Efficient Estimation and Computation in Generalized Varying Coefficient Models with Unknown Link and Variance Functions for Larg-Scale Data

Huazhen Lin[1*], Jiaxin Liu[1], Haoqi Li[2], Lixian Pan[1] and Yi Li[3]

Southwestern University of Finance and Economics,China[1]

Yangtze Normal University, China[2]

University of Michigan,USA[3]

Supplementary Material

S1 Notation

We denote by $\boldsymbol{\beta}, g$ and V the true coefficient, link and variance functions, respectively, and $\tilde{V}(\mu) = V(\mu) + \mu^2$. Denote $\dot{f}(x) = df(x)/dx$ and $\ddot{f}(x) = d^2 f(x)/dx^2$ for any $f(\cdot)$. Let $\boldsymbol{\zeta}(u) = (\zeta_1(u), \zeta_2(u), ..., \zeta_d(u))'$, $\boldsymbol{\gamma}(u) = (\gamma_1(u), \gamma_2(u), ..., \gamma_d(u))'$, $\boldsymbol{\delta}(u) = (\boldsymbol{\zeta}'(u), \boldsymbol{\gamma}'(u))'$, $\mathbf{g}(z) = (g_1(z), g_2(z))'$, $\mathbf{V}(\omega) = (V_1(\omega), V_2(\omega))'$, where u,z,ω are defined in Section 2.3. Let I_d be a $d \times d$

identity matrix and $\mathbf{0}_d$ be a $d \times d$ matrix with all entries being 0.

Define

$$\mathcal{C}_d = \{\boldsymbol{\zeta}(\cdot) : \boldsymbol{\zeta}(\cdot) \text{ is continuous on } [-1,1],\ \text{satisfying } ||\boldsymbol{\zeta}(-1)|| = 1,\ \zeta_1(-1) > 0\},$$

$$\mathcal{C}_1 = \{g_1(\cdot) : g_1(\cdot) \text{ is continuous on } [-1,1]\},$$

$$\mathcal{C}_2 = \{V_1(\cdot) : V_1(\cdot) > 0 \text{ and } V_1(\cdot) \text{ is continuous on } [-1,1]\}.$$

The following operators are used in Theorem 2:

$$\mathrm{H}_g(q)(u) = E\left(\mathbf{X}_i q\big\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\big\} \times \dot{g}\big\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\big\}\Big/ V\left[g\left\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\right\}\right]\Big| U_i = u\right) f_1(u),\ \forall\, q \in \mathcal{C}_1;$$

$$\mathrm{H}_{\boldsymbol{\beta}}(\mathbf{q})(u) = E\left(\mathbf{X}_i\Big[\dot{g}\big\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\big\}\Big]^2 \mathbf{X}_i'\mathbf{q}(U_i)\Big/ V\left[g\left\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\right\}\right]\Big|\, U_i = u\right) f_1(u),\ \forall\, \mathbf{q} \in \mathcal{C}_d;$$

$$\mathrm{H}_{\boldsymbol{\beta}g}(\mathbf{q})(z) = E\left(\mathbf{X}_i'\mathbf{q}(U_i)\dot{g}\big\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\big\}\Big/ V\left[g\left\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\right\}\right]\Big|\, \mathbf{X}_i'\boldsymbol{\beta}(U_i) = z\right) f_2(z),\ \forall\, \mathbf{q} \in \mathcal{C}_d;$$

$$\mathrm{H}_{\boldsymbol{\beta}V}(\mathbf{q})(\omega) = E\left(\dot{\tilde{V}}\left[g\{\mathbf{X}'\boldsymbol{\beta}(U_i)\}\right]\left[g\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\} - g\{\mathbf{X}_i'\boldsymbol{\beta}(U_i) + \mathbf{X}_i'\mathbf{q}(U_i)\}\right]\Big| g\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\} = \omega\right)$$
$$\times f_3(\omega),\ \forall\, \mathbf{q} \in \mathcal{C}_d;$$

$$\mathrm{H}_{gV}(q)(\omega) = E\left(q\big\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\big\}\dot{\tilde{V}}\left[g\big\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\big\}\right]\Big| g\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\} = \omega\right) f_3(\omega),\ \forall\, q \in \mathcal{C}_1;$$

where f_1, f_2 and f_3 are the density function of random variable U, $\mathbf{X}'\boldsymbol{\beta}(U)$ and $g(\mathbf{X}'\boldsymbol{\beta}(U))$, respectively. Define an operator matrix Ψ such that

$$\Psi(\mathbf{q}_1, q_2, q_3)(u, z, \omega) = \begin{pmatrix} \mathrm{H}_{\boldsymbol{\beta}}(\mathbf{q}_1)(u) & \mathrm{H}_g(q_2)(u) & 0 \\ \mathrm{H}_{\boldsymbol{\beta}g}(\mathbf{q}_1)(z) & f_2(z)/V\{g(z)\} & 0 \\ \mathrm{H}_{\boldsymbol{\beta}V}(\mathbf{q}_1)(\omega) & \mathrm{H}_{gV}(q_3)(\omega) & f_1(\omega) \end{pmatrix}.$$

Denote by $\mu_2 = \int_0^1 x^2 K(x)dx$, $\mathrm{B}(u, z, \omega) := (B_d'(u), B_{d+1}'(z), B_{d+2}'(\omega))'$

where $B_d(u) = \frac{1}{2}E\left(\mathbf{X}_i'\ddot{\boldsymbol{\beta}}(u)\mathbf{X}_i\left[\dot{g}\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\}\right]^2\Big/V\left[g\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\}\right]|U_i = u\right)f_1(u)\mu_2$, $B_{d+1}(z) = \frac{1}{2}\mu_2\ddot{g}(z)f_2(z)E(V^{-1}[g\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\}] \mid \mathbf{X}_i'\boldsymbol{\beta}(U_i) = z)$ and $B_{d+2}(\omega) = \frac{1}{2}\mu_2\{\ddot{V}(\omega) + 2\}f_3(\omega)$.

Let $\nu = \int_{-1}^{1} K^2(x)dx$. Then we can define $\mathbf{M}(u, z, \omega) = (M_{k,j}(u, z, \omega))$ to be a semipositive definite matrix with the following elements

$$
\begin{aligned}
M_{d,d}(u, z) &= E\Big[\mathbf{X}_i\mathbf{X}_i\frac{\dot{g}^2\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\}}{V\left[g\left\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\right\}\right]}\Big|U_i = u\Big]\nu f_1(u),\\
M_{d,d+1}(u, z) &= E\Big[\mathbf{X}_i\frac{\dot{g}\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\}}{V\left[g\left\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\right\}\right]}\Big|\mathbf{X}_i'\boldsymbol{\beta}(U_i) = z\Big]f_2(z)f_1(u),\\
M_{d+1,d+1}(u, z) &= \frac{\nu f_2(z)}{V\{g(z)\}},\\
M_{d+2,d+2}(\omega) &= E\Big[\left(Y_i^2 - V[g\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\}] - g^2\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\}\right)^2\Big|g\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\} = \omega\Big]\nu f_3(\omega),\\
M_{d,d+2}(u, \omega) &= E\Big\{\left(Y_i^2 - V\left[g\left\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\right\}\right] - g^2\left\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\right\}\right)\left[Y_i - g\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\}\right]\\
&\qquad\times\frac{\mathbf{X}_i\dot{g}\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\}}{V\left[g\left\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\right\}\right]}\Big|g\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\} = \omega\Big\}f_1(u)f_3(\omega),\\
M_{d+1,d+2}(z, \omega) &= E\Big\{\left(Y_i^2 - V[g\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\}] - g^2\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\}\right)\\
&\qquad\times\frac{\left[Y_i - g\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\}\right]}{V[g\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\}]}\Big|g\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\} = \omega\Big\}f_2(z)f_3(\omega).
\end{aligned}
$$

