Supplementary Materials

S1. The proof of theorems

This part contains the proofs of Theorem 1 and Theorem 2, which depend on two preliminary lemmas and first we will give the following Lemma 1, a direct result of de Boor (2001).

Lemma 1 There exists positive constants C, for any functions $g(z) \in \mathcal{H}(r)$ with $r < h-1$, there exists a function $g_0(z) \in \mathcal{S}_n^0$, such that

$$\sup_z |g(z) - g_0(z)| \leqslant CK_n^{-r}.$$

From Lemma 1, we can find $\boldsymbol{\zeta} = \{\zeta_{jk}, j = 1, \cdots, p_n, k = 1, \cdots, K_n\}^\top$, and an additive spline function $h(\boldsymbol{Z}) = \boldsymbol{B}^\top(\boldsymbol{Z})\boldsymbol{\zeta} \in \mathcal{G}_n$, such that

$$\sup_{\boldsymbol{Z}} |h(\boldsymbol{Z}) - h_0(\boldsymbol{Z})| \leqslant Cp_nK_n^{-r}. \tag{S1.1}$$

Denote $\tilde{\xi}_{ik} = \langle \hat{X}_i, \phi_k \rangle$, $m_0(\boldsymbol{\mathcal{F}}) = \int_0^1 X(t)\alpha_0(t)\mathrm{d}t + h_0(\boldsymbol{Z})$, $\boldsymbol{\mathcal{T}} = (\hat{\boldsymbol{\xi}}, \boldsymbol{Z})$ and $\tilde{m}_0(\boldsymbol{\mathcal{T}}) = \hat{\boldsymbol{\xi}}^\top\boldsymbol{\gamma}_0 + \boldsymbol{B}^\top(\boldsymbol{Z})\boldsymbol{\zeta}_0$, where $\boldsymbol{\gamma}_0$ and $\boldsymbol{\zeta}_0$ be the true parameter values of $\boldsymbol{\gamma}$ and $\boldsymbol{\zeta}$, and define

$$\boldsymbol{\Gamma}(\hat{\boldsymbol{\xi}}) = \frac{E[\boldsymbol{Z}\rho_2\{\tilde{m}_0(\boldsymbol{\mathcal{T}})\}|\hat{\boldsymbol{\xi}}]}{E[\rho_2\{\tilde{m}_0(\boldsymbol{\mathcal{T}})\}|\hat{\boldsymbol{\xi}}]}, \quad \tilde{\boldsymbol{B}}(\boldsymbol{Z}) = \boldsymbol{B}(\boldsymbol{Z}) - \boldsymbol{\Gamma}(\hat{\boldsymbol{\xi}}).$$

Lemma 2 Let $R_i = \tilde{m}_0(\boldsymbol{\mathcal{T}}_i) - m_0(\boldsymbol{\mathcal{F}}_i) = \hat{\boldsymbol{\xi}}_i^\top\boldsymbol{\gamma}_0 - \int_0^1 X_i(t)\alpha_0(t)\mathrm{d}t + \boldsymbol{B}^\top(\boldsymbol{Z}_i)\boldsymbol{\zeta}_0 - h_0(\boldsymbol{Z}_i), i = 1, 2, \ldots, n$. Suppose that conditions C1-C12 hold and the dimension of the

spline space $\mathcal{S}_n^0$ fulfils $K_n = O_p(n^{1/(1+2r)})$, then we have

$$\|R_i\|^2 = O_p\left(n^{-\frac{a+2b-1}{a+2b}}\right), \quad i = 1, 2, \ldots, n.$$

Proof of Lemma 2.

$$\begin{aligned}
\|R_i\|^2 &= \|\hat{\boldsymbol{\xi}}_i^\top \boldsymbol{\gamma}_0 - \int_0^1 X_i(t)\alpha_0(t)\mathrm{d}t + \boldsymbol{B}^\top(\boldsymbol{Z}_i)\boldsymbol{\zeta}_0 - h_0(\boldsymbol{Z}_i)\|^2 \\
&= \|\sum_{k=1}^{m_n} \hat{\xi}_{ik}\gamma_{0k} - \sum_{k=1}^{\infty} \xi_{ik}\gamma_{0k} + \boldsymbol{B}^\top(\boldsymbol{Z}_i)\boldsymbol{\zeta}_0 - h_0(\boldsymbol{Z}_i)\|^2 \\
&\leqslant 4\|\sum_{k=1}^{m_n}(\hat{\xi}_{ik} - \tilde{\xi}_{ik})\gamma_{0k}\|^2 + 4\|\sum_{k=1}^{m_n}(\tilde{\xi}_{ik} - \xi_{ik})\gamma_{0k}\|^2 \\
&\quad +4\|\sum_{k=m_n+1}^{\infty} \xi_{ik}\gamma_{0k}\|^2 + 4\|\boldsymbol{B}^\top(\boldsymbol{Z}_i)\boldsymbol{\zeta}_0 - h_0(\boldsymbol{Z}_i)\|^2,
\end{aligned}$$

