bigg([g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}-E\{g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}]\\
&\times\int_0^1\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_{ic}(t)\mathbf{w}_0^*(t)dt\Bigg)\Bigg\|_2\\
=\;& o_p(1).
\end{aligned}$$

As a result, by (S40) the population asymptotic forms of $\sqrt{n}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)$ can be written as

$$\begin{aligned}
&\sqrt{n}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)\\
=\;& -\{\mathbf{H}_0+\mathbf{G}_{013}^{\mathrm{T}}\mathbf{P}+o_p(1)\}^{-1}\{\mathbf{G}_{00}+\mathbf{G}_{013}^{\mathrm{T}}\mathbf{L}+o_p(1)\}\\
=\;& -\Bigg\{\mathbf{H}_0-E\Bigg[\Delta g_c^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\int_0^1\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_{ic}(t)\mathbf{w}_0^*(t)dt\boldsymbol{\alpha}_{c0}(\mathbf{X}_i)^{\mathrm{T}}\boldsymbol{\Theta}_{\boldsymbol{\beta}}^{\mathrm{T}}\Bigg]\\
&+o_p(1)\Bigg\}^{-1}\Bigg[\mathbf{G}_{00}-n^{-1/2}\sum_{i=1}^{n}\Bigg\{(g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\\
&-E[g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)])\int_0^1\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_{ic}(t)\mathbf{w}_0^*(t)dt\Bigg\}+o_p(1)\Bigg]\\
=\;& \mathbf{A}^{-1}\mathbf{B}+o_p(1),
\end{aligned}$$

where

$$\begin{aligned}
\mathbf{A} =\;& -\Bigg(E[\Delta g_c^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\{\boldsymbol{\Theta}_{\boldsymbol{\beta}}\boldsymbol{\alpha}_{c0}(\mathbf{X}_i)\}^{\otimes 2}]\\
&-E\left[\Delta g_c^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\int_0^1\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_{ic}(t)\mathbf{w}_0^*(t)dt\boldsymbol{\alpha}_{c0}(\mathbf{X}_i)^{\mathrm{T}}\boldsymbol{\Theta}_{\boldsymbol{\beta}}^{\mathrm{T}}\right]\Bigg)
\end{aligned}$$

and

$$\begin{aligned}
\mathbf{B} =\;& n^{-1/2}[g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}-E\{g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}]\\
&\times\left\{\boldsymbol{\Theta}_{\boldsymbol{\beta}}\boldsymbol{\alpha}_{c0}(\mathbf{X}_i)-\int_0^1\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_{ic}(t)\mathbf{w}_0^*(t)dt\right\}.
\end{aligned}$$

If $g^*=g$, then $E\{g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}]=0$,
$E[\Delta g_c^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}(\mathbf{X}_i)]=-E[g^{*2}\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}(\mathbf{X}_i)]$ and $\mathbf{w}_0^*(t)=\mathbf{w}_0(t)$ as defined in (S48). Hence we have

$$\sqrt{n}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})=\mathbf{A}_0^{-1}\mathbf{B}_0+o_p(1)$$

where

$$
\begin{aligned}
\mathbf{A}_0 &= E[g^2\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\{\boldsymbol{\Theta}_{\beta}\boldsymbol{\alpha}_{c0}(\mathbf{X}_i)\}^{\otimes 2}] \\
&\quad - E\left[g^2\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\int_0^1 \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_{ic}(t)\mathbf{w}_0(t)dt\boldsymbol{\alpha}_{c0}(\mathbf{X}_i)^{\mathrm{T}}\boldsymbol{\Theta}_{\boldsymbol{\beta}}^{\mathrm{T}}\right]
\end{aligned}
$$

and

$$
\begin{aligned}
\mathbf{B}_0 &= n^{-1/2}[g\{Y_i, \boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}]\left\{\boldsymbol{\Theta}_{\beta}\boldsymbol{\alpha}_{c0}(\mathbf{X}_i) - \int_0^1 \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_{ic}(t)\mathbf{w}_0(t)dt\right\} \\
&= \mathbf{S}_{o\mathrm{eff}},
\end{aligned}
$$

where $\mathbf{S}_{o\mathrm{eff}}$ is defined in (S47). With simple calculation, we can see that $\mathbf{A}_0 = \mathrm{var}(\mathbf{B}_0)$. Hence, when g^* is correctly specified, $\widehat{\boldsymbol{\beta}}$ is the efficient estimator. This proves the result. □

S3.3 Proof of Theorem 3

Proof:

$$
\begin{aligned}
& \sup_{t\in[0,1]} |\mathbf{B}_r(t)^{\mathrm{T}}\widehat{\boldsymbol{\gamma}}(\widehat{\boldsymbol{\beta}}) - \alpha_0(t)| \\
= & \sup_{t\in[0,1]} |\mathbf{B}_r(t)^{\mathrm{T}}\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\widehat{\boldsymbol{\gamma}}^{-}(\widehat{\boldsymbol{\beta}}) - \alpha_0(t)| \\
\le & \sup_{t\in[0,1]} |\mathbf{B}_r(t)^{\mathrm{T}}\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\widehat{\boldsymbol{\gamma}}^{-}(\widehat{\boldsymbol{\beta}}) - \mathbf{B}_r(t)^{\mathrm{T}}\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\widehat{\boldsymbol{\gamma}}^{-}(\boldsymbol{\beta}_0)| + \sup_{t\in[0,1]} |\mathbf{B}_r(t)^{\mathrm{T}}\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\widehat{\boldsymbol{\gamma}}^{-}(\boldsymbol{\beta}_0) - \mathbf{B}_r(t)^{\mathrm{T}}\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\boldsymbol{\gamma}_0^{-}| \\
& + \sup_{t\in[0,1]} |\mathbf{B}_r(t)^{\mathrm{T}}\boldsymbol{\gamma}_0 - \alpha_0(t)| \\
\le & \|\widehat{\boldsymbol{\gamma}}^{-}(\boldsymbol{\beta})/\partial\boldsymbol{\beta}^{\mathrm{T}}|_{\boldsymbol{\beta}=\boldsymbol{\beta}^*}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)\|_\infty + \|\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0) - \boldsymbol{\gamma}_0\|_\infty + O_p(h_b^q) \\
= & O_p(n^{-1/2}h_b^{-1/2} + h_b^q) + \|\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0) - \boldsymbol{\gamma}_0\|_\infty \\
= & O_p(n^{-1/2}h_b^{-1/2} + h_b^q) \\
= & O_p(n^{-1/2}h_b^{-1/2}). \qquad \text{(S46)}
\end{aligned}
$$

The third line holds by Lemma 1 and Condition (A5). The fourth line holds because

$$
\widehat{\boldsymbol{\gamma}}^{-}(\boldsymbol{\beta})/\partial\boldsymbol{\beta}^{\mathrm{T}}|_{\boldsymbol{\beta}=\boldsymbol{\beta}^*}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0) = \mathbf{P}_1 + o_p(\mathbf{P}_1)
$$

by Lemma 4, where

$$\begin{aligned}
\mathbf{P}_1 &= -\Bigg(\Bigg\{E\Bigg(\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\left[\mathbf{Z}_i - E\{\mathbf{Z}_i|\boldsymbol{\beta}_0\boldsymbol{\alpha}_0(\mathbf{X}_i\}\right]^{\mathrm{T}}\boldsymbol{\beta}_0 \\
&\quad\times\frac{\partial\left(g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\} - E[g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)]\right]\big)}{\partial\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)} \\
&\quad\times\boldsymbol{\beta}_0^{\mathrm{T}}[\mathbf{Z}_i - E\{\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}]\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\Bigg)\Bigg\}^{-1}\Bigg)\mathbf{M}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0).
\end{aligned}$$

where

$$\begin{aligned}
\mathbf{M} &= E(\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\left[\mathbf{Z}_i - E\{\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\right]^{\mathrm{T}}\boldsymbol{\beta}_0\boldsymbol{\alpha}_0(\mathbf{X}_i)^{\mathrm{T}}\boldsymbol{\Theta}_{\boldsymbol{\beta}}^{\mathrm{T}} \\
&\quad\times\partial[g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\} - E\{g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}]/\partial\{\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\})
\end{aligned}$$

is a $d_{\boldsymbol{\gamma}-1} \times d_{\boldsymbol{\beta}}$ dimensional matrix. Now note that $\|\mathbf{P}_1\|_\infty \le \|\mathbf{P}_1\|_2$ by the fact that the vector sup norm is less than its L_2 norm, we obtain

$$\begin{aligned}
&\|\mathbf{P}_1\|_\infty \\
\le\ & \Bigg\|\Bigg(\Bigg\{E\Bigg(\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\left[\mathbf{Z}_i - E\{\mathbf{Z}_i|\boldsymbol{\beta}_0\boldsymbol{\alpha}_0(\mathbf{X}_i\}\right]^{\mathrm{T}}\boldsymbol{\beta}_0 \\
&\times\frac{\partial\left(g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\} - E[g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)]\right]\big)}{\partial\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)} \\
&\times\boldsymbol{\beta}_0^{\mathrm{T}}[\mathbf{Z}_i - E\{\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}]\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\Bigg)\Bigg\}^{-1}\Bigg)\Bigg\|_2\|\mathbf{M}\|_2\|(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)\|_2 \\
=\ & O_p(h_b^{-1})O_p(h_b^{1/2})O_p(n^{-1/2}) \\
=\ & O_p(n^{-1/2}h_b^{-1/2}).
\end{aligned}$$

The second to the last equality holds because of (S55) and the root n convergence of $\widehat{\boldsymbol{\beta}}$. In addition, $\|\mathbf{M}\|_2 = O_p(h_b^{1/2})$ by (S56).