The following notation is used in Theorem 3. We define a 2×2 operator matrix Ψ_1 such that $\Psi_1(\mathbf{q}_1, q_2)(u, z) = \begin{pmatrix} \mathrm{H}_{\boldsymbol{\beta}}(\mathbf{q}_1)(u) & \mathrm{H}_g(q_2)(u) \\ \mathrm{H}_{\boldsymbol{\beta}g}(\mathbf{q}_1)(z) & f_2(z)/V\{g(z)\} \end{pmatrix}$. For any vector of function $\boldsymbol{\psi}'(u, z, \omega) = (\{\psi(u)\}_{j=1}^d, \psi(z))'$, due to linearity of Ψ_1, the following functions are well-defined: $(\{\phi(u)\}_{j=1}^d, \phi(z)) =$

$\boldsymbol{\psi}(u,z)\Psi_1^{-1}$. Let σ_v^2 be

$$\sigma_v^2 = E\left(\left[\dot{g}\left\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\right\}\sum_{j=1}^{d} X_{ij}\phi_j(U_i) + \phi_g\left\{\mathbf{X}'\boldsymbol{\beta}(U_i)\right\}\right]^2 \Big/ V\left[g\left\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\right\}\right]\right).$$

We first introduce two Lemmas, which are useful to prove Theorems.

Lemma 1 *Under (A1)-(A4) and suppose that $g(x,y,z)$ is a bounded and continuous function, then*

$$\sup_{x\in[-1,1]} |c_n(x) - Ec_n(x)| = O_p((\log n)^{1/2}(nh)^{-1/2}),$$

where $c_n(x) = \frac{1}{n}\sum_{i=1}^n g(X_i,(X_i-x)/h,x)K_h(X_i-x)$.

This Lemma is similar to Lemma 4 of Chen et al. (2010) and follows from Theorem 37 and Example 38 in Chapter 2 of Pollard (1984).

Note that Theorem 37 of Pollard (1984) requires bounded random functions, which is no longer valid, for example, when Y_i is unbounded. To solve this problem, we need to introduce another concentration inequality for unbounded random variables in a Hilbert space (Pinelis and Sakhanenko, 1985).

Lemma 2 *Let ξ_i $(i=1,...,n)$ be independent random variables with values in a Hilbert space such that $\mathbb{E}\xi_i = 0$. If for some constants $M, V > 0$, the bound $\mathbb{E}\|\xi_i\|^\ell \le \frac{1}{2}\ell! M^{\ell-2}V$ holds for every $2 \le \ell < \infty$, then*

$$Prob\left\{\Big\|\sum_{i=1}^n \xi_i\Big\| \ge \varepsilon\right\} \le 2\exp\left\{-\frac{\varepsilon^2}{2(\varepsilon M + Vm)}\right\} \quad \forall\varepsilon > 0.$$

S2 Proofs of Theorems 1-3

Proof of Theorem 1. For any vector functions $\boldsymbol{\delta}(\cdot)$, $\mathbf{g}(\cdot)$ and $\mathbf{V}(\cdot)$, set

$$S_{\boldsymbol{\beta}}(\boldsymbol{\delta},\mathbf{g},V_1;u)=\frac{1}{n}\sum_{i=1}^{n}\left(Y_i-g_1\Big[\mathbf{X}_i'\big\{\boldsymbol{\zeta}(u)+\boldsymbol{\gamma}(u)(U_i-u)\big\}\Big]\right)\Big/V_1\big[g_1\left\{\mathbf{X}_i'\boldsymbol{\zeta}(U_i)\right\}\big]$$
$$\times\Upsilon_i(u)g_2\big\{\mathbf{X}_i'\boldsymbol{\zeta}(U_i)\big\}K_{h_1}(U_i-u),$$
$$S_g(\boldsymbol{\zeta},\mathbf{g},V_1;z)=\frac{1}{n}\sum_{i=1}^{n}\Big[Y_i-g_1(z)-g_2(z)\big\{\mathbf{X}_i'\boldsymbol{\zeta}(U_i)-z\big\}\Big]\Big/V_1\big[g_1\left\{\mathbf{X}_i'\boldsymbol{\zeta}(U_i)\right\}\big]$$
$$\times K_{h_2}\big\{\mathbf{X}_i'\boldsymbol{\zeta}(U_i)-z\big\}W_i(\boldsymbol{\zeta};z),$$
$$S_V(\boldsymbol{\zeta},g_1,\mathbf{V};\omega)=\frac{1}{n}\sum_{i=1}^{n}\Big(Y_i^2-\{V_1(\omega)+\omega^2\}-V_2(\omega)\big[g_1\{\mathbf{X}_i'\boldsymbol{\zeta}(U_i)\}-\omega\big]\Big)\Omega_i(\omega;\boldsymbol{\zeta},g_1)$$
$$\times K_{h_3}\Big(g_1\{\mathbf{X}_i'\boldsymbol{\zeta}(U_i)\}-\omega\Big),$$

where $W_i(\boldsymbol{\zeta};z)$, $\Upsilon_i(u)$ and $\Omega_i(\omega;\boldsymbol{\zeta},g)$ is defined in Section 2. Then the deterministic terms of $S_{\boldsymbol{\beta}}(\boldsymbol{\delta},\mathbf{g},V_1;u)$, $S_g(\boldsymbol{\zeta},\mathbf{g},V_1;z)$ and $S_V(\boldsymbol{\zeta},g_1,\mathbf{V};\omega)$ are given by $\mathbf{s}_{\boldsymbol{\beta}}(\boldsymbol{\zeta},\mathbf{g},V_1;u)$, $\mathbf{s}_g(\boldsymbol{\zeta},g_1,V_1;z)$ and $\mathbf{s}_V(\boldsymbol{\zeta},g_1,V_1;\omega)$ which are defined in Section 3, respectively.

Denote $S_{\boldsymbol{\beta}1}(\boldsymbol{\delta},\mathbf{g},V_1;u)$ and $S_{\boldsymbol{\beta}2}(\boldsymbol{\delta},\mathbf{g},V_1;u)$ be the first and last d component of $S_{\boldsymbol{\beta}}$, respectively. Define

$$\mathbf{S}(\boldsymbol{\delta},\mathbf{g},\mathbf{V};u,z,\omega)=\begin{pmatrix}(S_{\boldsymbol{\beta}1}(\boldsymbol{\delta},\mathbf{g},V_1;u),S_{\boldsymbol{\beta}2}(\boldsymbol{\delta},\mathbf{g},V_1;u))\\ S_g(\boldsymbol{\zeta},\mathbf{g},V_1;z)'\\ S_V(\boldsymbol{\zeta},g_1,\mathbf{V};\omega)'\end{pmatrix},$$

Then, the proposed iterative algorithm, under model (1.1), leads to $\mathbf{S}([\widehat{\boldsymbol{\beta}}, \widehat{\dot{\boldsymbol{\beta}}}], \widehat{\mathbf{g}}, \widehat{\mathbf{V}}; u, z, \omega) = 0$, and the model show that $\mathbf{s}(\boldsymbol{\beta}, [g, \dot{g}_2], V; u, z, \omega) = 0$ for any bounded function g_2.