By (S1.1), the last term on the right hand side of the above is $\|\boldsymbol{B}^\top(\boldsymbol{Z}_i)\boldsymbol{\zeta}_0 - h_0(\boldsymbol{Z}_i)\|^2 = O_p(p_n^2 K_n^{-2r}) = o(1)$. Similar to the derivation of Lemma 2 in Liu et al. (2024), we get

$$\|R_i\|^2 = O_p\left(n^{-\frac{a+2b-1}{a+2b}}\right).$$

□

Proof of Theorem 1 . Denote $\boldsymbol{\zeta} = \boldsymbol{\zeta}_0 + \delta_n \boldsymbol{T}_1, \boldsymbol{\gamma} = \boldsymbol{\gamma}_0 + \delta_n \boldsymbol{T}_2$ and $\boldsymbol{T} = (\boldsymbol{T}_1^\top, \boldsymbol{T}_2^\top)^\top$. We first prove the following inequality, that is, for any given $\varepsilon > 0$, there exists a large constant C such that

$$P\left\{\sup_{\|\boldsymbol{T}\|=C} PL(\boldsymbol{\zeta}_0 + \delta_n \boldsymbol{T}_1, \boldsymbol{\gamma}_0 + \delta_n \boldsymbol{T}_2) < PL\left(\boldsymbol{\zeta}_0, \boldsymbol{\gamma}_0\right)\right\} \geq 1 - \varepsilon. \tag{S1.2}$$

It can be seen that, there is local maximum $(\hat{\boldsymbol{\zeta}}, \hat{\boldsymbol{\gamma}})$ in the ball $\{(\boldsymbol{\zeta}_0 + \delta_n \boldsymbol{T}_1, \boldsymbol{\gamma}_0 + \delta_n \boldsymbol{T}_2) : \|(\boldsymbol{T}_1^\top, \boldsymbol{T}_2^\top)^\top\| \leq C\}$ with probability tending to 1. Then we have

$$\|\hat{\boldsymbol{\zeta}} - \boldsymbol{\zeta}_0\| = O_p(\delta_n), \|\hat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0\| = O_p(\delta_n). \tag{S1.3}$$

Denote $\boldsymbol{U}_i = \{\boldsymbol{B}^\top(\boldsymbol{Z}_i), \hat{\boldsymbol{\xi}}_i^\top\}^\top$. Then, since $p_{\lambda_n}(0) = 0$ and $p_{\lambda_n}(\|\boldsymbol{\zeta}\|) \geq 0$ for all $\boldsymbol{\zeta}$, we find that

$$
\begin{aligned}
& PL(\boldsymbol{\zeta}_0 + \delta_n \boldsymbol{T_1}, \boldsymbol{\gamma}_0 + \delta_n \boldsymbol{T_2}) - PL\left(\boldsymbol{\zeta}_0, \boldsymbol{\gamma}_0\right) \\
\leq\ & L(\boldsymbol{\zeta}_0 + \delta_n \boldsymbol{T}_1, \boldsymbol{\gamma}_0 + \delta_n \boldsymbol{T}_2) - L\left(\boldsymbol{\zeta}_0, \boldsymbol{\gamma}_0\right) - n \sum_{j=1}^{s} \{p_{\lambda_n}(\|\boldsymbol{\zeta}_{0j} + \delta_n \boldsymbol{T}_{1j}\|) - p_{\lambda_n}(\|\boldsymbol{\zeta}_{0j}\|)\} \\
=\ & \sum_{i=1}^{n} [Q\{g^{-1}(\tilde{m}_0(\boldsymbol{\mathcal{T}}_i) + \delta_n(\boldsymbol{B}^\top(\boldsymbol{Z}_i)\boldsymbol{T_1} + \hat{\boldsymbol{\xi}}_i^\top \boldsymbol{T_2})), Y_i\} - Q\{g^{-1}(\tilde{m}_0(\boldsymbol{\mathcal{T}}_i)), Y_i\}] \\
& - n \sum_{j=1}^{s} \{p_{\lambda_n}(\|\boldsymbol{\zeta}_{0j} + \delta_n \boldsymbol{T}_{1j}\|) - p_{\lambda_n}(\|\boldsymbol{\zeta}_{0j}\|)\} \\
\leq\ & \sum_{i=1}^{n} [Q\{g^{-1}(\tilde{m}_0(\boldsymbol{\mathcal{T}}_i) + \delta_n \boldsymbol{T}^\top \boldsymbol{U}_i), Y_i\} - Q\{g^{-1}(\tilde{m}_0(\boldsymbol{\mathcal{T}}_i)), Y_i\}] \\
& - n \sum_{j=1}^{s} [\delta_n p'_{\lambda_n}(\|\boldsymbol{\zeta}_{0j}\|) \frac{\boldsymbol{\zeta}_{0j}^\top}{\|\boldsymbol{\zeta}_{0j}\|} \boldsymbol{T}_{1j} + \frac{1}{2} \delta_n^2 p''_{\lambda_n}(\|\boldsymbol{\zeta}_{0j}\|) \boldsymbol{T}_{1j}^\top \boldsymbol{T}_{1j} \{1 + o_p(1)\}],
\end{aligned}
$$