The fifth equality in (S46) holds because by choosing $\mathbf{a}$ to be a unit vector with 1 at one entry in Theorem 1, we can obtain that each entry of $\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0) - \boldsymbol{\gamma}_0$ is of order $O_p(n^{-1/2}h_b^{-1/2})$, and thus $\|\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0) - \boldsymbol{\gamma}_0\|_\infty = O_p(n^{-1/2}h_b^{-1/2})$. The last equality in (S46) holds by Condition (A4). $\square$

S4 Necessary propositions and lemmas

Proposition S4.1. *The efficient score function of model (2) is* $\mathbf{S}_{\text{eff}} = (\mathbf{S}_{\text{eff}\beta}^{\mathrm{T}}, \mathbf{S}_{\text{eff}\gamma}^{\mathrm{T}})^{\mathrm{T}}$, *where*

$$\begin{aligned}
\mathbf{S}_{\text{eff}\beta} &= \frac{f_2'(Y, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma})}{f(Y, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma})}(\mathbf{I}_{J-1}, \mathbf{0})\{\mathbf{Z} - E(\mathbf{Z} \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma})\}\boldsymbol{\gamma}, \\
\mathbf{S}_{\text{eff}\gamma} &= \frac{f_2'(Y, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma})}{f(Y, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma})}(\mathbf{0}, \mathbf{I}_{d_\gamma-1})\{\mathbf{Z} - E(\mathbf{Z} \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma})\}^{\mathrm{T}}\boldsymbol{\beta}.
\end{aligned}$$

Here f_2' *stands for the derivative of* f *with respect to the second variable.*

Proof: We first calculate the score vector $\mathbf{S} = (\mathbf{S}_\beta^{\mathrm{T}}, \mathbf{S}_\gamma^{\mathrm{T}})^{\mathrm{T}}$. Taking derivative of the log likelihood of one observation with respect to the parameters in $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, we obtain

$$\begin{aligned}
\mathbf{S}_\beta &= \frac{f_2'(Y, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma})}{f(Y, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma})}(\mathbf{I}_{J-1}, \mathbf{0})\mathbf{Z}\boldsymbol{\gamma}, \\
\mathbf{S}_\gamma &= \frac{f_2'(Y, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma})}{f(Y, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma})}(\mathbf{0}, \mathbf{I}_{d_\gamma-1})\mathbf{Z}^{\mathrm{T}}\boldsymbol{\beta}.
\end{aligned}$$

We first show that $\mathbf{S}_{\text{eff}} \in \Lambda^\perp$. This is verified by noting that

$$\mathbf{S}_{\text{eff}} = \mathbf{S} - E(\mathbf{S} \mid Y, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma}),$$

and $\mathbf{S}$ satisfies $E(\mathbf{S} \mid \mathbf{Z}) = \mathbf{0} = E(\mathbf{S} \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma})$. In addition, we also have $E(\mathbf{S} \mid Y, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma}) \in \Lambda_f \subset \Lambda$ since $E\{E(\mathbf{S} \mid Y, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma}) \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\gamma}\} = \mathbf{0}$. $\square$

Proposition S4.2. *Under model (1), the efficient score function for* $\boldsymbol{\beta}$ *is*

$$\mathbf{S}_{o\text{eff}} = g\{Y, \boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{\alpha}(\mathbf{X})\}\left\{\boldsymbol{\Theta}_{\boldsymbol{\beta}}\boldsymbol{\alpha}_c(\mathbf{X}) - \int_0^1 \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_c(t)\mathbf{w}_0(t)dt\right\} \tag{S47}$$

where $\mathbf{w}_0(t)$ *is a* $(J-1)$*-dimensional function that satisfies*

$$\begin{aligned}
&\boldsymbol{\Theta}_{\boldsymbol{\beta}}E[\boldsymbol{\alpha}_c(\mathbf{X})g^2\{Y, \boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{\alpha}(\mathbf{X})\}\mathbf{X}_c^{\mathrm{T}}(t)]\boldsymbol{\beta} \qquad \text{(S48)} \\
= &\int_0^1 E[\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_c(s)g^2\{Y, \boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{\alpha}(\mathbf{X})\}\mathbf{X}_c^{\mathrm{T}}(t)\boldsymbol{\beta}]\mathbf{w}_0(s)ds.
\end{aligned}$$

Here we use the subindex $_o$ *to indicate quantities calculated under the original model (1).*

Proof: We first make the observation that $\Lambda^{\perp}$ is the nuisance tangent space orthogonal complement of the B-spline approximated model (2). Follow the same arguments as those lead to Proposition 3, it can be shown that the spaces corresponding to Λ_x, Λ_f for model (1) are

$$\begin{aligned} \Lambda_{ox} &= [\mathbf{f}\{\mathbf{X}(t)\} : \forall \mathbf{f} \text{ such that } E(\mathbf{f}) = \mathbf{0}], \\ \Lambda_{of} &= \left[\mathbf{f}\left\{Y, \boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{\alpha}(\mathbf{X})\right\} : \forall \mathbf{f} \text{ such that } E\{\mathbf{f} \mid \mathbf{X}(t)\} = E\left\{\mathbf{f} \mid \boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{\alpha}(\mathbf{X})\right\} = \mathbf{0}\right]. \end{aligned}$$

In addition, the nuisance tangent space with respect to $\alpha(t)$ is

$$\widetilde{\Lambda}_{o\alpha} = \left[g\left\{Y, \boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{\alpha}(\mathbf{X})\right\} \int_0^1 \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}(t)\mathbf{w}(t)dt : \forall \mathbf{w}(t) \in \mathcal{R}^{J-1}\right].$$

After projecting $\widetilde{\Lambda}_{o\alpha}$ to Λ_{ox} and Λ_{of} to get the residual, we obtain

$$\begin{aligned} \Lambda_{o\alpha} &\equiv \widetilde{\Lambda}_{o\alpha} - \Pi(\widetilde{\Lambda}_{o\alpha} \mid \Lambda_{ox} + \Lambda_{of}) \\ &= \left[g\left\{Y, \boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{\alpha}(\mathbf{X})\right\} \int_0^1 \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_c(t)\mathbf{w}(t)dt : \forall \mathbf{w}(t) \in \mathcal{R}^{J-1}\right]. \end{aligned}$$

We can easily see that with respect to the model in (1), the nuisance tangent space is $\Lambda = \Lambda_{ox} \oplus \Lambda_{of} \oplus \Lambda_{o\alpha}$.

We can see that the score function with respect to the parameters in $\boldsymbol{\beta}$ is

$$\begin{aligned} \mathbf{S}_{o\beta} &= g\{Y, \boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{\alpha}(\mathbf{X})\}\boldsymbol{\Theta}_{\boldsymbol{\beta}}\boldsymbol{\alpha}(\mathbf{X}) \\ &= g\{Y, \boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{\alpha}(\mathbf{X})\}\left\{\boldsymbol{\Theta}_{\boldsymbol{\beta}}\boldsymbol{\alpha}_c(\mathbf{X}) - \int_0^1 \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_c(t)\mathbf{w}_0(t)dt\right\} \\ &\quad + g\{Y, \boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{\alpha}(\mathbf{X})\}\int_0^1 \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_c(t)\mathbf{w}_0(t)dt \\ &\quad + g\{Y, \boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{\alpha}(\mathbf{X})\}\boldsymbol{\Theta}_{\boldsymbol{\beta}}E\{\boldsymbol{\alpha}(\mathbf{X}) \mid \boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{\alpha}(\mathbf{X})\}, \end{aligned}$$

where the second and third summands of $\mathbf{S}_{o\beta}$ belong to $\Lambda_{o\alpha}$ and Λ_{of} respectively. We can also verify that the first summand of $\mathbf{S}_{o\beta}$ is simultaneously orthogonal to $\Lambda_{ox}, \Lambda_{of}$ and $\Lambda_{o\alpha}$, hence the efficient score of model (1) is indeed as given in (S47). $\square$

<u>**Lemma**</u> **1.** *There is a constant $D_r > 0$ such that for each spline $\sum_{k=1}^{d_\gamma} c_k B_{rk}$, and for each $1 \leq p \leq \infty$*

$$D_r\|\mathbf{c}'\|_p \leq \|\sum_{k=1}^{d_\gamma} c_k B_{rk}\|_p \leq \|\mathbf{c}'\|_p,$$

where $\mathbf{c}' = \{c_k\{(t_k - t_{k-r})/r\}^{1/p}, k = 1, \ldots, d_{\boldsymbol{\gamma}}\}^{\mathrm{T}}$.