Define

$$\mathcal{B}_n^d = \{f : \|f\|_\infty \leq C, \|f(z_1) - f(z_2)\| \leq c\|z_1 - z_2\| + b_{1n},\ z_1, z_2 \in [-1,1]^d\},$$

$$\mathcal{B}_n^{d+1} = \{f : \|f\|_\infty \leq C, \|f(z_1) - f(z_2)\| \leq c\|z_1 - z_2\| + b_{2n},\ z_1, z_2 \in [-1,1]\},$$

$$\mathcal{B}_n^{d+2} = \{f : \|f\|_\infty \leq C, \|f(z_1) - f(z_2)\| \leq c\|z_1 - z_2\| + b_{3n},\ z_1, z_2 \in [-1,1]\},$$

for some constants $C > 0$ and $c > 0$, where $b_{1n} = \{h_1 + (nh_1)^{-1/2}(\log n)^{1/2}\}$, $b_{2n} = \{h_2 + (nh_2)^{-1/2}(\log n)^{1/2}\}$, $b_{3n} = \{h_3 + (nh_3)^{-1/2}(\log n)^{1/2}\}$.

To show the uniform consistency of $\widehat{\beta}$, $\widehat{g}$ and $\hat{V}$, it suffices to prove the following:

(i) For any continuous function vectors $\boldsymbol{\zeta}$, g_1, V_1 and bounded functions $\boldsymbol{\gamma}, g_2, \dot{V}_1$,

$$\sup_{u,z,\omega\in[-1,1]^3} \|\mathbf{S}(\boldsymbol{\delta}, \mathbf{g}, \mathbf{V}; u, z, \omega) - \mathbf{s}(\boldsymbol{\zeta}, \mathbf{g}, V_1; u, z, \omega)(1, 0)\| = o_p(1).$$

(ii) $\sup_{u,z,\omega\in[-1,1]^3} \|\mathbf{S}(\boldsymbol{\delta}, \mathbf{g}, \mathbf{V}; u, z, \omega) - \mathbf{s}(\boldsymbol{\zeta}, \mathbf{g}, V_1; u, z, \omega)(1, 0)\| = o_p(1)$ uniformly holds over $\boldsymbol{\zeta} \in \mathcal{B}_n^d, g_1 \in \mathcal{B}_n^1, V_1 \in \mathcal{B}_n^2$ and bounded $\boldsymbol{\gamma}, g_2, \dot{V}_1$.

(iii) $P\{\widehat{\boldsymbol{\beta}} \in \mathcal{B}_n^d, \widehat{g} \in \mathcal{B}_n^{d+1}, \widehat{V} \in \mathcal{B}_n^{d+2}\} \to 1$.

Once (i)-(iii) are established, applying the Arzela-Ascoli theorem in $\mathcal{B}_n^k$ ($k = d, d+1, d+2$) for all the estimators, we can show that for any subsequence of $\{\widehat{\boldsymbol{\beta}}, \widehat{g}, \widehat{V}\}$, there exists convergence subsequences $\{\widehat{\boldsymbol{\beta}}, \widehat{g}, \widehat{V}\}_{nk}$, such that uniformly over $u \in [-1,1]$, $z \in [-1,1]$ and $\omega \in [-1,1]$, $\{\widehat{\boldsymbol{\beta}}, \widehat{g}, \widehat{V}\}_{nk} \to \{\boldsymbol{\beta}^*, g^*, V^*\}$ in probability, and it follows that $\boldsymbol{\beta}^* \in \mathcal{C}_d$ and $g^* \in \mathcal{C}_1, V^* \in \mathcal{C}_2$, where $\mathcal{C}$ is the continuous function class. Note that

$$\mathbf{s}(\boldsymbol{\beta}^*, [g^*, \widehat{\dot{g}}], V^*; u, z, \omega)(1,0) = \mathbf{s}(\boldsymbol{\beta}^*, [g^*, \widehat{\dot{g}}], V^*; u, z, \omega)(1,0) - \mathbf{s}(\{\widehat{\boldsymbol{\beta}}, \widehat{\mathbf{g}}, \widehat{V}\}_{nk}; u, z, \omega)(1,0)$$
$$+\mathbf{s}(\{\widehat{\boldsymbol{\beta}}, \widehat{\mathbf{g}}, \widehat{V}\}_{nk}; u, z, \omega)(1,0) - \mathbf{S}(\{[\widehat{\boldsymbol{\beta}}, \widehat{\dot{\boldsymbol{\beta}}}], \widehat{\mathbf{g}}, \widehat{V}\}_{nk}; u, z, \omega).$$

It also follows from (ii) and (iii) that $\mathbf{s}(\boldsymbol{\beta}^*, [g^*, \widehat{\dot{g}}], V^*; u, z, \omega) = 0$ over $u \in [-1,1]$, $z \in [-1,1]$ and $\omega \in [-1,1]$. Since $\mathbf{s}(\boldsymbol{\zeta}, [g_1, \widehat{\dot{g}}], V; u, z, \omega) = 0$ has a unique root at $[\boldsymbol{\beta}, g, V]$, we have $[\boldsymbol{\beta}, g, V] = [\boldsymbol{\beta}^*, g^*, V^*]$, which ensures the uniform consistency of $\widehat{\boldsymbol{\beta}}$, $\widehat{g}$ and $\widehat{V}$. This completes the proof of Theorem 1.

Proof of (i). We only give the proof of $\|S_{\boldsymbol{\beta}}(\boldsymbol{\delta}, \mathbf{g}, V_1; u) - (I_d, \mathbf{0}_d)'\mathbf{s}_{\boldsymbol{\beta}}(\boldsymbol{\zeta}, \mathbf{g}, V_1; u)\|$. Similar arguments result in the conclusions about $S_g(\boldsymbol{\zeta}, \mathbf{g}, V_1; z)$ and $S_V(\boldsymbol{\zeta}, g_1, \mathbf{V}; \omega)$. To estimate $S_{\boldsymbol{\beta}}(\boldsymbol{\delta}, \mathbf{g}, V_1; u) - (I_d, \mathbf{0}_d)'\mathbf{s}_{\boldsymbol{\beta}}(\boldsymbol{\zeta}, \mathbf{g}, V_1; u)$, we consider the following decomposition:

$$S_{\boldsymbol{\beta}}(\boldsymbol{\delta}, \mathbf{g}, V_1; u) - (I_d, \mathbf{0}_d)'\mathbf{s}_{\boldsymbol{\beta}}(\boldsymbol{\zeta}, \mathbf{g}, V_1; u)$$
$$= \{S_{\boldsymbol{\beta}}(\boldsymbol{\delta}, \mathbf{g}, V_1; u) - S_{\boldsymbol{\beta}}([\boldsymbol{\zeta}, 0], \mathbf{g}, V_1; u)\} + \{S_{\boldsymbol{\beta}}([\boldsymbol{\zeta}, 0], \mathbf{g}, V_1; u) - \widetilde{\mathbf{s}}_{\boldsymbol{\beta}}(\boldsymbol{\zeta}, \mathbf{g}, V_1; u)\}$$
$$+ [\widetilde{\mathbf{s}}_{\boldsymbol{\beta}}(\boldsymbol{\zeta}, \mathbf{g}, V_1; u) - (I_d, \mathbf{0}_d)'\mathbf{s}_{\boldsymbol{\beta}}(\boldsymbol{\zeta}, \mathbf{g}, V_1; u)]$$
$$\equiv I_1 + I_2 + I_3,$$

where $\widetilde{\mathbf{s}}_{\boldsymbol{\beta}}$ is the mean of $S_{\boldsymbol{\beta}}([\boldsymbol{\zeta}, 0], \mathbf{g}, V_1; u)$.