where s is the number of nonzero additive components. Using standard Taylor expansion argument, the first term

$$
\begin{aligned}
& \sum_{i=1}^{n} [Q\{g^{-1}(\tilde{m}_0(\boldsymbol{\mathcal{T}}_i) + \delta_n \boldsymbol{T}^\top \boldsymbol{U}_i), Y_i\} - Q\{g^{-1}(\tilde{m}_0(\boldsymbol{\mathcal{T}}_i)), Y_i\}] \\
=\ & \sum_{i=1}^{n} q_1\{\tilde{m}_0(\boldsymbol{\mathcal{T}}_i), Y_i\} \delta_n \boldsymbol{T}^\top \boldsymbol{U}_i + \frac{n}{2} \delta_n^2 \boldsymbol{T}^\top \left[\frac{1}{n} \sum_{i=1}^{n} q_2\{\tilde{m}_0(\boldsymbol{\mathcal{T}}_i) + \varsigma_i, Y_i\} \boldsymbol{U}_i \boldsymbol{U}_i^\top\right] \boldsymbol{T} \\
=\ & \sum_{i=1}^{n} q_1\{\tilde{m}_0(\boldsymbol{\mathcal{T}}_i), Y_i\} \delta_n \boldsymbol{T}^\top \boldsymbol{U}_i \\
& + \frac{n}{2} \delta_n^2 \boldsymbol{T}^\top \left[\frac{1}{n} \sum_{i=1}^{n} \left\{(Y_i - g^{-1}(\tilde{m}_0(\boldsymbol{\mathcal{T}}_i) + \varsigma_i)) \rho'_1(\tilde{m}_0(\boldsymbol{\mathcal{T}}_i) + \varsigma_i) - \rho_2(\tilde{m}_0(\boldsymbol{\mathcal{T}}_i) + \varsigma_i)\right\} \boldsymbol{U}_i \boldsymbol{U}_i^\top\right] \boldsymbol{T} \\
=\ & \delta_n \boldsymbol{T}^\top \sum_{i=1}^{n} q_1\{\tilde{m}_0(\boldsymbol{\mathcal{T}}_i), Y_i\} \boldsymbol{U}_i - \frac{n}{2} \delta_n^2 \boldsymbol{T}^\top E[\rho_2\{m_0(\boldsymbol{\mathcal{F}})\} \boldsymbol{U} \boldsymbol{U}^\top] \boldsymbol{T} + o_p(1),
\end{aligned}
$$

where ς_i is independent of Y_i, for $i = 1, \cdots, n$, and ς_i is between 0 and $\delta_n \boldsymbol{T}^\top \boldsymbol{U}_i$. Using

arguments similar to those in Liu et al. (2024), we can deduce further that

$$\begin{aligned} & PL(\boldsymbol{\zeta}_0 + \delta_n \boldsymbol{T_1}, \boldsymbol{\gamma}_0 + \delta_n \boldsymbol{T_2}) - PL\left(\boldsymbol{\zeta}_0, \boldsymbol{\gamma}_0\right) \\ \leq \; & \delta_n \boldsymbol{T}^\top \sum_{i=1}^{n} q_1\{m_0(\boldsymbol{\mathcal{F}}_i), Y_i\}\boldsymbol{U}_i - \frac{n}{2}\delta_n^2 \boldsymbol{T}^\top E[\rho_2\{m_0(\boldsymbol{\mathcal{F}})\}\boldsymbol{U}\boldsymbol{U}^\top]\boldsymbol{T} \\ & -n\sum_{j=1}^{s}[\delta_n p'_{\lambda_n}(\|\boldsymbol{\zeta}_{0j}\|)\frac{\boldsymbol{\zeta}_{0j}^\top}{\|\boldsymbol{\zeta}_{0j}\|}\boldsymbol{T}_{1j} + \frac{1}{2}\delta_n^2 p''_{\lambda_n}(\|\boldsymbol{\zeta}_{0j}\|)\boldsymbol{T}_{1j}^\top \boldsymbol{T}_{1j}\{1 + o_p(1)\}], \end{aligned}$$

where the orders of the first and second term are $O_p\{n^{1/2}\delta_n(m_n+K_n)\}$ and $O_p\{n\delta_n^2(m_n+K_n)^2\}$, respectively. For enough large C, we can deduce that the second term dominates in magnitude in $\|\boldsymbol{T}\| = C$. The third and fourth terms are bounded by $\sqrt{s}n\delta_n a_n\|\boldsymbol{T}_1\| + n\delta_n^2 b_n\|\boldsymbol{T}_1\|^2 = Cn\delta_n^2(\sqrt{s} + b_n C)$. As $b_n \to 0$, this is also dominated by the second term in $\|\boldsymbol{T}_1\| = C$. Hence, (S1.2) holds by letting C large enough.