Proof: This is a direct consequence of Theorem 5.4.2 on page 145 in DeVore and Lorentz (1993). □

Lemma 2. *Let* $\mathbf{u}$ *be a* $d_{\boldsymbol{\gamma}}$*-dimensional vector with* $\|\mathbf{u}\|_2 = 1$*. There exist positive constants* D_1, D_2, D_3, D_4 *such that*

$$\begin{aligned} D_1 h_b \leq \mathbf{u}^{\mathrm{T}}\mathbf{C}(\boldsymbol{\beta})\mathbf{u} \leq D_2 h_b, \\ D_2^{-1}h_b^{-1} \leq \mathbf{u}^{\mathrm{T}}\mathbf{C}(\boldsymbol{\beta})^{-1}\mathbf{u} \leq D_1^{-1}h_b^{-1}, \end{aligned}$$

and

$$\begin{aligned} D_3 h_b \leq \mathbf{u}^{\mathrm{T}}\widehat{\mathbf{C}}(\boldsymbol{\beta})\mathbf{u} \leq D_4 h_b, \\ D_4^{-1}h_b^{-1} \leq \mathbf{u}^{\mathrm{T}}\widehat{\mathbf{C}}(\boldsymbol{\beta})^{-1}\mathbf{u} \leq D_3^{-1}h_b^{-1} \end{aligned}$$

in probability.

Proof: First note that by the Cauchy-Schwartz inequality, Lemma 1 and Condition (A6), we have

$$\begin{aligned} & E\left[\left\{\mathbf{u}^{\mathrm{T}}\int_0^1 \mathbf{B}_r(t)\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}(t)dt\right\}^2\right] \\ \leq\ & E\left[\int_0^1\left\{\sum_{k=1}^{d_\gamma} u_k B_{rk}(t)\right\}^2 dt \int_0^1 \{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}(t)\}^2 dt\right] \\ =\ & \int_0^1\left\{\sum_{k=1}^{d_\gamma} u_k B_{rk}(t)\right\}^2 dt E\left\{\int_0^1 \{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}(t)\}^2 dt\right\} \\ \leq\ & \|\mathbf{u}'\|_2^2 O(1) \\ =\ & O(h_b), \end{aligned}$$

where $\mathbf{u}' = \{u_k\{(t_k - t_{k-r})/r\}^{1/2}, k = 1, \ldots, d_{\boldsymbol{\gamma}}\}^{\mathrm{T}}$, whose L_2 norm is of order $O(h_b^{1/2})$. Thus we have $\mathbf{u}^{\mathrm{T}}\mathbf{C}(\boldsymbol{\beta})\mathbf{u} \leq D_2 h_b$ for some positive constant $D_2 < \infty$.

As shown in (28) in Cardot et al. (2003), since the eigenvalues of the covariance operator $\Gamma(\boldsymbol{\beta})$ are strictly positive, there is a positive constant D such that

$$< \Gamma(\boldsymbol{\beta})\phi, \phi > \geq D\|\phi\|^2, \text{ for } \phi \in \mathcal{H}.$$

Note that $\mathbf{B}_r^{\mathrm{T}}\mathbf{u} \in \mathcal{H}$, so we have

$$\mathbf{u}^{\mathrm{T}}\mathbf{C}(\boldsymbol{\beta})\mathbf{u} =< \Gamma(\boldsymbol{\beta})\mathbf{B}_r^{\mathrm{T}}\mathbf{u}, \mathbf{B}_r^{\mathrm{T}}\mathbf{u} >\geq D\|\mathbf{B}_r^{\mathrm{T}}\mathbf{u}\|_2 \geq D_1\|\mathbf{u}\|^2 h_b = D_1 h_b$$

for a positive constant D_1 by Lemma 1 and Condition (A4). Therefore, $D_1 h_b \leq \mathbf{u}^{\mathrm{T}}\mathbf{C}(\boldsymbol{\beta})\mathbf{u} \leq D_2 h_b$. And so $D_2^{-1}h_b^{-1} \leq \mathbf{u}^{\mathrm{T}}\mathbf{C}^{-1}(\boldsymbol{\beta})\mathbf{u} \leq D_1^{-1}h_b^{-1}$. Further, with Theorem 1.19 in Chatelin (1983), we have

$$\|\widehat{\mathbf{C}}(\boldsymbol{\beta}) - \mathbf{C}(\boldsymbol{\beta})\|_2 \leq \sup_{1\leq l\leq d_\gamma} \sum_{k=1}^{d_\gamma} \|\Gamma_n - \Gamma\|_2| < B_{rk}, B_{rl} > |$$

As shown in Cardot et al. (2003), Lemma 5.3 in Cardot et al. (1999) implies $\|\Gamma_n - \Gamma\|_2 = o_p(n^{(h_b-1)/2})$. Further, by the property of B-spline basis, we have when $|k-l| > r$, $B_{rk}B_{rl} = 0$. Therefore, $\sup_{1\leq l\leq d_\gamma}\sum_{k=1}^{d_\gamma} | < B_{rk}, B_{rl} > | = O(h_b)$, which implies

$$\|\widehat{\mathbf{C}}(\boldsymbol{\beta}) - \mathbf{C}(\boldsymbol{\beta})\|_2 \leq o_p(h_b n^{(h_b-1)/2}). \tag{S49}$$

Now because $h_b < 1$, combine with the result that $D_1 h_b \leq \mathbf{u}^{\mathrm{T}}\mathbf{C}\mathbf{u} \leq D_2 h_b$, by the triangular inequality we obtain

$$\begin{aligned}
\mathbf{u}^{\mathrm{T}}\widehat{\mathbf{C}}(\boldsymbol{\beta})\mathbf{u} &= \mathbf{u}^{\mathrm{T}}\mathbf{C}(\boldsymbol{\beta})\mathbf{u} + \mathbf{u}^{\mathrm{T}}\{\widehat{\mathbf{C}}(\boldsymbol{\beta}) - \mathbf{C}(\boldsymbol{\beta})\}\mathbf{u} \\
&\leq D_2 h_b + \|\mathbf{u}\|_2\|\{\widehat{\mathbf{C}}(\boldsymbol{\beta}) - \mathbf{C}(\boldsymbol{\beta})\}\mathbf{u}\|_2 \\
&\leq D_2 h_b + \|\{\widehat{\mathbf{C}}(\boldsymbol{\beta}) - \mathbf{C}(\boldsymbol{\beta})\}\|_2\|\mathbf{u}\|_2 \\
&= D_2 h_b + o_p(h_b)
\end{aligned}$$

and

$$\begin{aligned}
\mathbf{u}^{\mathrm{T}}\widehat{\mathbf{C}}(\boldsymbol{\beta})\mathbf{u} &= \mathbf{u}^{\mathrm{T}}\mathbf{C}(\boldsymbol{\beta})\mathbf{u} + \mathbf{u}^{\mathrm{T}}\{\widehat{\mathbf{C}}(\boldsymbol{\beta}) - \mathbf{C}(\boldsymbol{\beta})\}\mathbf{u} \\
&\geq D_1 h_b - \|\mathbf{u}\|_2\|\{\widehat{\mathbf{C}}(\boldsymbol{\beta}) - \mathbf{C}(\boldsymbol{\beta})\}\mathbf{u}\|_2 \\
&\geq D_1 h_b - \|\{\widehat{\mathbf{C}}(\boldsymbol{\beta}) - \mathbf{C}(\boldsymbol{\beta})\}\|_2\|\mathbf{u}\|_2 \\
&= D_1 h_b + o_p(h_b).
\end{aligned}$$

Thus, $D_3 h_b \leq \mathbf{u}^{\mathrm{T}}\widehat{\mathbf{C}}(\boldsymbol{\beta})\mathbf{u} \leq D_4 h_b$ in probability for some positive constant $D_3, D_4 < \infty$. And so $D_4^{-1}h_b \leq \mathbf{u}^{\mathrm{T}}\widehat{\mathbf{C}}(\boldsymbol{\beta})^{-1}\mathbf{u} \leq D_3^{-1}h_b^{-1}$ in probability. This proves the result. $\square$

The result in Lemma 2 can be generalized to any vector valued random function with almost surely bounded L_2 norm and positive definite second moment operator. Therefore, we have the following corollary.