First we consider I_1. Let $\mathbf{w}_{g_1}(h)$ be the modulus of continuity of g_1. Observing that

$$\|I_1\| \leq \frac{1}{n}\sum_{i=1}^{n} \mathbf{w}_{g_1}\Big[\mathbf{X}_i'\big\{\boldsymbol{\gamma}(\mathbf{u})(U_i-u)\big\}\Big]\|\Upsilon_i(u)\| \times |g_2\big\{\mathbf{X}_i'\boldsymbol{\zeta}(\mathbf{U}_i)\big\}|\frac{K_{h_1}(U_i-u)}{V_1\big[g_1\left\{\mathbf{X}_i'\boldsymbol{\zeta}(U_i)\right\}\big]}.$$

For any given $i \in \{1, 2, \ldots, n\}$ and any bounded function $\boldsymbol{\gamma}$, note that

$$\begin{aligned}
&\mathbf{w}_{g_1}\Big[\mathbf{X}_i'\big\{\boldsymbol{\gamma}(u)(U_i-u)\big\}\Big]\|\Upsilon_i(u)\| \times |g_2\big\{\mathbf{X}_i'\boldsymbol{\zeta}(U_i)\big\}|\frac{K_{h_1}(U_i-u)}{V_1\big[g_1\left\{\mathbf{X}_i'\boldsymbol{\zeta}(U_i)\right\}\big]}\\
&\leq \mathbf{w}_{g_1}\Big[\mathbf{X}_i'\big\{\mathbf{C}(U_i-u)\big\}\Big]\|\Upsilon_i(u)\| \times |g_2\big\{\mathbf{X}_i'\boldsymbol{\zeta}(U_i)\big\}|\frac{K_{h_1}(U_i-u)}{V_1\big[g_1\left\{\mathbf{X}_i'\boldsymbol{\zeta}(U_i)\right\}\big]},
\end{aligned}$$

where $\mathbf{C}$ is a constant vector. And it is easy to show that

$$\int_{-1}^{1} \mathbf{w}_{g_1}(U_i-u)K_{h_1}(U_i-u)f_1(U_i)dU_i = O_p\big\{\mathbf{w}_{g_1}(h_1)\big\}.$$

for $u \in [-1, 1]$.

For any bounded functions $\boldsymbol{\zeta}$ and $\mathbf{g}$, it follows that

$$E(\|I_1\|) \leq O_p\big\{\mathbf{w}_{g_1}(h_1)\big\}.$$

Hence, Lemma 1 implies that, for any given continuous functions $\boldsymbol{\zeta}$ and $\mathbf{g}$,

$$\sup_{u\in[-1,1]} \|I_1\| = O_p(\mathbf{w}_{g_1}(h_1)) + \mathcal{O}_p((\log n)^{1/2}(nh_1)^{-1/2}) \to 0 \quad \text{as } n \to \infty. \tag{S2.1}$$

To estimate I_2, it suffices to verify the conditions given in Lemma 2. Condition (A1) means that K_{h_1} lies in a Sobolev space denoted by H^2 with the property: $\|f\|_\infty \le c\|f\|_{\mathrm{H}^2}$, for any $f \in \mathrm{H}^2$. Then using Condition (A3), we have

$$\sup_{u\in[-1,1]} \|I_2\| \le O_p((\log n)^{1/2}(nh_1)^{-1/2}). \tag{S2.2}$$

Next consider I_3. By replacing $\boldsymbol{\zeta}(u)$ with $\boldsymbol{\zeta}(U_i)$ in $\widetilde{\mathbf{s}}_{\boldsymbol{\beta}}(\boldsymbol{\zeta}, \mathbf{g}, V_1; u)$, a difference controlled by $\mathbf{w}_{(g_1\circ\boldsymbol{\zeta})}$ is caused for all $u \in [-1,1]$ we note that

$$\begin{aligned}
&\int_{[-1,1]} \mathbf{X}_i'\Big[g\big\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\big\} - g_1\big\{\mathbf{X}_i'\boldsymbol{\zeta}(U_i)\big\}\Big] \\
&\qquad\qquad\qquad \times \frac{g_2\big\{\mathbf{X}_i'\boldsymbol{\zeta}(U_i)\big\}}{V_1\big[g_1\left\{\mathbf{X}_i'\boldsymbol{\zeta}(U_i)\right\}\big]} K_{h_1}(U_i-u)dF(U_i) \\
&= \int_{-1}^{1} H(u_j)K_{h_1}(u_j-u)f_1(u_j)du_j \to \mathbf{s}_{\boldsymbol{\beta}}(\boldsymbol{\zeta}, \mathbf{g}, V_1; u) \text{ as } h_1 \to 0,
\end{aligned}$$

where F is the joint distribution function of U_i,

$$H(u_j) = E\bigg(\mathbf{X}_i'\Big[g\big\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\big\} - g_1\big\{\mathbf{X}_i'\boldsymbol{\zeta}(U_i)\big\}\Big] \times \Big(g_2\big\{\mathbf{X}_i'\boldsymbol{\zeta}(U_i)\big\}/V_1\big[g_1\left\{\mathbf{X}_i'\boldsymbol{\zeta}(U_i)\right\}\big]\Big)\Big|U_i = u_j\bigg).$$

By Lemma 1, it can be shown that

$$\sup_{u\in[-1,1]} \|I_3\| = O_p\big\{\mathbf{w}_{(g_1 o\boldsymbol{\zeta})}(h_1)\big\}. \tag{S2.3}$$

We complete the proof of (i) by combining (S2.1), (S2.2) with (S2.3).

Proof of (ii). Noting that u, z, ω are bounded, the arguments used to prove (ii) is essentially the same as those in Chen et al. (2010).

Proof of (iii). We only give the proof for $\widehat{\boldsymbol{\beta}} \in \mathcal{B}_n^d$.

Given any $u_1, u_2 \in [-1,1]$ with $|u_1 - u_2| \le h_1$. Since $S_{\boldsymbol{\beta}}(\widehat{\boldsymbol{\beta}}, \widehat{\mathbf{g}}, \widehat{V}; u_1) = 0$ and $S_{\boldsymbol{\beta}}(\widehat{\boldsymbol{\beta}}, \widehat{\mathbf{g}}, \widehat{V}; u_2) = 0$, by the Taylor expansion and Condition (A1) it follows that

$$
\begin{aligned}
&S_{\boldsymbol{\beta}}(\widehat{\boldsymbol{\beta}}, \widehat{\mathbf{g}}, \widehat{V}; u_1) - S_{\boldsymbol{\beta}}(\widehat{\boldsymbol{\beta}}, \widehat{\mathbf{g}}, \widehat{V}; u_2) \\
&\quad = \frac{1}{n}\sum_{i=1}^{n} \mathbf{X}_i'\Big(Y_i - \widehat{g}_1\Big[\mathbf{X}_i'\{\widehat{\boldsymbol{\beta}}(u_1) + \widehat{\dot{\boldsymbol{\beta}}}(u_1)(U_i - u_1)\}\Big]\Big)\frac{\widehat{g}_2\big\{\mathbf{X}_i'\widehat{\boldsymbol{\beta}}(U_i)\big\}}{\widehat{V}\Big[\widehat{g}_1\big\{\mathbf{X}_i'\widehat{\boldsymbol{\beta}}(U_i)\big\}\Big]} \\
&\quad \times [K_{h_1}(U_i - u_1) - K_{h_1}(U_i - u_2)] + \frac{1}{n}\sum_{i=1}^{n}\Bigg\{\mathbf{X}_i'\widehat{g}_2\Big[\mathbf{X}_i'\Big\{\widehat{\boldsymbol{\beta}}(u_1) + \widehat{\dot{\boldsymbol{\beta}}}(u_1)(U_i - u_1)\Big\}\Big] \\
&\quad \times \Bigg(\mathbf{X}_i'\Big[\widehat{\boldsymbol{\beta}}(u_2) - \widehat{\boldsymbol{\beta}}(u_1) + (\widehat{\dot{\boldsymbol{\beta}}}(u_2) - \widehat{\dot{\boldsymbol{\beta}}}(u_1))(U_i - u_1) - \widehat{\dot{\boldsymbol{\beta}}}(u_2)(u_2 - u_1)\Big]\Bigg)\Bigg\}\frac{\widehat{g}_2\big\{\mathbf{X}_i'\widehat{\boldsymbol{\beta}}(U_i)\big\}}{\widehat{V}\Big[\widehat{g}_1\big\{\mathbf{X}_i'\widehat{\boldsymbol{\beta}}(U_i)\big\}\Big]} \\
&\quad \times \big\{K_{h_1}(U_i - u_2)\big\} + O\Bigg((u_2 - u_1)^2 + \Big[\widehat{\boldsymbol{\beta}}(u_2) - \widehat{\boldsymbol{\beta}}(u_1)\Big]^2 + b_{1n}\{\widehat{\boldsymbol{\beta}}(u_2) - \widehat{\boldsymbol{\beta}}(u_1)\} + b_{1n}^2\Bigg),
\end{aligned}
$$