For estimator $\hat{h}(\boldsymbol{Z})$, from (S1.3), it is sufficient to prove that

$$\begin{aligned} \| \hat{h}(\boldsymbol{Z}) - h_0(\boldsymbol{Z}) \|^2 &= \left\|(\hat{\boldsymbol{\zeta}} - \boldsymbol{\zeta}_0)^\top \boldsymbol{B}\right\|^2 \\ &= (\hat{\boldsymbol{\zeta}} - \boldsymbol{\zeta}_0)^\top E\{n^{-1}\sum_{i=1}^{n}\mathbf{B}(\boldsymbol{Z}_i)\mathbf{B}^\top(\boldsymbol{Z}_i)\}(\hat{\boldsymbol{\zeta}} - \boldsymbol{\zeta}_0) \\ &\leq C\|\hat{\boldsymbol{\zeta}} - \boldsymbol{\zeta}_0\|^2 \\ &= O_p(\delta_n^2). \end{aligned}$$

Since $\|\hat{\phi}_j - \phi_j\|^2 = O_p(n^{-1}j^2)$ and (S1.3), using arguments similar to those in Liu et al.

(2024), we can see that

$$\begin{aligned}\|\hat{\alpha}(t)-\alpha_0(t)\|^2 &= \|\sum_{j=1}^{m_n}\hat{\gamma}_j\hat{\phi}_j(t)-\sum_{j=1}^{\infty}\gamma_{0j}\phi_j(t)\|^2 \\ &\leq 4\|\hat{\boldsymbol{\gamma}}-\boldsymbol{\gamma}_0\|^2+4\|\sum_{j=1}^{m_n}\gamma_{0j}\{\hat{\phi}_j(t)-\phi_j(t)\}\|^2+2\sum_{j=m_n+1}^{\infty}{\gamma_{0j}}^2 \\ &= O_p\left(\delta_n^2\right).\end{aligned}$$

This completes the proof of Theorem 1. □

Proof of Theorem 2 . We use proof by contradiction. Suppose that there exists a k_0, $s+1 \leq k_0 \leq p_n$ such that $P\{h_{k_0}(z)=0\} \nrightarrow 0$. Then, there exists a $\eta > 0$ such that, $P(\hat{\boldsymbol{\zeta}}_{k_0} \neq \mathbf{0}) = P\{\hat{h}_{k_0}(z) \neq 0\} \geq \eta$. Denote $\hat{\boldsymbol{\zeta}}^* = (\hat{\boldsymbol{\zeta}}_1^\top, \cdots, \hat{\boldsymbol{\zeta}}_s^\top, \mathbf{0}^\top, \cdots, \mathbf{0}^\top)^\top$. It will be shown that there exists a $\delta > 0$ such that $PL(\hat{\boldsymbol{\zeta}}, \hat{\boldsymbol{\gamma}}) - PL(\hat{\boldsymbol{\zeta}}^*, \hat{\boldsymbol{\gamma}}) < 0$ with probability at least δ, which contradicts with the fact that $PL(\hat{\boldsymbol{\zeta}}, \hat{\boldsymbol{\gamma}}) - PL(\hat{\boldsymbol{\zeta}}^*, \hat{\boldsymbol{\gamma}}) \geq 0$.

By Theorem 1 , we have $\|\hat{\boldsymbol{\zeta}}_k - \boldsymbol{\zeta}_k\| = O_p\{\sqrt{(m_n+K_n)n^{-(\frac{2r}{1+2r}-\frac{2b-1}{a+2b})}}\}$. Since $\boldsymbol{\zeta}_k = \mathbf{0}$ for $k = s+1, \cdots, p_n$, we have $\|\hat{\boldsymbol{\zeta}}_k\| = O_p\{\sqrt{(m_n+K_n)n^{-(\frac{2r}{1+2r}-\frac{2b-1}{a+2b})}}\}$ for $k = s+1, \cdots, p_n$. So $\|\hat{\boldsymbol{\zeta}}_{k_0}\| = O_p\{\sqrt{(m_n+K_n)n^{-(\frac{2r}{1+2r}-\frac{2b-1}{a+2b})}}\}$. With probability tending to one, $\|\hat{\boldsymbol{\zeta}}_{k_0}\| \leq \lambda_n$, since $\lambda_n(n^{-(\frac{2r}{1+2r}-\frac{2b-1}{a+2b})})^{-\frac{1}{2}}/\sqrt{m_n+J_n} \to \infty$. By the definition of $p_{\lambda_n}(\cdot)$, we have $P\{p_{\lambda_n}(\|\hat{\boldsymbol{\zeta}}_{k_0}\|) = \lambda_n\|\hat{\boldsymbol{\zeta}}_{k_0}\|\} \to 1$. Mimicking the proof for Theorem 1