Corollary 1. *Let* $\mathbf{Q}_i(t)$ *be a* d_q*-dimensional vector valued random function with bounded* L_2 *norm, and let* $\mathbf{v}$ *be an arbitrary* d_q*-dimensional vector. Let* $\Gamma_{\mathbf{Q}}(\mathbf{v})$ *be the second moment operator generated by* $\mathbf{v}^{\mathrm{T}}\mathbf{Q}_i(\cdot)$*, i.e.,* $\Gamma_{\mathbf{Q}}(\mathbf{v})\phi(t) = E\{< \mathbf{v}^{\mathrm{T}}\mathbf{Q}_i, \phi > \mathbf{v}^{\mathrm{T}}\mathbf{Q}_i(t)\}$*. Assume* $\Gamma_{\mathbf{Q}}(\mathbf{v})$ *is strictly positive definite. We define the matrix* $\mathbf{C}_{\mathbf{Q}}(\mathbf{v})$ *to be a* $d_{\boldsymbol{\gamma}} \times d_{\boldsymbol{\gamma}}$ *matrix with its* (k,l) *element*

$$E\{\int_0^1 B_{rk}(t)\mathbf{v}^{\mathrm{T}}\mathbf{Q}_i(t)dt \int_0^1 B_{rl}(t)\mathbf{v}^{\mathrm{T}}\mathbf{Q}_i(t)dt\},$$

and $\widehat{\mathbf{C}}_{\mathbf{Q}}(\mathbf{v})$ *to be a* $d_{\boldsymbol{\gamma}} \times d_{\boldsymbol{\gamma}}$ *matrix with its* (k,l) *element*

$$n^{-1}\sum_{i=1}^{n}\int_0^1 B_{rk}(t)\mathbf{v}^{\mathrm{T}}\mathbf{Q}_i(t)dt \int_0^1 B_{rl}(t)\mathbf{v}^{\mathrm{T}}\mathbf{Q}_i(t)dt.$$

Let $\mathbf{u}$ *be a* $d_{\boldsymbol{\gamma}}$*-dimensional vector with* $\|\mathbf{u}\|_2 = 1$*. Then there are positive constants* D_5, D_6, D_7, D_8 *such that*

$$\begin{aligned} &D_5 h_b \le \mathbf{u}^{\mathrm{T}}\mathbf{C}_{\mathbf{Q}}(\mathbf{v})\mathbf{u} \le D_6 h_b, \\ &D_6^{-1}h_b^{-1} \le \mathbf{u}^{\mathrm{T}}\mathbf{C}_{\mathbf{Q}}^{-1}(\mathbf{v})\mathbf{u} \le D_5^{-1}h_b^{-1}, \end{aligned}$$

and

$$\begin{aligned} &D_7 h_b \le \mathbf{u}^{\mathrm{T}}\widehat{\mathbf{C}}_{\mathbf{Q}}(\mathbf{v})\mathbf{u} \le D_8 h_b, \\ &D_8^{-1}h_b^{-1} \le \mathbf{u}^{\mathrm{T}}\widehat{\mathbf{C}}_{\mathbf{Q}}^{-1}(\mathbf{v})\mathbf{u} \le D_7^{-1}h_b^{-1}, \end{aligned}$$

in probability. And

$$\|\widehat{\mathbf{C}}_{\mathbf{Q}}(\mathbf{v}) - \mathbf{C}_{\mathbf{Q}}(\mathbf{v})\|_2 \le o_p(h_b n^{(h_b-1)/2}).$$

The proof of Corollary 1 follows the same arguments as those in the proof of Lemma 2 hence is omitted.

<u>Lemma</u> 3. *Assume $C_i(\cdot)$ is a continuous random function of $t \in [0,1]$. At each t, $|E\{C_i(t)\}| < \infty$. $\|C_i(\cdot)\|_2 < \infty$ a.s.. Then*

$$|n^{-1}\sum_{i=1}^{n}\int_0^1 B_{rk}(t)C_i(t)dt| = O_p(h_b)$$

if $E\{C_i(t)\} \neq 0$ and

$$|n^{-1}\sum_{i=1}^{n}\int_0^1 B_{rk}(t)C_i(t)dt| = O_p\{\sqrt{h_b n^{-1} log(n)}\}$$

if $E\{C_i(t)\} = 0$.

Proof: By the Bernstein's inequality in Bosq (1998), we have

$$\begin{aligned} & \left| 1/n\sum_{i=1}^{n}\int_0^1 B_{rk}(t)C_i(t)dt - E\left\{\int_0^1 B_{rk}(t)C_i(t)dt\right\}\right| \\ = \; & O_p\left[\left[\sum_{i=1}^{n} E\left\{1/n\int_0^1 B_{rk}(t)C_i(t)dt\right\}^2 \log n\right]^{1/2}\right] \\ = \; & O_p\{\sqrt{h_b n^{-1}\log(n)}\}. \end{aligned}$$

The last equality holds from Corollary 1, by choosing $\mathbf{v}^{\mathrm{T}}\mathbf{Q}_i(t) = C_i(t)$ and setting $\mathbf{u}$ with $u_k = 1$ and $u_l = 0$, for $l \neq k$. Now if $E\{C_i(t)\} \neq 0$, then because B_{rk} is positive in the interval (t_{k-r}, t_k), and is 0 otherwise (page 88 in De Boor (1978))

$$\begin{aligned} \left|E\left\{\int_0^1 B_{rk}(t)C_i(t)dt\right\}\right| &= \left|\int_0^1 B_{rk}(t)E\{C_i(t)\}dt\right| \\ &= |E\{C_i(\xi)\}\int_0^1 B_{rk}(t)dt| \\ &= |E\{C_i(\xi)\}|\int_0^1 B_{rk}(t)dt \\ &\leq D_9(t_k - t_{k-r}) \\ &= O_p(h_b). \end{aligned} \tag{S50}$$

where ξ is a point in the interval $[0,1]$, D_9 is a finite constant.The second equality holds by the assumption that $C_i(\cdot)$ is continuous function in t and the mean value theorem. The inequality holds because the support of B_{rk} is the interval (t_{k-r}, t_k) and $|E\{C_i(t)\}| < \infty$ for any $t \in [0,1]$. Therefore, by Condition (A4) that $N^{-1}n(\log n)^{-1} \to \infty$ we have

$$\|n^{-1}\sum_{i=1}^{n}\int_0^1 B_{rk}(t)C_i(t)dt\|_2 = O_p(h_b)$$

for $E\{C_i(t)\} \neq 0$ and

$$\|n^{-1}\sum_{i=1}^{n}\int_0^1 B_{rk}(t)C_i(t)dt\|_2 = \sqrt{h_b n^{-1}\log(n)}$$

for $E\{C_i(t)\} = 0$. This proves the results. $\square$

<u>**Lemma**</u> **4.** *Assume* $\widehat{\boldsymbol\beta} - \boldsymbol\beta_0 = o_p(1)$. *Then*

$$\begin{aligned}
&\widehat{\boldsymbol\gamma}^{-}(\boldsymbol\beta)/\partial\boldsymbol\beta^{\mathrm T}|_{\boldsymbol\beta^*}\\
= \ & -\Bigg(\Bigg\{E\Bigg(\boldsymbol\Theta_{\boldsymbol\gamma}\left[\mathbf Z_i - E\{\mathbf Z_i|\boldsymbol\beta_0\boldsymbol\alpha_0(\mathbf X_i\}\right]^{\mathrm T}\boldsymbol\beta_0\\
&\times\frac{\partial\left(g^*\{Y_i,\boldsymbol\beta_0^{\mathrm T}\boldsymbol\alpha_0(\mathbf X_i)\} - E[g^*\{Y_i,\boldsymbol\beta_0^{\mathrm T}\boldsymbol\alpha_0(\mathbf X_i)\}|\boldsymbol\beta_0^{\mathrm T}\boldsymbol\alpha_0(\mathbf X_i)]]\right)}{\partial\boldsymbol\beta_0^{\mathrm T}\boldsymbol\alpha_0(\mathbf X_i)}\\
&\times\boldsymbol\beta_0^{\mathrm T}[\mathbf Z_i - E\{\mathbf Z_i|\boldsymbol\beta_0^{\mathrm T}\boldsymbol\alpha_0(\mathbf X_i)\}]\boldsymbol\Theta_{\boldsymbol\gamma}^{\mathrm T}\Bigg)\Bigg\}^{-1}\Bigg)\Bigg\{E(\boldsymbol\Theta_{\boldsymbol\gamma}\left[\mathbf Z_i - E\{\mathbf Z_i|\boldsymbol\beta_0^{\mathrm T}\boldsymbol\alpha_0(\mathbf X_i)\}\right]^{\mathrm T}\\
&\times\boldsymbol\beta_0\boldsymbol\alpha_0(\mathbf X_i)^{\mathrm T}\frac{\partial[g^*\{Y_i,\boldsymbol\beta_0^{\mathrm T}\boldsymbol\alpha_0(\mathbf X_i)\} - E\{g^*\{Y_i,\boldsymbol\beta_0^{\mathrm T}\boldsymbol\alpha_0(\mathbf X_i)\}|\boldsymbol\beta_0^{\mathrm T}\boldsymbol\alpha_0(\mathbf X_i)\}]}{\partial\{\boldsymbol\beta_0^{\mathrm T}\boldsymbol\alpha_0(\mathbf X_i))\}}\Bigg\}\\
&\times\{1+o_p(1)\},
\end{aligned}$$

where $\boldsymbol\beta^*$ *is a point on the line connecting* $\widehat{\boldsymbol\beta}, \boldsymbol\beta_0$. *In addition,* $\|\widehat{\boldsymbol\gamma}^{-}(\boldsymbol\beta)/\partial\boldsymbol\beta^{\mathrm T}|_{\boldsymbol\beta=\boldsymbol\beta^*}\|_2 = O_p(h_b^{-1/2})$.