By the similar discussion in Cai et al. (2000), we have

$$
\begin{aligned}
&S_{\boldsymbol{\beta}}(\widehat{\boldsymbol{\beta}}, \widehat{\mathbf{g}}, \widehat{V}; u_1) - S_{\boldsymbol{\beta}}(\widehat{\boldsymbol{\beta}}, \widehat{\mathbf{g}}, \widehat{V}; u_2) \\
&\quad = O_p\Big(u_2 - u_1\Big) + O_p\Big(\mathbf{X}_i'\Big\{\widehat{\boldsymbol{\beta}}(u_2) - \widehat{\boldsymbol{\beta}}(u_1)\Big\} + \widehat{\dot{\boldsymbol{\beta}}}(u_2)(u_2 - u_1)\Big) + O_p\Big(b_n\{\widehat{\dot{\boldsymbol{\beta}}}(u_2) - \widehat{\dot{\boldsymbol{\beta}}}(u_1)\}\Big) \\
&\qquad + O\Bigg((u_2 - u_1)^2 + \Big[\widehat{\boldsymbol{\beta}}(u_2) - \widehat{\boldsymbol{\beta}}(u_1)\Big]^2 + b_{1n}\{\widehat{\boldsymbol{\beta}}(u_2) - \widehat{\boldsymbol{\beta}}(u_1)\} + b_{1n}^2\Bigg).
\end{aligned}
$$

Note that $\widehat{\dot{\boldsymbol{\beta}}}$ is bounded, (iii) is held immediately.

Proof of Theorem 2. For convenience of notation, denote

$$a_n = \max_{1\le k\le 3}\{h_k^2 + (nh_k)^{-1/2}(\log n)^{1/2}\},\ c_{1n} = \sup_{u\in[-1,1]} \|\widehat{\boldsymbol\beta}(u) - \boldsymbol\beta(u)\|,$$

$$c_{2n} = \sup_{u\in[-1,1]} \|h_1\widehat{\dot{\boldsymbol\beta}}(u) - h_1\dot{\boldsymbol\beta}(u)\|,\ d_{1n} = \sup_{z\in[-1,1]} |\widehat{g}(z) - g(z)|,$$

$$d_{2n} = \sup_{z\in[-1,1]} |h_2\widehat{\dot{g}}(z) - h_2\dot{g}(z)|,\ e_{1n} = \sup_{\omega\in[-1,1]} |\widehat{V}^{-1}(\omega) - V^{-1}(\omega)|,$$

$$d_{3n} = \sup_{z\in[-1,1]} |\widehat{\dot{g}}(z) - \dot{g}(z)|,\ \ e_{2n} = \sup_{\omega\in[-1,1]} |h_3\widehat{\dot{V}}^{-1}(\omega) - h_3\dot{V}^{-1}(\omega)|,$$

$$a_{kn} = h_k^2 + (nh_k)^{-1/2}(\log n)^{1/2},\ b_{kn} = \{h_k + (nh_k)^{-1/2}(\log n)^{1/2}\}, \text{for } k = 1, 2, 3;$$

$$\mu_2 = \int_0^1 x^2K(x)dx, \quad \mathbb{Q} = \begin{pmatrix} 1 & 0 \\ 0 & \mu_2 \end{pmatrix}.$$

First, we claim that uniformly over $u \in [-1, 1]$ and let $S_{\boldsymbol\beta 1}$ be the first d component of $S_{\boldsymbol\beta}$. We have

$$\begin{aligned} &S_{\boldsymbol\beta 1}([\hat{\boldsymbol\beta}, \hat{\dot{\boldsymbol\beta}}], [\hat g, \hat{\dot g}], \hat V; u) - S_{\boldsymbol\beta 1}([\boldsymbol\beta, \dot{\boldsymbol\beta}], [g, \dot g], V; u) \\ &= -\Big\{\mathrm{H}_{\boldsymbol\beta}(\widehat{\boldsymbol\beta} - \boldsymbol\beta)(u) + \mathrm{H}_g(\widehat g - g)(u)\Big\} + O_p\Big[(c_{1n} + e_{1n} + d_{1n})a_{1n} + c_{1n}(e_{1n} + d_{1n} + c_{1n}) \\ &\qquad + c_{2n}b_{1n} + d_{2n}c_{1n} + d_{1n}(e_{1n} + d_{1n} + d_{2n}) + d_{2n}b_{1n}\Big], \end{aligned} \tag{S2.4}$$

where $\mathrm{H}_{\boldsymbol\beta}$ is an integral-type map from $\mathcal{C}_d$ to $\mathcal{C}_1$ and H_g is an integral operator on $\mathcal{C}_1$, both of which are defined in 1 Notations.

To prove (S2.4), we write

$$S_{\boldsymbol\beta 1}([\hat{\boldsymbol\beta}, \hat{\dot{\boldsymbol\beta}}], [\hat g, \hat{\dot g}], \hat V; x) - S_{\boldsymbol\beta 1}([\boldsymbol\beta, \dot{\boldsymbol\beta}], [g, \dot g], V; u) \equiv J_1 + J_2 + J_3$$

where

$$J_1 = S_{\boldsymbol{\beta}1}([\hat{\boldsymbol{\beta}}, \hat{\dot{\boldsymbol{\beta}}}], [\hat{g}, \hat{\dot{g}}], \hat{V}; u) - S_{\boldsymbol{\beta}1}([\boldsymbol{\beta}, \dot{\boldsymbol{\beta}}], [\hat{g}, \hat{\dot{g}}], \hat{V}; u),$$

$$J_2 = S_{\boldsymbol{\beta}1}([\boldsymbol{\beta}, \dot{\boldsymbol{\beta}}], [\hat{g}, \hat{\dot{g}}], \hat{V}; u) - S_{\boldsymbol{\beta}1}([\boldsymbol{\beta}, \dot{\boldsymbol{\beta}}], [g, \dot{g}], \hat{V}; u),$$

$$J_3 = S_{\boldsymbol{\beta}1}([\boldsymbol{\beta}, \dot{\boldsymbol{\beta}}], [g, \dot{g}], \hat{V}; u) - S_{\boldsymbol{\beta}1}([\boldsymbol{\beta}, \dot{\boldsymbol{\beta}}], [g, \dot{g}], V; u).$$

Similar to the proof of Theorem 1, we can show that $\hat{\dot{\boldsymbol{\beta}}}$, $\hat{\dot{g}}$ and $\hat{\dot{V}}$ are both bounded with probability going to 1, furthermore $\|\hat{\dot{\boldsymbol{\beta}}} - \dot{\boldsymbol{\beta}}\| \to 0$, $\|\hat{\dot{g}} - \dot{g}\| \to 0$ and $\|\hat{\dot{V}} - \dot{V}\| \to 0$. Thus, by the uniform law of large numbers and Taylor expansion of $g(\cdot)$, we conclude that

$$J_1 = -\mathrm{H}_{\boldsymbol{\beta}}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})(u) + O_p\Big[(c_{1n} + e_{1n} + d_{1n})a_{1n} + c_{1n}(e_{1n} + d_{1n} + c_{1n}) + c_{2n}b_{1n} + d_{2n}c_{1n}\}\Big].$$

Similarly, we can obtain that

$$J_2 = -\mathrm{H}_g(\hat{g} - g)(u) + O_p\Big[(a_{1n} + d_{1n})(e_{1n} + d_{1n} + d_{2n}) + d_{2n}b_{1n}\Big],$$

and

$$J_3 = O(a_n e_{1n}).$$

Consequently, this, together with J_1, J_2, J_3, yields the conclusion of (S2.4).