indicates that

$$
\begin{aligned}
& PL(\hat{\boldsymbol{\zeta}}, \hat{\boldsymbol{\gamma}}) - PL(\hat{\boldsymbol{\zeta}}^*, \hat{\boldsymbol{\gamma}}) \\
= & \sum_{i=1}^{n} \Big[Q\{g^{-1}(\boldsymbol{B}^\top(\boldsymbol{Z}_i)\hat{\boldsymbol{\zeta}} + \hat{\boldsymbol{\xi}}_i^\top \hat{\boldsymbol{\gamma}}), Y_i\} - Q\{g^{-1}(\boldsymbol{B}^\top(\boldsymbol{Z}_i)\hat{\boldsymbol{\zeta}}^* + \hat{\boldsymbol{\xi}}_i^\top \hat{\boldsymbol{\gamma}}), Y_i\} \Big] \\
& -n \sum_{k=1}^{s} \{ p_{\lambda_n}(\|\hat{\boldsymbol{\zeta}}_k\|) - p_{\lambda_n}(\|\hat{\boldsymbol{\zeta}}_k^*\|) \} \\
= & \sum_{i=1}^{n} q_1\{m_0(\boldsymbol{\mathcal{F}}_i), Y_i\}(\hat{\boldsymbol{\zeta}} - \hat{\boldsymbol{\zeta}}^*) \\
& -\frac{n}{2}(\hat{\boldsymbol{\zeta}} - \hat{\boldsymbol{\zeta}}^*)^\top E[\rho_2\{m_0(\boldsymbol{\mathcal{F}})\}](\hat{\boldsymbol{\zeta}} - \hat{\boldsymbol{\zeta}}^*) - n\lambda_n \|\hat{\boldsymbol{\zeta}}_{k_0}\| + o_p(1) \\
= & O_p(n^{1/2}\|\hat{\boldsymbol{\zeta}}_{k_0}\|) + O_p(n\|\hat{\boldsymbol{\zeta}}_{k_0}\|^2) - n\lambda_n \|\hat{\boldsymbol{\zeta}}_{k_0}\| \\
= & n\lambda_n^2 \{ -{\lambda_n}^{-1}\|\hat{\boldsymbol{\zeta}}_{k_0}\| + O_p(\frac{\|\hat{\boldsymbol{\zeta}}_{k_0}\|^2}{\lambda_n^2}) \} \\
= & n\lambda_n^2 [-{\lambda_n}^{-1}\|\hat{\boldsymbol{\zeta}}_{k_0}\| + O_p\{ \frac{(m_n + K_n)(n^{-(\frac{2r}{1+2r} - \frac{2b-1}{a+2b})})}{\lambda_n^2} \}]. \qquad \text{(S1.4)}
\end{aligned}
$$

By the fact that $\lambda_n(n^{-(\frac{2r}{1+2r} - \frac{2b-1}{a+2b})})^{-\frac{1}{2}}/\sqrt{m_n + K_n} \to \infty$, we can conclude that (S1.4) is dominated by $\lambda_n^{-1}\|\hat{\boldsymbol{\zeta}}_{k_0}\|$, so $PL(\hat{\boldsymbol{\zeta}}, \hat{\boldsymbol{\gamma}}) - PL(\hat{\boldsymbol{\zeta}}^*, \hat{\boldsymbol{\gamma}}) < 0$, which contradicts to $PL(\hat{\boldsymbol{\zeta}}, \hat{\boldsymbol{\gamma}}) - PL(\hat{\boldsymbol{\zeta}}^*, \hat{\boldsymbol{\gamma}}) \geq 0$. The proof of the theorem 2 (a) is complete.

Now, for part (b) of Theorem 2, we prove the asymptotic normality of the estimate $\hat{\boldsymbol{\zeta}}_{\mathcal{S}}$. By Theorem 2 (a), we know that $\hat{\boldsymbol{\zeta}}_{\mathcal{N}} = \mathbf{0}$ with probability tending to 1. Denote $\hat{\boldsymbol{\theta}} = \delta_n^{-1}(\hat{\boldsymbol{\zeta}}_{\mathcal{S}} - \boldsymbol{\zeta}_{0\mathcal{S}})$ and $\tilde{m}_0(\boldsymbol{\mathcal{T}}_{i\mathcal{S}}) = \boldsymbol{B}^\top(\boldsymbol{Z}_{i\mathcal{S}})\boldsymbol{\zeta}_{0\mathcal{S}} + \hat{\boldsymbol{\xi}}_i^\top \boldsymbol{\gamma}_0$. For any $\boldsymbol{v} \in R^s$, define $\tilde{m}_{\boldsymbol{v}} = \tilde{m}(\boldsymbol{\mathcal{T}}_{\mathcal{S}}) + \boldsymbol{v}^\top \tilde{\boldsymbol{B}}(\boldsymbol{Z}_{\mathcal{S}}) = \hat{\boldsymbol{\xi}}^\top \boldsymbol{\gamma} - \boldsymbol{v}^\top \boldsymbol{\Gamma}_1(\hat{\boldsymbol{\xi}}) + (\hat{\boldsymbol{\zeta}}_{\mathcal{S}} + \boldsymbol{v})^\top \boldsymbol{B}(\boldsymbol{Z}_{\mathcal{S}})$. Note that $\tilde{m}_{\boldsymbol{v}}$ maximizes the function $l_n(\tilde{m}) = \frac{1}{n}\sum_{i=1}^{n} Q[g^{-1}\{\tilde{m}(\boldsymbol{\mathcal{T}}_{i\mathcal{S}})\}, Y_i] - \sum_{j=1}^{s} p_{\lambda_n}(\|\hat{\boldsymbol{\zeta}}_j + v_j\|)$ when $\boldsymbol{v} = \mathbf{0}$,