Proof: $\widehat{\boldsymbol\gamma}^{-}(\boldsymbol\beta)$ satisfies

$$\begin{aligned}
&\sum_{i=1}^{n}\widehat{\mathbf S}_{\mathrm{eff}\boldsymbol\gamma}\{Y_i,\mathbf Z_i,\boldsymbol\beta,\widehat{\gamma}(\boldsymbol\beta),g^*\}\\
= \ &\sum_{i=1}^{n}\left\{g^*\{Y_i,\boldsymbol\beta^{\mathrm T}\mathbf Z_i\widehat{\gamma}(\boldsymbol\beta)\} - \frac{\sum_{j=1}^{J}K_h\{\boldsymbol\beta^{\mathrm T}\mathbf Z_j\widehat{\gamma}(\boldsymbol\beta) - \boldsymbol\beta^{\mathrm T}\mathbf Z_i\widehat{\gamma}(\boldsymbol\beta)\}g^*\{Y_j,\boldsymbol\beta^{\mathrm T}\mathbf Z_j\widehat{\gamma}(\boldsymbol\beta)\}}{\sum_{j=1}^{J}K_h\{\boldsymbol\beta^{\mathrm T}\mathbf Z_j\widehat{\gamma}(\boldsymbol\beta) - \boldsymbol\beta^{\mathrm T}\mathbf Z_i\widehat{\gamma}(\boldsymbol\beta)\}}\right\}\\
&\boldsymbol\Theta_{\boldsymbol\gamma}\left\{\mathbf Z_i - \frac{\sum_{j=1}^{J}K_h\{\boldsymbol\beta^{\mathrm T}\mathbf Z_j\widehat{\gamma}(\boldsymbol\beta) - \boldsymbol\beta^{\mathrm T}\mathbf Z_i\widehat{\gamma}(\boldsymbol\beta)\}\mathbf Z_j}{\sum_{j=1}^{J}K_h\{\boldsymbol\beta^{\mathrm T}\mathbf Z_j\widehat{\gamma}(\boldsymbol\beta) - \boldsymbol\beta^{\mathrm T}\mathbf Z_i\widehat{\gamma}(\boldsymbol\beta)\}}\right\}^{\mathrm T}\boldsymbol\beta
\end{aligned}$$

$$= \mathbf{0}$$

for any $\boldsymbol\beta$. Taking the derivative with respect to $\boldsymbol\beta$ on both sides, we have

$$\sum_{i=1}^n \partial\left\{g^*\{Y_i, \boldsymbol\beta^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta)\} - \frac{\sum_{j=1}^J K_h\{\boldsymbol\beta^{\mathrm T}\mathbf{Z}_j\widehat{\boldsymbol\gamma}(\boldsymbol\beta) - \boldsymbol\beta^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta)\}g^*\{Y_j, \boldsymbol\beta^{\mathrm T}\mathbf{Z}_j\widehat{\boldsymbol\gamma}(\boldsymbol\beta)\}}{\sum_{j=1}^J K_h\{\boldsymbol\beta^{\mathrm T}\mathbf{Z}_j\widehat{\boldsymbol\gamma}(\boldsymbol\beta) - \boldsymbol\beta^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta)\}}\right\}$$
$$\boldsymbol\Theta_\gamma\left\{\mathbf{Z}_i - \frac{\sum_{j=1}^J K_h\{\boldsymbol\beta^{\mathrm T}\mathbf{Z}_j\widehat{\boldsymbol\gamma}(\boldsymbol\beta) - \boldsymbol\beta^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta)\}\mathbf{Z}_j}{\sum_{j=1}^J K_h\{\boldsymbol\beta^{\mathrm T}\mathbf{Z}_j\widehat{\boldsymbol\gamma}(\boldsymbol\beta) - \boldsymbol\beta^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta)\}}\right\}^{\mathrm T} \boldsymbol\beta/\partial\boldsymbol\beta^{\mathrm T}$$
$$+\sum_{i=1}^n \partial\left\{g^*\{Y_i, \boldsymbol\beta^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta)\} - \frac{\sum_{j=1}^J K_h\{\boldsymbol\beta^{\mathrm T}\mathbf{Z}_j\widehat{\boldsymbol\gamma}(\boldsymbol\beta) - \boldsymbol\beta^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta)\}g^*\{Y_j, \boldsymbol\beta^{\mathrm T}\mathbf{Z}_j\widehat{\boldsymbol\gamma}(\boldsymbol\beta)\}}{\sum_{j=1}^J K_h\{\boldsymbol\beta^{\mathrm T}\mathbf{Z}_j\widehat{\boldsymbol\gamma}(\boldsymbol\beta) - \boldsymbol\beta^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta)\}}\right\}$$
$$\boldsymbol\Theta_\gamma\left\{\mathbf{Z}_i - \frac{\sum_{j=1}^J K_h\{\boldsymbol\beta^{\mathrm T}\mathbf{Z}_j\widehat{\boldsymbol\gamma}(\boldsymbol\beta) - \boldsymbol\beta^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta)\}\mathbf{Z}_j}{\sum_{j=1}^J K_h\{\boldsymbol\beta^{\mathrm T}\mathbf{Z}_j\widehat{\boldsymbol\gamma}(\boldsymbol\beta) - \boldsymbol\beta^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta)\}}\right\}^{\mathrm T} \boldsymbol\beta/\partial\widehat{\boldsymbol\gamma}(\boldsymbol\beta)^{\mathrm T}\partial\widehat{\boldsymbol\gamma}(\boldsymbol\beta)/\partial\boldsymbol\beta^{\mathrm T} = \mathbf{0}.$$

Therefore,

$$\widehat{\boldsymbol\gamma}^-(\boldsymbol\beta)/\partial\boldsymbol\beta^{\mathrm T}|_{\boldsymbol\beta^*} = \mathbf{U}\mathbf{V},$$

where

$$\begin{aligned}
\mathbf{U} &= -\Bigg(n^{-1}\sum_{i=1}^n \partial\Bigg\{g^*\{Y_i, \boldsymbol\beta^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta)\} \\
&\quad - \frac{\sum_{j=1}^J K_h\{\boldsymbol\beta^{\mathrm T}\mathbf{Z}_j\widehat{\boldsymbol\gamma}(\boldsymbol\beta) - \boldsymbol\beta^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta)\}g^*\{Y_j, \boldsymbol\beta^{\mathrm T}\mathbf{Z}_j\widehat{\boldsymbol\gamma}(\boldsymbol\beta)\}}{\sum_{j=1}^J K_h\{\boldsymbol\beta^{\mathrm T}\mathbf{Z}_j\widehat{\boldsymbol\gamma}(\boldsymbol\beta) - \boldsymbol\beta^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta)\}}\Bigg\} \\
&\quad \boldsymbol\Theta_\gamma\left\{\mathbf{Z}_i - \frac{\sum_{j=1}^J K_h\{\boldsymbol\beta^{\mathrm T}\mathbf{Z}_j\widehat{\boldsymbol\gamma}(\boldsymbol\beta) - \boldsymbol\beta^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta)\}\mathbf{Z}_j}{\sum_{j=1}^J K_h\{\boldsymbol\beta^{\mathrm T}\mathbf{Z}_j\widehat{\boldsymbol\gamma}(\boldsymbol\beta) - \boldsymbol\beta^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta)\}}\right\}^{\mathrm T} \boldsymbol\beta/\partial\widehat{\boldsymbol\gamma}^-(\boldsymbol\beta)^{\mathrm T}\Bigg|_{\boldsymbol\beta^*}\Bigg)^{-1} \\
&= -\Bigg(n^{-1}\sum_{i=1}^n \partial\left\{g^*(Y_i, \boldsymbol\beta^{\mathrm T}\mathbf{Z}_i\boldsymbol\gamma) - \frac{\sum_{j=1}^J K_h(\boldsymbol\beta^{\mathrm T}\mathbf{Z}_j\boldsymbol\gamma - \boldsymbol\beta^{\mathrm T}\mathbf{Z}_i\boldsymbol\gamma)g^*(Y_j, \boldsymbol\beta^{\mathrm T}\mathbf{Z}_j\boldsymbol\gamma)}{\sum_{j=1}^J K_h(\boldsymbol\beta^{\mathrm T}\mathbf{Z}_j\boldsymbol\gamma - \boldsymbol\beta^{\mathrm T}\mathbf{Z}_i\boldsymbol\gamma)}\right\} \\
&\quad \boldsymbol\Theta_\gamma\left\{\mathbf{Z}_i - \frac{\sum_{j=1}^J K_h(\boldsymbol\beta^{\mathrm T}\mathbf{Z}_j\boldsymbol\gamma - \boldsymbol\beta^{\mathrm T}\mathbf{Z}_i\boldsymbol\gamma)\mathbf{Z}_j}{\sum_{j=1}^J K_h(\boldsymbol\beta^{\mathrm T}\mathbf{Z}_j\boldsymbol\gamma - \boldsymbol\beta^{\mathrm T}\mathbf{Z}_i\boldsymbol\gamma)}\right\}^{\mathrm T} \boldsymbol\beta/\partial(\boldsymbol\beta^{\mathrm T}\mathbf{Z}_i\boldsymbol\gamma)\Bigg|_{\boldsymbol\gamma=\widehat{\boldsymbol\gamma}(\boldsymbol\beta^*)} \boldsymbol\beta_0^{\mathrm T}\mathbf{Z}_i\boldsymbol\Theta_\gamma^{\mathrm T}\Bigg)^{-1}, \\
\mathbf{V} &= n^{-1}\sum_{i=1}^n \partial\Bigg\{g^*\{Y_i, \boldsymbol\beta^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta)\} \\
&\quad - \frac{\sum_{j=1}^J K_h\{\boldsymbol\beta^{\mathrm T}\mathbf{Z}_j\widehat{\boldsymbol\gamma}(\boldsymbol\beta) - \boldsymbol\beta^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta)\}g^*\{Y_j, \boldsymbol\beta^{\mathrm T}\mathbf{Z}_j\widehat{\boldsymbol\gamma}(\boldsymbol\beta)\}}{\sum_{j=1}^J K_h\{\boldsymbol\beta^{\mathrm T}\mathbf{Z}_j\widehat{\boldsymbol\gamma}(\boldsymbol\beta) - \boldsymbol\beta^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta)\}}\Bigg\} \\
&\quad \boldsymbol\Theta_\gamma\left\{\mathbf{Z}_i - \frac{\sum_{j=1}^J K_h\{\boldsymbol\beta^{\mathrm T}\mathbf{Z}_j\widehat{\boldsymbol\gamma}(\boldsymbol\beta) - \boldsymbol\beta^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta)\}\mathbf{Z}_j}{\sum_{j=1}^J K_h\{\boldsymbol\beta^{\mathrm T}\mathbf{Z}_j\widehat{\boldsymbol\gamma}(\boldsymbol\beta) - \boldsymbol\beta^{\mathrm T}\mathbf{Z}_i\widehat{\boldsymbol\gamma}(\boldsymbol\beta)\}}\right\}^{\mathrm T} \boldsymbol\beta/\partial\boldsymbol\beta^{\mathrm T}\Bigg|_{\boldsymbol\beta^*}.
\end{aligned}$$