Next we consider $S_g(\widehat{\boldsymbol{\beta}}, [\widehat{g}, \widehat{\dot{g}}], \hat{V}; z) - S_g(\boldsymbol{\beta}, [g, \dot{g}], V; z)$, which can be decomposed as $J_4 + J_5 + J_6$, where

$$J_4 := S_g(\widehat{\boldsymbol{\beta}}, [\widehat{g}, \widehat{\dot{g}}], \hat{V}; z) - S_g(\widehat{\boldsymbol{\beta}}, [g, \dot{g}], \hat{V}; z),$$

$$J_5 := S_g(\widehat{\boldsymbol{\beta}}, [g, \dot{g}], \hat{V}; z) - S_g(\boldsymbol{\beta}, [g, \dot{g}], \hat{V}; z),$$

and

$$J_6 := S_g(\boldsymbol{\beta}, [g, \dot{g}], \hat{V}; z) - S_g(\boldsymbol{\beta}, [g, \dot{g}], V; z).$$

By a simple calculation, we can easily obtain that

$$J_4 = -\mathbb{Q}\Big[\hat{g}(z) - g(z), h_2^2\big\{\hat{\dot{g}}(z) - \dot{g}(z)\big\}\Big]' f_2(z)/V\big\{g(z)\big\} + \begin{pmatrix} O_p\Big[d_{1n}c_{1n} + (d_{1n} + e_{1n})a_{2n} + d_{2n}b_{2n} + c_{1n}e_{1n}\Big] \\ O_p\Big[h_2 d_{2n}c_{1n} + h_2 d_{1n}b_{2n} + h_2 d_{2n}a_{2n}\Big] \end{pmatrix}.$$

Similarly, we can have that

$$J_5 = -H_{\boldsymbol{\beta}g}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) f_2(z)(0,1)' + \begin{pmatrix} O_p\Big\{(e_{1n} + c_{1n})a_{2n} + e_{1n}c_{1n}\Big\} \\ O_p\Big\{h_2 c_{1n}b_{2n} + c_{1n}a_{2n} + c_{1n}^2 h_2\Big\} \end{pmatrix},$$

and

$$J_6 = \begin{pmatrix} O_p(a_{2n}e_{1n}) \\ O_p(h_2 e_{1n}b_{2n}) \end{pmatrix}.$$

Hence, combining with J_4, J_5, J_6 we have that

$$\begin{aligned} & S_g(\widehat{\boldsymbol{\beta}}, [\widehat{g}, \widehat{\dot{g}}], \hat{V}; z) - S_g(\boldsymbol{\beta}, [g, \dot{g}], V; z) \\ &= -\mathbb{Q}\Big[\hat{g}(z) - g(z), h_2^2\big\{\hat{\dot{g}}(z) - \dot{g}(z)\big\}\Big]' f_2(z)/V\big\{g(z)\big\} - \mathrm{H}_{\boldsymbol{\beta}g}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})(z)(1,0)' \\ &\quad + \begin{pmatrix} O_p\Big[d_{1n}c_{1n} + (d_{1n} + c_{1n} + e_{1n})a_{2n} + d_{2n}b_{2n} + c_{1n}e_{1n}\Big] \\ \\ O_p\Big[h_2 d_{2n}c_{1n} + h_2 d_{1n}b_{2n} + (h_2 d_{2n} + h_{2n}c_{1n})a_{2n} + c_{1n}^2 h_2\Big] \end{pmatrix}. \end{aligned} \tag{S2.5}$$

Similarly, we can get

$$
\begin{aligned}
&S_V(\hat{\boldsymbol{\beta}}, \hat{g}, [\hat{V}, \widehat{\dot{\hat{V}}}]; \omega) - S_V(\boldsymbol{\beta}, g, [V, \dot{\hat{V}}]; \omega) \\
&= -\mathbb{Q}\Big[\hat{V}(\omega) - V(\omega), h_3^2\big\{\widehat{\dot{\hat{V}}}(\omega) - \dot{\hat{V}}(\omega)\big\}\Big]' f_3(\omega) - H_{\boldsymbol{\beta} V}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})(1,0)' - H_{gV}(\hat{g} - g)(1,0)' \\
&+ \begin{pmatrix} O_p\Big\{c_{1n}(b_{3n} + d_{1n}) + (c_{1n} + d_{1n} + e_{1n})a_{3n} + (c_{1n} + d_{2n})e_{1n} + e_{2n}b_{3n}\Big\} \\ O_p\Big[c_{1n}h_3 b_{3n} + h_3 c_{1n}^2 + d_{1n}h_3 b_{3n} + (c_{1n} + d_{2n})e_{2n}h_3 + h_3 e_{2n}a_{3n} + h_3 e_{1n}b_{3n}\Big] \end{pmatrix}.
\end{aligned}
\tag{S2.6}
$$

On the other hand, using Condition (A1), Lemma 1 implies that

$$
\sup_{(u,z,\omega)\in[-1,1]^3} \|\mathbf{S}([\boldsymbol{\beta}, \dot{\boldsymbol{\beta}}], [g, \dot{g}], [V, \dot{V}]; u, z, \omega)\| = O_p(a_n).
$$

Note that $\mathbf{S}([\hat{\boldsymbol{\beta}}, \hat{\dot{\boldsymbol{\beta}}}], [\hat{g}, \hat{\dot{g}}], [\hat{V}, \hat{\dot{V}}]; u, z, \omega) = 0$. From (S2.4), we can get $c_{2n} = a_{1n} + (d_{1n} + c_{1n})b_{1n} + c_{2n}a_{1n} + e_{1n}b_{1n}$. Following the second components in (S2.5) and (S2.6), we can get $d_{2n} = a_{2n} + d_{2n}c_{1n} + d_{1n}b_{2n} + (d_{2n} + c_{1n})a_{2n} + c_{1n}^2$ and $e_{2n} = a_{3n} + c_{1n}b_{3n} + c_{1n}^2 + d_{1n}b_{3n} + (c_{1n} + d_{2n})e_{2n} + e_{2n}a_{3n} + e_{1n}b_{3n}$.