thus we have

$$\begin{aligned}
\mathbf{0} &= \frac{\partial}{\partial \boldsymbol{v}} l_n(\tilde{m}_{\boldsymbol{v}})|_{\boldsymbol{v}=\mathbf{0}} \\
&= \frac{1}{n}\sum_{i=1}^{n}\left[q_1\left\{\tilde{m}(\boldsymbol{\mathcal{T}}_{i\mathcal{S}}), Y_i\right\}\tilde{\boldsymbol{B}}(\boldsymbol{Z}_{i\mathcal{S}})\right] \\
&\quad - \left\{p'_{\lambda_n}\left(\left\|\hat{\boldsymbol{\zeta}}_j\right\|\right)\frac{\hat{\boldsymbol{\zeta}}_j^{\top}}{\left\|\hat{\boldsymbol{\zeta}}_j\right\|}\right\}_{j=1}^{s} + o_p(1). \qquad \text{(S1.5)}
\end{aligned}$$

For the first term in (S1.5), we get

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n}\left[q_1\left\{\tilde{m}(\boldsymbol{\mathcal{T}}_{i\mathcal{S}}), Y_i\right\}\tilde{\boldsymbol{B}}(\boldsymbol{Z}_{i\mathcal{S}})\right] \\
= &\frac{1}{n}\sum_{i=1}^{n}\left[q_1\left\{\tilde{m}_0(\boldsymbol{\mathcal{T}}_{i\mathcal{S}}), Y_i\right\}\tilde{\boldsymbol{B}}(\boldsymbol{Z}_{i\mathcal{S}})\right] \\
&+\frac{1}{n}\sum_{i=1}^{n}\left[q_2\left\{\tilde{m}_0(\boldsymbol{\mathcal{T}}_{i\mathcal{S}}), Y_i\right\}\left\{\tilde{m}(\boldsymbol{\mathcal{T}}_{i\mathcal{S}}) - \tilde{m}_0(\boldsymbol{\mathcal{T}}_{i\mathcal{S}})\right\}\tilde{\boldsymbol{B}}(\boldsymbol{Z}_{i\mathcal{S}})\right] \\
&+\frac{1}{n}\sum_{i=1}^{n}\left[q_2'\{\bar{m}(\boldsymbol{\mathcal{T}}_{i\mathcal{S}}), Y_i\}\{\tilde{m}(\boldsymbol{\mathcal{T}}_{i\mathcal{S}}) - \tilde{m}_0(\boldsymbol{\mathcal{T}}_{i\mathcal{S}})\}^2\tilde{\boldsymbol{B}}(\boldsymbol{Z}_{i\mathcal{S}})\right] \\
= &\ I_1 + I_2 + I_3. \qquad \text{(S1.6)}
\end{aligned}$$

We decompose I_2 into two terms I_{21} and I_{22} as follows

$$\begin{aligned}
I_2 &= \frac{1}{n}\sum_{i=1}^{n} q_2\left\{\tilde{m}_0(\boldsymbol{\mathcal{T}}_{i\mathcal{S}}), Y_i\right\}\tilde{\boldsymbol{B}}(\boldsymbol{Z}_{i\mathcal{S}})\hat{\boldsymbol{\xi}}_i^{\top}(\hat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0) \\
&\quad +\frac{1}{n}\sum_{i=1}^{n} q_2\left\{\tilde{m}_0(\boldsymbol{\mathcal{T}}_{i\mathcal{S}}), Y_i\right\}\tilde{\boldsymbol{B}}(\boldsymbol{Z}_{i\mathcal{S}})\boldsymbol{B}^{\top}(\boldsymbol{Z}_{i\mathcal{S}})(\hat{\boldsymbol{\zeta}}_{\mathcal{S}} - \boldsymbol{\zeta}_{0\mathcal{S}}) \\
&= I_{21} + I_{22}.
\end{aligned}$$

Mimicking the proof for Theorem 1 indicates that

$$\begin{aligned} I_{21} &= -\frac{1}{n}\sum_{i=1}^{n}\rho_2\left\{\tilde{m}_0(\boldsymbol{\mathcal{T}}_{i\mathcal{S}})\right\}\tilde{\boldsymbol{B}}(\boldsymbol{Z}_{i\mathcal{S}})\hat{\boldsymbol{\xi}}_i^{\top}(\hat{\boldsymbol{\gamma}}-\boldsymbol{\gamma}_0)+o_p(1) \\ &= -E[\rho_2\left\{\tilde{m}_0(\boldsymbol{\mathcal{T}}_{\mathcal{S}})\right\}\tilde{\boldsymbol{B}}(\boldsymbol{Z}_{\mathcal{S}})\hat{\boldsymbol{\xi}}^{\top}(\hat{\boldsymbol{\gamma}}-\boldsymbol{\gamma}_0)]+o_p(1). \end{aligned}$$