Further, the convergence of $\boldsymbol{\beta}^*$ implies $\sup_{t\in[0,1]}\{\mathbf{B}_r(t)^{\mathrm{T}}\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}^*)-\mathbf{B}_r(t)^{\mathrm{T}}\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}_0)\} = o_p(1)$, and thus $\sup_{t\in[0,1]}\{\mathbf{B}_r(t)^{\mathrm{T}}\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}^*) - \alpha_0(t)\} = o_p(1)$. Now by the uniform convergence of the kernel estimator, and the consistency of $\mathbf{B}_r(\cdot)^{\mathrm{T}}\widehat{\boldsymbol{\gamma}}(\boldsymbol{\beta}^*)$ and $\boldsymbol{\beta}^*$, $\mathbf{U}$ is asymptotic equivalent to $-\mathbf{T}^{-1}$, where $\mathbf{T}$ is defined in (S22), and in turn

$$
\begin{aligned}
\mathbf{U} &= -\mathbf{T}_{00}^{-1} + o_p(\mathbf{T}_{00}^{-1}) \\
&= -\Bigg(\Bigg\{E\Bigg(\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\left[\mathbf{Z}_i - E\{\mathbf{Z}_i|\boldsymbol{\beta}_0\boldsymbol{\alpha}_0(\mathbf{X}_i\}\right]^{\mathrm{T}}\boldsymbol{\beta}_0 \\
&\quad\times\frac{\partial\left(g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\} - E[g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)]\right]\big)}{\partial\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)} \\
&\quad\times\boldsymbol{\beta}_0^{\mathrm{T}}[\mathbf{Z}_i - E\{\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}]\boldsymbol{\Theta}_{\gamma}^{\mathrm{T}}\Bigg)\Bigg\}^{-1} \\
&\quad+o_p\Bigg[\Bigg\{E\Bigg(\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\left[\mathbf{Z}_i - E\{\mathbf{Z}_i|\boldsymbol{\beta}_0\boldsymbol{\alpha}_0(\mathbf{X}_i\}\right]^{\mathrm{T}}\boldsymbol{\beta}_0 \\
&\quad\times\frac{\partial\left(g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\} - E[g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)]\right]\big)}{\partial\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)} \\
&\quad\times\boldsymbol{\beta}_0^{\mathrm{T}}[\mathbf{Z}_i - E\{\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}]\boldsymbol{\Theta}_{\gamma}^{\mathrm{T}}\Bigg)\Bigg\}^{-1}\Bigg]\Bigg).
\end{aligned} \tag{S51}
$$

Also, use the similar argument as those led to (S51), we have

$$
\begin{aligned}
\mathbf{V} &= n^{-1}\sum_{i=1}^{n}\partial[g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\} \\
&\quad -E\{g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}]\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\left[\mathbf{Z}_i - E\{\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\right]^{\mathrm{T}}\boldsymbol{\beta}_0/\partial\boldsymbol{\beta}^{\mathrm{T}} \\
&\quad +o_p\Bigg(n^{-1}\sum_{i=1}^{n}\partial[g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\} \\
&\quad -E\{g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}]\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\left[\mathbf{Z}_i - E\{\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\right]^{\mathrm{T}}\boldsymbol{\beta}_0/\partial\boldsymbol{\beta}^{\mathrm{T}}\Bigg) \\
&= n^{-1}\sum_{i=1}^{n}\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\left[\mathbf{Z}_i - E\{\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\right]^{\mathrm{T}}\boldsymbol{\beta}_0\boldsymbol{\alpha}_0(\mathbf{X}_i)^{\mathrm{T}} \\
&\quad\times\partial[g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\} - E\{g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}]/\partial\{\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\} \\
&\quad +o_p\Bigg(n^{-1}\sum_{i=1}^{n}\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\left[\mathbf{Z}_i - E\{\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\right]^{\mathrm{T}}\boldsymbol{\beta}_0\boldsymbol{\alpha}_0(\mathbf{X}_i)^{\mathrm{T}} \\
&\quad\times\partial[g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\} - E\{g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}]/\partial\{\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\Bigg).
\end{aligned}
$$

Therefore, we can write

$$
\begin{aligned}
& \widehat{\boldsymbol{\gamma}}^{-}(\boldsymbol{\beta})/\partial\boldsymbol{\beta}^{\mathrm{T}}|_{\boldsymbol{\beta}^*} \\
= & -\Bigg(\Bigg\{E\Bigg(\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\left[\mathbf{Z}_i - E\{\mathbf{Z}_i|\boldsymbol{\beta}_0\boldsymbol{\alpha}_0(\mathbf{X}_i\}\right]^{\mathrm{T}}\boldsymbol{\beta}_0 \\
& \times\frac{\partial\left(g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\} - E[g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)]\right])}{\partial\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)} \\
& \times\boldsymbol{\beta}_0^{\mathrm{T}}[\mathbf{Z}_i - E\{\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}]\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\Bigg)\Bigg\}^{-1}\Bigg)\Bigg(n^{-1}\sum_{i=1}^{n}\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\left[\mathbf{Z}_i - E\{\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\right]^{\mathrm{T}} \\
& \boldsymbol{\beta}_0\boldsymbol{\alpha}_0(\mathbf{X}_i)^{\mathrm{T}}\partial[g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\} - E\{g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}] \\
& /\partial\{\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\Bigg)\{1+o_p(1)\} \\
= & -\Bigg(\Bigg\{E\Bigg(\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\left[\mathbf{Z}_i - E\{\mathbf{Z}_i|\boldsymbol{\beta}_0\boldsymbol{\alpha}_0(\mathbf{X}_i\}\right]^{\mathrm{T}}\boldsymbol{\beta}_0 \\
& \times\frac{\partial\left(g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\} - E[g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)]\right])}{\partial\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)} \\
& \times\boldsymbol{\beta}_0^{\mathrm{T}}[\mathbf{Z}_i - E\{\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}]\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\Bigg)\Bigg\}^{-1}\Bigg)\Bigg\{E(\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\left[\mathbf{Z}_i - E\{\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\right]^{\mathrm{T}} \\
& \times\boldsymbol{\beta}_0\boldsymbol{\alpha}_0(\mathbf{X}_i)^{\mathrm{T}}\partial[g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\} - E\{g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}] \\
& /\partial\{\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\})\Bigg\}\{1+o_p(1)\}.
\end{aligned}
$$

Now we can write

$$
\partial\widehat{\boldsymbol{\gamma}}^{-}(\boldsymbol{\beta})/\partial\boldsymbol{\beta}^{\mathrm{T}}|_{\boldsymbol{\beta}=\boldsymbol{\beta}^*} \quad = \quad \mathbf{P} + o_p(\mathbf{P}), \tag{S52}
$$

where

$$
\begin{aligned}
\mathbf{P} \quad = \quad & -\Bigg(\Bigg\{E\Bigg(\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\left[\mathbf{Z}_i - E\{\mathbf{Z}_i|\boldsymbol{\beta}_0\boldsymbol{\alpha}_0(\mathbf{X}_i\}\right]^{\mathrm{T}}\boldsymbol{\beta}_0 \qquad \text{(S53)} \\
& \times\frac{\partial\left(g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\} - E[g^*\{Y_i,\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)]\right])}{\partial\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)} \\
& \times\boldsymbol{\beta}_0^{\mathrm{T}}[\mathbf{Z}_i - E\{\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}]\boldsymbol{\Theta}_{\boldsymbol{\gamma}}^{\mathrm{T}}\Bigg)\Bigg\}^{-1}\Bigg)\mathbf{M},
\end{aligned}
$$