Let S_{g1} be the first component of S_g. It follows from (S2.5) that

$$
\begin{aligned}
f_2(z)/V\big\{g(z)\big\}\big(\hat{g} - g\big)(z) + \big(\mathbf{H}_{\boldsymbol{\beta} g}\big)(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})(z) &= S_{g1}(\boldsymbol{\beta}, [g, \dot{g}], V; z) \\
&+ O_p\Big[d_{1n}c_{1n} + (d_{1n} + c_{1n} + e_{1n})a_{2n} + d_{2n}b_{2n} + c_{1n}e_{1n}\Big].
\end{aligned}
\tag{S2.7}
$$

Let S_{v1} be the first component of S_V. It also follows from (S2.6) that

$$f_3(\omega)(\hat{V}-V)(\omega)+\big(\mathbf{H}_{gV}\big)(\widehat{g}-g)(\omega)+\big(\mathbf{H}_{\boldsymbol{\beta}V}\big)(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})(\omega)=S_{v1}(\boldsymbol{\beta},g,[V,\dot{\hat{V}}];\omega)$$
$$+O_p\Big\{c_{1n}(b_{3n}+d_{1n})+(c_{1n}+d_{1n}+e_{1n})a_{3n}+(c_{1n}+d_{2n})e_{1n}+e_{2n}b_{3n}\Big\}$$
$$. \tag{S2.8}$$

This together with (S2.4) and (S2.7), implies that

$$\begin{aligned}&\Big[\mathrm{H}_{\boldsymbol{\beta}}-V\big\{g(z)\big\}/f_2(z)\mathrm{H}_g\circ\mathrm{H}_{\boldsymbol{\beta}g}\Big](\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})(u)\\&=S_{\boldsymbol{\beta}1}([\boldsymbol{\beta},\dot{\boldsymbol{\beta}}],[g,\dot{g}],V;u)-V\big\{g(z)\big\}/f_2(z)\mathrm{H}_g\big\{S_{g1}(\boldsymbol{\beta},[g,\dot{g}],V;z)\big\}(u)\\&\quad+O_p\Big\{(c_{1n}+e_{1n}+d_{1n})a_{1n}+c_{1n}(e_{1n}+d_{1n}+c_{1n})+c_{2n}b_{1n}+d_{2n}c_{1n}\\&\qquad+d_{1n}(e_{1n}+d_{1n})+a_{2n}b_{1n}+(d_{1n}+c_{1n}+e_{1n})a_{2n}+a_{2n}b_{2n}\Big\},\end{aligned} \tag{S2.9}$$

which holds because H_g is a bounded operator on $\mathcal{C}_1$.

Following Condition (A7), $\big[\mathrm{H}_{\boldsymbol{\beta}}-V\big\{g(z)\big\}/f_2(z)\mathrm{H}_g\circ\mathrm{H}_{\boldsymbol{\beta}g}\big]^{-1}$ exists and is bounded on $\mathcal{C}_d$, and hence the supremum norm of the left-side hand of (S2.9) is equivalent to c_{1n}. Lemma 1 further implies that

$$\Big\|S_{\boldsymbol{\beta}1}([\boldsymbol{\beta},\dot{\boldsymbol{\beta}}],[g,\dot{g}],V;u)-V\big\{g(z)\big\}/f_2(z)\mathrm{H}_g\{S_{g1}(\boldsymbol{\beta},[g,\dot{g}],V;z)\}\Big\|=O_p(a_{1n}+a_{2n}).$$

By (S2.9), we have $c_{1n}=O_p(a_{1n}+a_{2n})+O_p\Big\{(e_{1n}+d_{1n})a_{1n}+(d_{1n}+$

$e_{1n})a_{2n} + d_{2n}c_{1n}\Big\}$. Consequently, we can get

$$\begin{aligned} c_{1n} &= O_p(a_{1n} + a_{2n}) = O_p(d_{1n}) \\ e_{1n} &= O_p(a_{1n} + a_{2n} + a_{3n}) = O_p(a_n). \end{aligned} \tag{S2.10}$$

By Condition (A7), we note that Ψ is linear and so Ψ^{-1} is linear and bounded on $\mathcal{C}_d \times \mathcal{C}_1 \times \mathcal{C}_2$. Combining (S2.10) with (S2.4), (S2.7) and (S2.8), we have

$$\Psi \begin{pmatrix} \widehat{\boldsymbol{\beta}} - \boldsymbol{\beta} \\ \widehat{g} - g \\ \widehat{V} - V \end{pmatrix} (u, z, \omega) = \begin{pmatrix} S_{\boldsymbol{\beta}1}([\boldsymbol{\beta}, \dot{\boldsymbol{\beta}}], [g, \dot{g}], V; u) \\ S_{g1}(\boldsymbol{\beta}, [g, \dot{g}], V; z) \\ S_{v1}(\boldsymbol{\beta}, g, [V, \dot{\hat{V}}]; \omega) \end{pmatrix} + \begin{pmatrix} O_p(a_{1n}a_n + a_{1n}b_{1n}) \\ O_p(a_{2n}a_n + a_{2n}b_{2n}) \\ O_p(a_{3n}a_n + a_nb_{3n}) \end{pmatrix}. \tag{S2.11}$$

On the other hand, note that $S_{\boldsymbol{\beta}1}([\boldsymbol{\beta}, \dot{\boldsymbol{\beta}}], [g, \dot{g}], V; u)$ can be expressed as

$$S_{\boldsymbol{\beta}1}([\boldsymbol{\beta}, \dot{\boldsymbol{\beta}}], [g, \dot{g}], V; u) = V_{n,d}(u) + B_{n,d}(u)$$

where

$$\begin{aligned} V_{n,d}(u) &= \frac{1}{n}\sum_{i=1}^{n}\Big[Y_i - g\big\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\big\}\Big]\dot{g}\big\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\big\}\frac{K_{h_1}(U_i - u)}{V\left[g\big\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\big\}\right]}\mathbf{X}_i, \\ B_{n,d}(u) &= \frac{1}{2n}\sum_{i=1}^{n}\Big[\dot{g}\big\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\big\}\Big]^2\mathbf{X}_i'\ddot{\boldsymbol{\beta}}(u)(U_i - u)^2\frac{K_{h_1}(U_i - u)}{V\left[g\big\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\big\}\right]}\mathbf{X}_i + o_p(h_1^2). \end{aligned}$$

We apply Lemma 1 to show that

$$\begin{aligned} B_{n,d}(u) = \frac{1}{2}E\bigg(\mathbf{X}_i'\ddot{\boldsymbol{\beta}}(u)\mathbf{X}_i\Big[\dot{g}\big\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\big\}\Big]^2\Big/V\left[g\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\}\right]|U_i = u\bigg)f_1(u)\mu_2h_1^2 \\ +O_p(h_1^2(nh_1)^{-1/2}(\log n)^{1/2}). \end{aligned}$$

Similarly, we can obtain

$$S_{g1}(\boldsymbol{\beta},[g,\dot{g}];z) = \frac{1}{n}\sum_{i=1}^{n}\Big[Y_i - g\big\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\big\}\Big]\frac{K_{h_2}\big\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)-z\big\}}{V\left[g\big\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\big\}\right]}$$
$$+B_{d+1}(z)h_2^2 + O_p(h_2^2(nh_2)^{-1/2}(\log n)^{1/2}) + o_p(h_2^2),$$
$$S_{v1}(\boldsymbol{\beta},[g,\dot{g}];\omega) = \frac{1}{n}\sum_{i=1}^{n}\bigg(Y_i^2 - V\Big[g\big\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\big\}\Big] - g^2\big\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\big\}\bigg)K_{h_3}\Big[g\big\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\big\}-\omega\Big]$$
$$+B_{d+2}(\omega)h_3^2 + O_p(h_3^2(nh_3)^{-1/2}(\log n)^{1/2}) + o_p(h_3^2),$$

where $B_{d+1}(\cdot)$ and $B_{d+2}(\cdot)$ are defined in Supplementary materials A.