By the definition of $\tilde{\boldsymbol{B}}(\boldsymbol{Z}_{\mathcal{S}})$, for any measurable function ϕ, $E[\phi(\hat{\boldsymbol{\xi}})\rho_2\left\{\tilde{m}_0(\boldsymbol{\mathcal{T}}_{\mathcal{S}})\right\}\tilde{\boldsymbol{B}}(\boldsymbol{Z}_{\mathcal{S}})] = \mathbf{0}$. Hence $I_{21} = o_p(1)$. Using similar arguments, we can show that

$$\begin{aligned} I_{22} &= -\frac{1}{n}\sum_{i=1}^{n}\rho_2\left\{\tilde{m}_0(\boldsymbol{\mathcal{T}}_{i\mathcal{S}}),Y_i\right\}\tilde{\boldsymbol{B}}(\boldsymbol{Z}_{i\mathcal{S}})\boldsymbol{B}^{\top}(\boldsymbol{Z}_{i\mathcal{S}})(\hat{\boldsymbol{\zeta}}_{\mathcal{S}}-\boldsymbol{\zeta}_{0\mathcal{S}})+o_p(1) \\ &= -\frac{1}{n}\sum_{i=1}^{n}\rho_2\left\{\tilde{m}_0(\boldsymbol{\mathcal{T}}_{i\mathcal{S}}),Y_i\right\}\tilde{\boldsymbol{B}}(\boldsymbol{Z}_{i\mathcal{S}})\tilde{\boldsymbol{B}}^{\top}(\boldsymbol{Z}_{i\mathcal{S}})(\hat{\boldsymbol{\zeta}}_{\mathcal{S}}-\boldsymbol{\zeta}_{0\mathcal{S}})+o_p(1). \end{aligned}$$

According to (S1.2) and condition (C6), we have

$$\begin{aligned} I_3 &= \frac{1}{n}\sum_{i=1}^{n}\left[q_2'\{\bar{m}(\boldsymbol{\mathcal{T}}_{i\mathcal{S}}),Y_i\}\{\tilde{m}(\boldsymbol{\mathcal{T}}_{i\mathcal{S}})-\tilde{m}_0(\boldsymbol{\mathcal{T}}_{i\mathcal{S}})\}^2\tilde{\boldsymbol{B}}(\boldsymbol{Z}_{i\mathcal{S}})\right] \\ &\leq C\parallel\tilde{m}(\boldsymbol{\mathcal{T}}_{\mathcal{S}})-\tilde{m}_0(\boldsymbol{\mathcal{T}}_{\mathcal{S}})\parallel^2 = O_p(\delta_n^2). \end{aligned}$$

For the second term in (S1.5), we get

$$\begin{aligned} &\left\{p_{\lambda_n}'\left(\left\|\hat{\boldsymbol{\zeta}}_j\right\|\right)\frac{\hat{\boldsymbol{\zeta}}_j^{\top}}{\left\|\hat{\boldsymbol{\zeta}}_j\right\|}\right\}_{j=1}^{s} \\ &= \left\{p_{\lambda_n}'\left(\|\boldsymbol{\zeta}_{0j}\|\right)\frac{\boldsymbol{\zeta}_{0j}^{\top}}{\|\boldsymbol{\zeta}_{0j}\|}\right\}_{j=1}^{s} + \{\sum_{j=1}^{s}p_{\lambda_n}''(\|\boldsymbol{\zeta}_{0j}\|)+o_p(1)\}(\hat{\boldsymbol{\zeta}}_j-\boldsymbol{\zeta}_{0j}). \end{aligned} \tag{S1.7}$$

From (S1.5)-(S1.7), we obtain that

$$\begin{aligned} \mathbf{0} &= \frac{1}{n}\sum_{i=1}^{n}\left[q_1\left\{\tilde{m}_0(\boldsymbol{\mathcal{T}}_{i\mathcal{S}}),Y_i\right\}\tilde{\boldsymbol{B}}(\boldsymbol{Z}_{i\mathcal{S}})\right]-\left\{p_{\lambda_n}'\left(\|\boldsymbol{\zeta}_{0j}\|\right)\frac{\boldsymbol{\zeta}_{0j}^{\top}}{\|\boldsymbol{\zeta}_{0j}\|}\right\}_{j=1}^{s} \\ &\quad -\left[E[\rho_2\{\tilde{m}_0(\boldsymbol{\mathcal{T}}_{\mathcal{S}})\}\tilde{\boldsymbol{B}}(\boldsymbol{Z}_{\mathcal{S}})\tilde{\boldsymbol{B}}^{\top}(\boldsymbol{Z}_{\mathcal{S}})]+o_p(1)\right](\hat{\boldsymbol{\zeta}}_{\mathcal{S}}-\boldsymbol{\zeta}_{0\mathcal{S}}) \\ &\quad -\{\sum_{j=1}^{s}p_{\lambda_n}''(\|\boldsymbol{\zeta}_{0j}\|)+o_p(1)\}(\hat{\boldsymbol{\zeta}}_j-\boldsymbol{\zeta}_{0j})+O_p(\delta_n^2). \end{aligned}$$