where

$$
\mathbf{M} \quad = \quad E(\boldsymbol{\Theta}_{\boldsymbol{\gamma}}\left[\mathbf{Z}_i - E\{\mathbf{Z}_i|\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{\alpha}_0(\mathbf{X}_i)\}\right]^{\mathrm{T}}\boldsymbol{\beta}_0\boldsymbol{\alpha}_0(\mathbf{X}_i)^{\mathrm{T}}\boldsymbol{\Theta}_{\boldsymbol{\beta}}^{\mathrm{T}}
$$

$$\times\partial[g^*\{Y_i,\boldsymbol\beta_0^{\rm T}\boldsymbol\alpha_0(\mathbf X_i)\}-E\{g^*\{Y_i,\boldsymbol\beta_0^{\rm T}\boldsymbol\alpha_0(\mathbf X_i)\}|\boldsymbol\beta_0^{\rm T}\boldsymbol\alpha_0(\mathbf X_i)\}]/\partial\{\boldsymbol\beta_0^{\rm T}\boldsymbol\alpha_0(\mathbf X_i)\})$$

is a $(d_{\boldsymbol\gamma-1}\times J)$-dimensional matrix. We further have

$$\begin{aligned}
&\|\mathbf P\|_2\\
\le\ & \Bigg\|\Bigg(\Bigg\{E\Bigg(\boldsymbol\Theta_{\boldsymbol\gamma}\left[\mathbf Z_i-E\{\mathbf Z_i|\boldsymbol\beta_0\boldsymbol\alpha_0(\mathbf X_i)\}\right]^{\rm T}\boldsymbol\beta_0\\
&\times\frac{\partial\left(g^*\{Y_i,\boldsymbol\beta_0^{\rm T}\boldsymbol\alpha_0(\mathbf X_i)\}-E[g^*\{Y_i,\boldsymbol\beta_0^{\rm T}\boldsymbol\alpha_0(\mathbf X_i)\}|\boldsymbol\beta_0^{\rm T}\boldsymbol\alpha_0(\mathbf X_i)]\right])}{\partial\boldsymbol\beta_0^{\rm T}\boldsymbol\alpha_0(\mathbf X_i)}\\
&\times\boldsymbol\beta_0^{\rm T}[\mathbf Z_i-E\{\mathbf Z_i|\boldsymbol\beta_0^{\rm T}\boldsymbol\alpha_0(\mathbf X_i)\}]\boldsymbol\Theta_{\boldsymbol\gamma}^{\rm T}\Bigg)\Bigg\}^{-1}\Bigg)\Bigg\|_2\|\mathbf M\|_2\\
=\ & O_p(h_b^{-1})O_p(h_b^{1/2})\\
=\ & O_p(h_b^{-1/2}). \qquad\qquad (S54)
\end{aligned}$$

The second to the last equality holds because of the root n convergence of $\widehat{\boldsymbol\beta}$, and because

$$\begin{aligned}
&\Bigg\|\Bigg(\Bigg\{E\Bigg(\boldsymbol\Theta_{\boldsymbol\gamma}\left[\mathbf Z_i-E\{\mathbf Z_i|\boldsymbol\beta_0\boldsymbol\alpha_0(\mathbf X_i)\}\right]^{\rm T}\boldsymbol\beta_0\\
&\times\frac{\partial\left(g^*\{Y_i,\boldsymbol\beta_0^{\rm T}\boldsymbol\alpha_0(\mathbf X_i)\}-E[g^*\{Y_i,\boldsymbol\beta_0^{\rm T}\boldsymbol\alpha_0(\mathbf X_i)\}|\boldsymbol\beta_0^{\rm T}\boldsymbol\alpha_0(\mathbf X_i)]\right])}{\partial\boldsymbol\beta_0^{\rm T}\boldsymbol\alpha_0(\mathbf X_i)}\\
&\times\boldsymbol\beta_0^{\rm T}[\mathbf Z_i-E\{\mathbf Z_i|\boldsymbol\beta_0^{\rm T}\boldsymbol\alpha_0(\mathbf X_i)\}]\boldsymbol\Theta_{\boldsymbol\gamma}^{\rm T}\Bigg)\Bigg\}^{-1}\Bigg)\Bigg\|_2=O_p(h_b^{-1}) \qquad\qquad (S55)
\end{aligned}$$

by (S28). In addition,

$$\|\mathbf M\|_2\le\sqrt N\|\mathbf M\|_\infty.$$

Now since J is of finite dimension, the order of $\|\mathbf M\|_\infty$ is the same as the order of each element in $\mathbf M$. Further

$$\mathbf M=E\left[\int_0^1\{B_{rk}(t),k=2,\dots,d_{\boldsymbol\gamma}\}^{\rm T}\mathbf C_{14i}(t)dt\right]$$

with

$$\begin{aligned}
\mathbf C_{14i}(t)\ =\ & \left[\mathbf X_i(t)-E\{\mathbf X_i(t)|\boldsymbol\beta_0^{\rm T}\boldsymbol\alpha_0(\mathbf X_i)\}\right]^{\rm T}\boldsymbol\beta_0\boldsymbol\alpha_0(\mathbf X_i)^{\rm T}\partial[g^*\{Y_i,\boldsymbol\beta_0^{\rm T}\boldsymbol\alpha_0(\mathbf X_i)\}\\
&-E\{g^*\{Y_i,\boldsymbol\beta_0^{\rm T}\boldsymbol\alpha_0(\mathbf X_i)\}|\boldsymbol\beta_0^{\rm T}\boldsymbol\alpha_0(\mathbf X_i)\}]/\partial\{\boldsymbol\beta_0^{\rm T}\boldsymbol\alpha_0(\mathbf X_i)\}.
\end{aligned}$$

Each (k, l)th element of $\mathbf{M}$, M_{kl}, can be written as

$$M_{kl} = E\left[\int_0^1 \{B_{rk}(t)\} C_{14li}(t) dt\right],$$

where $C_{14li}(t)$ is the lth element of $C_{14i}(t)$. Since $E\{C_{14li}(t)\} \neq 0$, $|E\{C_{14li}(t)\}| < \infty$ and $\|C_{14li}(\cdot)\|_2 < \infty$. By (S50), we have $|M_{kl}| = O_p(h_b)$, $\|\mathbf{M}\|_\infty = O_p(h_b)$ and in turn

$$\|\mathbf{M}\|_2 \leq \sqrt{N} O_p(h_b) = O_p(h_b^{1/2}). \tag{S56}$$

Hence $\|\widehat{\boldsymbol{\gamma}}^{-}(\boldsymbol{\beta})/\partial \boldsymbol{\beta}^{\mathrm{T}}|_{\boldsymbol{\beta}=\boldsymbol{\beta}^*}\|_2 = O_p(\|\mathbf{P}\|_2) = O_p(h_b^{-1/2})$. This proves the result.

$\square$

S5 Correlation coefficients between the four air pollutants CO, NO_2, O_3 and SO_2

$$\begin{array}{c c}
 & \begin{array}{cccc} \mathrm{CO} & \mathrm{NO_2} & \mathrm{O_3} & \mathrm{SO_2} \end{array} \\
\begin{array}{c} \mathrm{CO} \\ \mathrm{NO_2} \\ \mathrm{O_3} \\ \mathrm{SO_2} \end{array} & \begin{pmatrix} 1.00 & 0.69 & -0.32 & 0.35 \\ 0.69 & 1.00 & -0.15 & 0.38 \\ -0.32 & -0.15 & 1.00 & -0.16 \\ 0.35 & 0.38 & -0.16 & 1.00 \end{pmatrix}
\end{array}$$

S6 Figures

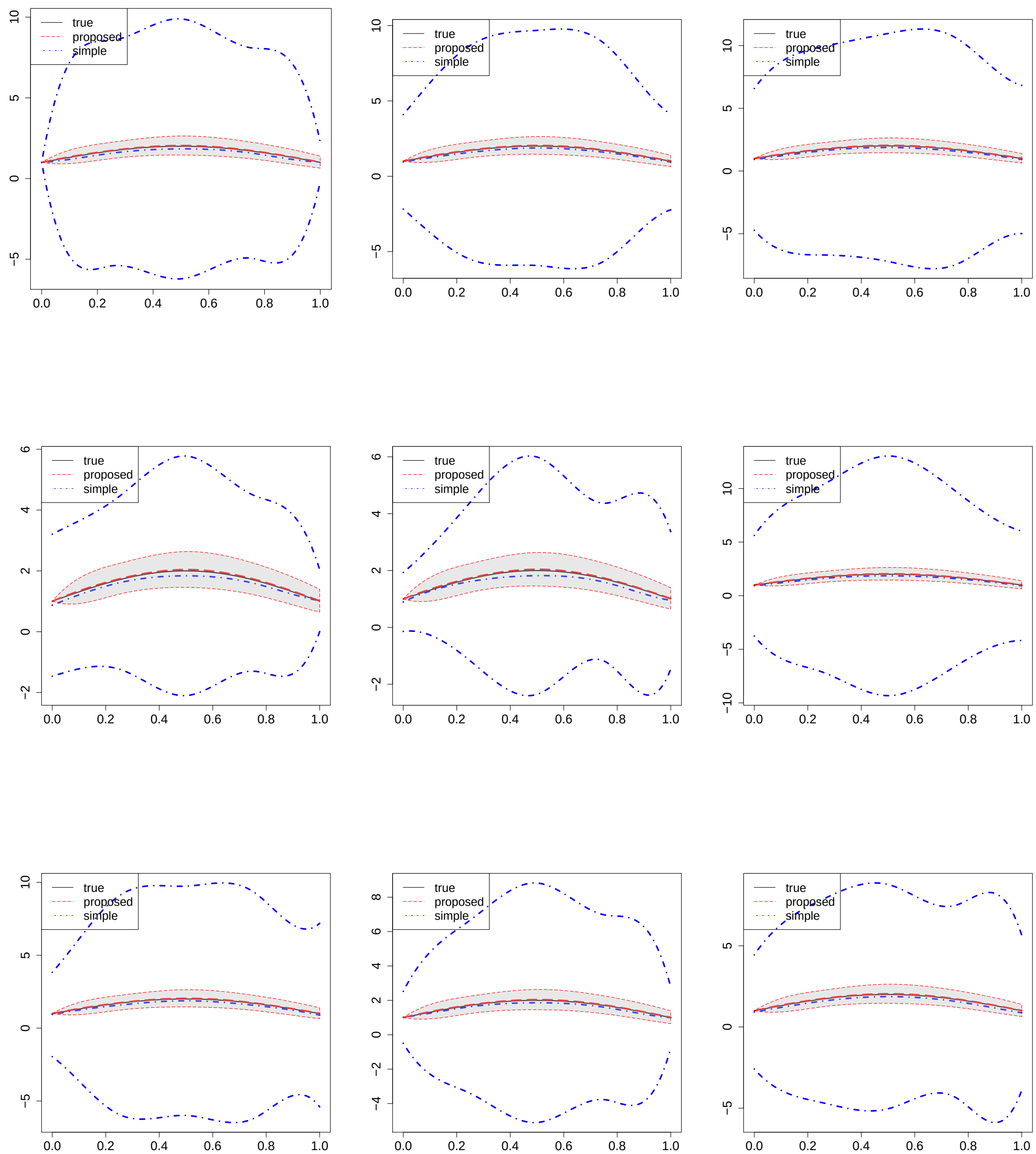

Figure S1: Simulation 2: Comparison of estimators of $\beta_k \alpha(t)$ based on (1) and $\beta_k(t)$ based on the model $f_{Y|\mathbf{X}(t)}\{Y, \mathbf{X}(t)\} = f\{Y, \int_0^1 \boldsymbol{\beta}(t)^{\mathrm{T}} \mathbf{X}(t) dt\}$ for $k = 1, \ldots, 9$.

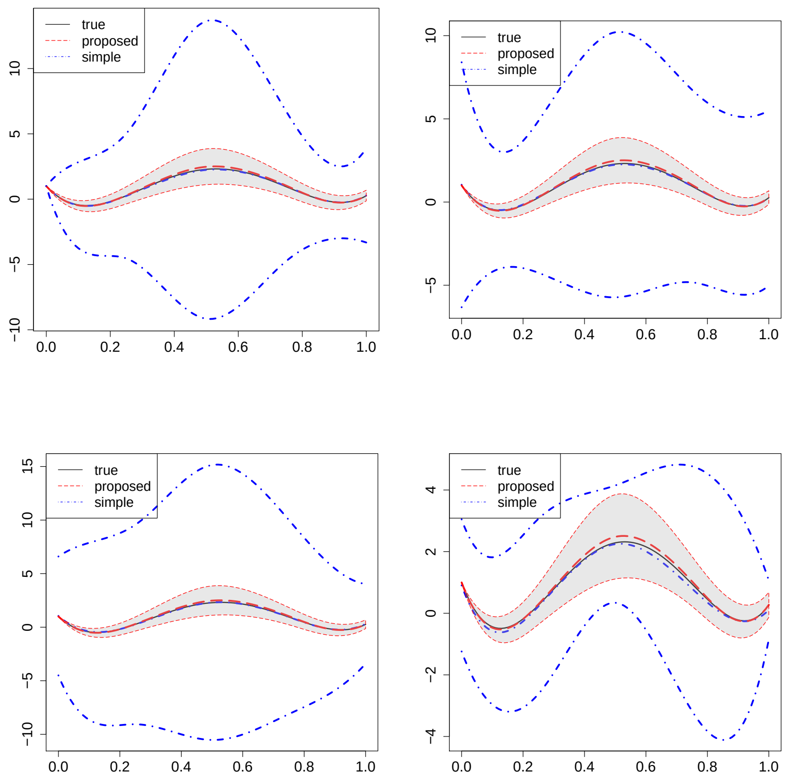

Figure S2: Simulation 3: Comparison of estimators of $\beta_k\alpha(t)$ based on (1) and $\beta_k(t)$ based on the model $f_{Y|\mathbf{X}(t)}\{Y,\mathbf{X}(t)\} = f\{Y,\int_0^1 \boldsymbol{\beta}(t)^{\mathrm{T}}\mathbf{X}(t)dt\}$ for $k=1,\ldots,4$.

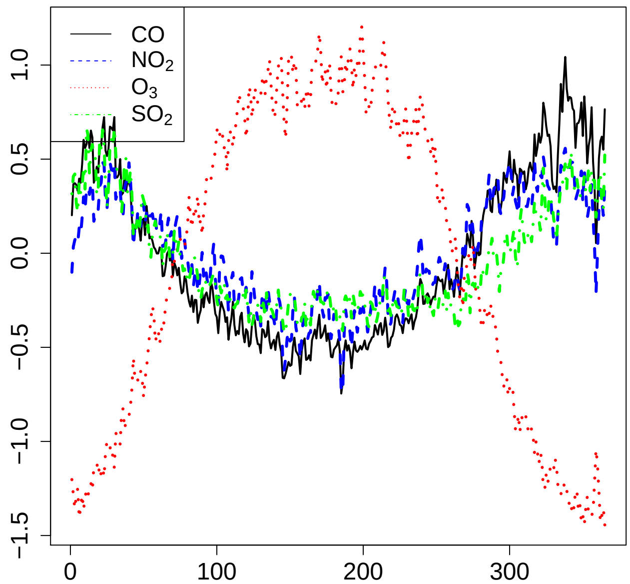

Figure S3: The point-wise mean trajectories of four air pollutants.

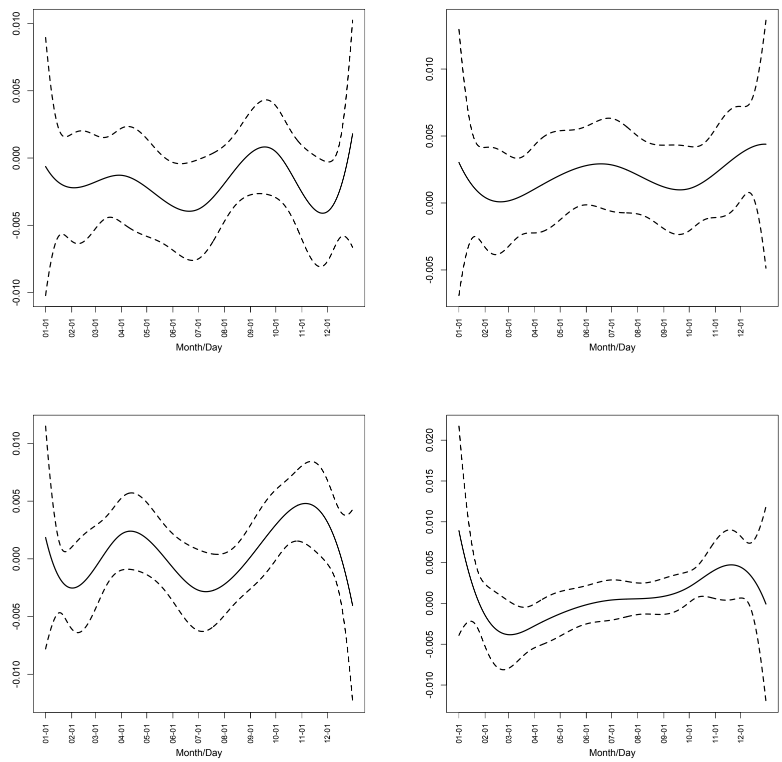

Figure S4: The estimated functional coefficient $\hat{\eta}(t)$ when fitting the functional linear model $E(Y) = \int_0^1 \eta(t)X(t)dt$ in the air pollution data, where the response variable Y is the annual CVD death rate, and the functional covariate $X(t)$ is the daily concentration of the air pollutant CO (the top left panel), NO_2 (the top right panel), SO_2 (the bottom left panel), and O_3 (the bottom right panel). The dashed lines are the pointwise 95% confidence intervals for $\eta(t)$.

References

Bosq, D. (1998), *Nonparametric statistics for stochastic processes : estimation and prediction D. Bosq.*, Lecture notes in statistics: 110, New York: Springer.

Cardot, H., Ferraty, F., and Sarda, P. (1999), "Functional linear model," *Statistics & Probability Letters*, 45, 11–22.

— (2003), "Spline estimators for the functional linear model," *Statistica Sinica*, 13, 571–592.

Chatelin, F. (1983), *Spectral Approximation of Linear Operators*, New York: SIAM.

De Boor, C. (1978), *A practical guide to splines*, vol. 27, New York: Springer-Verlag.

DeVore, R. A. and Lorentz, G. G. (1993), *Constructive approximation*, vol. 303, New York: Springer Science & Business Media.