Therefore,

$$\begin{aligned}\mathbf{0} &= \frac{1}{n}\sum_{i=1}^{n}\left[q_1\left\{\tilde{m}_0(\boldsymbol{\mathcal{T}}_{i\mathcal{S}}), Y_i\right\}\tilde{\boldsymbol{B}}(\boldsymbol{Z}_{i\mathcal{S}})\right] - \boldsymbol{c}_n \\ &\quad -\{\boldsymbol{\Sigma}_{\lambda_n} + \Omega_{\mathcal{S}} + o_p(1)\}(\hat{\boldsymbol{\zeta}}_{\mathcal{S}} - \boldsymbol{\zeta}_{0\mathcal{S}}) + O_p(\delta_n^2).\end{aligned}$$

Note that $q_1\left\{\tilde{m}_0(\boldsymbol{\mathcal{T}}_{i\mathcal{S}}), Y_i\right\} = [Y - g^{-1}\{\tilde{m}_0(\boldsymbol{\mathcal{T}}_{i\mathcal{S}})\}]\rho_1\{\tilde{m}_0(\boldsymbol{\mathcal{T}}_{i\mathcal{S}})\}$ and

$$\begin{aligned}&E\left[\rho_1^2\{\tilde{m}_0(\boldsymbol{\mathcal{T}}_{\mathcal{S}})\}[Y - g^{-1}\{\tilde{m}_0(\boldsymbol{\mathcal{T}}_{\mathcal{S}})\}]^2\tilde{\boldsymbol{B}}(\boldsymbol{Z}_{\mathcal{S}})\tilde{\boldsymbol{B}}^{\top}(\boldsymbol{Z}_{\mathcal{S}})\right] \\ &= E\left[E[\{Y - g^{-1}(\tilde{m}_0(\boldsymbol{\mathcal{T}}_{\mathcal{S}}))\}^2|\boldsymbol{\mathcal{T}}_{\mathcal{S}}]\rho_1^2\{\tilde{m}_0(\boldsymbol{\mathcal{T}}_{\mathcal{S}})\}\tilde{\boldsymbol{B}}(\boldsymbol{Z}_{\mathcal{S}})\tilde{\boldsymbol{B}}^{\top}(\boldsymbol{Z}_{\mathcal{S}})\right] \\ &= E[\rho_2\{\tilde{m}_0(\boldsymbol{\mathcal{T}}_{\mathcal{S}})\}\tilde{\boldsymbol{B}}(\boldsymbol{Z}_{\mathcal{S}})\tilde{\boldsymbol{B}}^{\top}(\boldsymbol{Z}_{\mathcal{S}})] = E(\Omega_{\mathcal{S}}),\end{aligned} \tag{S1.8}$$

thus

$$\begin{aligned}&\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left[q_1\left\{\tilde{m}_0(\boldsymbol{\mathcal{T}}_{i\mathcal{S}}), Y_i\right\}\tilde{\boldsymbol{B}}(\boldsymbol{Z}_{i\mathcal{S}})\right] \\ = \;& \sqrt{n}(\boldsymbol{\Sigma}_{\lambda_n} + \Omega_{\mathcal{S}})(\hat{\boldsymbol{\zeta}}_{\mathcal{S}} - \boldsymbol{\zeta}_{0\mathcal{S}}) + \sqrt{n}\boldsymbol{c}_n + o_p(1).\end{aligned}$$

Following a Lindeberg-Feller argument, we have

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left[q_1\left\{\tilde{m}_0(\boldsymbol{\mathcal{T}}_{i\mathcal{S}}), Y_i\right\}\tilde{\boldsymbol{B}}(\boldsymbol{Z}_{i\mathcal{S}})\right] \xrightarrow{d} N(0, E[q_1^2\left\{\tilde{m}_0(\boldsymbol{\mathcal{T}}_{\mathcal{S}}), Y\right\}\tilde{\boldsymbol{B}}(\boldsymbol{Z}_{\mathcal{S}})\tilde{\boldsymbol{B}}^{\top}(\boldsymbol{Z}_{\mathcal{S}})]).$$

According to (S1.8), we have

$$\sqrt{n}(\boldsymbol{\Sigma}_{\lambda_n} + \Omega_{\mathcal{S}})\{\hat{\boldsymbol{\zeta}}_{\mathcal{S}} - \boldsymbol{\zeta}_{0\mathcal{S}} + (\boldsymbol{\Sigma}_{\lambda_n} + \Omega_{\mathcal{S}})^{-1}\boldsymbol{c}_n\} \xrightarrow{d} N(0, E(\Omega_{\mathcal{S}})).$$

The proof of the Theorem 2 (b) is complete. □

References

de Boor, C. (2001). *A Practical Guide to Splines*. New York: Springer.

Liu, Y. X., Z. H. Wang, M. Z. Tian, and K. Yu (2024). Estimation and variable selection for generalized functional partially varying coefficient hybrid models. *Statistical Papers 65*, 93–119.