Define $V_{n,d+1}(z)$,$V_{n,d+2}(\omega)$ by

$$V_{n,d+1}(z) = \frac{1}{n}\sum_{i=1}^{n}\Big[Y_i - g\big\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\big\}\Big]\frac{K_{h_2}\big\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)-z\big\}}{V\left[g\big\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\big\}\right]},$$
$$V_{n,d+2}(\omega) = \frac{1}{n}\sum_{i=1}^{n}\bigg(Y_i^2 - V\Big[g\big\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\big\}\Big] - g^2\big\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\big\}\bigg)K_{h_3}\Big[g\big\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\big\}-\omega\Big].$$

We write $\mathrm{A}_n(u,z,w) = (V_{n,d}(u), V_{n,d+1}(z), V_{n,d+2}(\omega))'$ and $\mathrm{B}(u,z,\omega) = (B_d(u), B_{d+1}(u), B_{d+2}(u))'$.

Then it follows from (S2.11) that

$$\Psi\begin{pmatrix}\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}\\ \widehat{g}-g\\ \widehat{V}-V\end{pmatrix}(u,z,\omega) = \mathrm{A}_n(u,z,\omega)+H^2\mathrm{B}(u,z,\omega)+o_p(h_1^2+h_2^2+h_3^2)+\begin{pmatrix}O_p(a_{1n}a_n+a_{1n}b_{1n})\\ O_p(a_{2n}a_n+a_{2n}b_{2n})\\ O_p(a_{3n}a_n+a_nb_{3n})\end{pmatrix},$$

where $H = diag(\underbrace{h_1,\cdots,h_1}_{d}, h_2, h_3)$ and $\mathrm{B}(u,z,\omega)$ is defined in Supplementary materials A. The Central Limit Theorem implies that $(nH)^{1/2}\mathrm{A}_n(u,z,\omega)$ is asymptotically normal with mean 0 and variance-covariance matrix $\mathbf{M}$ where $\mathbf{M}$ is defined in Supplementary materials. Thus the proof of Theorem 2 is completed.

Proof of Theorem 3.

First, we derive the asymptotic variance of $\sum_{j=1}^d \int_{-1}^1 \widehat{\beta}_j(u)\psi_j(u)du + \int_{-1}^1 \widehat{g}(z)\psi_g(z)dz$. Conditioned on h_1 and h_2, from (S2.11), we have

$$\sum_{j=1}^d \int_{-1}^1 \{\widehat{\beta}_j(u) - \beta_j(u)\}\psi_j(u)du + \int_{-1}^1 (\widehat{g} - g)(z)\psi_g(z)dz$$
$$= \frac{1}{n}\sum_{i=1}^n \frac{\Big[Y_i - g\big\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\big\}\Big]}{V\left[g\big\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\big\}\right]}\dot{g}\left\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\right\}\sum_{j=1}^d X_{ij}\phi_j(U_i) + \frac{1}{n}\sum_{i=1}^n \frac{\Big[Y_i - g\big\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\big\}\Big]}{V\left[g\big\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\big\}\right]}\phi_g\big\{\mathbf{X}'\boldsymbol{\beta}(U_i)\big\}$$
$$+O_p(a_{1n}a_n + a_{1n}b_{1n} + a_{2n}a_n + a_{2n}b_{2n}).$$

By Central Limit Theorem, we get

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n \frac{\Big[Y_i - g\big\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\big\}\Big]}{V\left[g\big\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\big\}\right]}\Big[\dot{g}\left\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\right\}\sum_{j=1}^d X_{ij}\phi_j(U_i) + \phi_g\big\{\mathbf{X}'\boldsymbol{\beta}(U_i)\big\}\Big]$$
$$\xrightarrow{d} N(0, \sigma_v^2),$$

where σ_v^2 is

$$\sigma_v^2 = E\Bigg(\Big[\dot{g}\left\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\right\}\sum_{j=1}^d X_{ij}\phi_j(U_i) + \phi_g\big\{\mathbf{X}'\boldsymbol{\beta}(U_i)\big\}\Big]^2 \Big/ V\left[g\big\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\big\}\right]\Bigg).$$

Therefore, we can get

$$\sqrt{n}\sum_{j=1}^d \int_{-1}^1 \{\widehat{\beta}_j(u) - \beta_j(u)\}\psi_j(u)du + \sqrt{n}\int_{-1}^1 (\widehat{g} - g)(z)\psi_g(z)dz \to N(0, \bar{\sigma}^2). \tag{S2.12}$$

To show the asymptotic efficiency of $\sum_{j=1}^d \int_{-1}^1 \hat{\beta}_j(u)\psi_j(u)du + \int_{-1}^1 \widehat{g}(z)\psi_g(z)dz$, we consider the following parametric submodel with unknown parametric

γ,

$$\big\{\boldsymbol{\beta}'(u,z,\gamma), g(z,\gamma)\big\} = \big\{\boldsymbol{\beta}'(u), g(z)\big\} + \gamma\big\{\Phi_{\boldsymbol{\beta}}(u,z), \Phi_g(z)\big\}.$$

where

$$\Phi'_{\boldsymbol{\beta}}(u,z) = \big\{\Phi_1(u,z),\ldots,\Phi_d(u,z)\big\}', \quad \Phi_g(z) = \phi_g^2(z)/[V\{g(z)\}\psi_g(z)],$$

$$\Phi_k(u,z) = \Big[\big\{\sum_{j=1}^d \phi_j(u)EX_{ij}\big\}^2\dot{g}^2(z)+2\sum_{j=1}^d EX_{ij}\phi_j(u)\phi_g(z)\dot{g}(z)\Big]/[dV\{g(z)\}\psi_k(z)].$$

Obviously, $\gamma_0 = 0$ is the true value of γ. Based on the definition of the quasi-likelihood (2.2), the score of this parametric submodel at γ_0 is

$$\frac{1}{n}\sum_{i=1}^n \frac{\Big[Y_i - g\big\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\big\}\Big]}{V\,\big[g\big\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\big\}\big]}\Big[\dot{g}\,\{\mathbf{X}_i'\boldsymbol{\beta}(U_i)\}\sum_{j=1}^d X_{ij}\phi_j(U_i) + \phi_g\big\{\mathbf{X}'\boldsymbol{\beta}(U_i)\big\}\Big],$$

whose variance is σ_v^2. Thus, the maximum likelihood estimator of γ, denoted by $\widetilde{\gamma}$, satisfies

$$\sqrt{n}(\widetilde{\gamma} - \gamma_0) \to N(0, (\sigma_v^2)^{-1}).$$

For any vector functions $\boldsymbol{\psi}'(u,z) = (\{\psi(u)\}_{j=1}^d, \psi(z))'$, we observe that

$$\begin{aligned}&\int_{[u,z]\in[-1,1]^2}\Big[\big\{\boldsymbol{\beta}'(u,z,\widetilde{\gamma}), g(z,\widetilde{\gamma})\big\} - \big\{\boldsymbol{\beta}'(u,z,\gamma_0), g(z,\gamma_0)\big\}\Big]\boldsymbol{\psi}'(u,z)dudz \\ &\quad = (\widetilde{\gamma}-\gamma_0)\int_{[u,z]\in[-1,1]^2}\big\{\Phi_{\boldsymbol{\beta}}(u,z), \Phi_g(z)\big\}\boldsymbol{\psi}'(u,z)dudz. \qquad \text{(S2.13)}\end{aligned}$$

Moreover, we observe that

$$\int_{[u,z]\in[-1,1]^2}\big\{\Phi_{\boldsymbol{\beta}}(u,z), \Phi_g(z)\big\}\boldsymbol{\psi}'(u,z)dudz = \sigma_v^2.$$

Then it follows from (S2.13) that

$$\sqrt{n}\int_{[u,z]\in[-1,1]^2}\Big[\big\{\boldsymbol{\beta}'(u,z,\widetilde{\gamma}),g(z,\widetilde{\gamma})\big\}-\big\{\boldsymbol{\beta}'(u,z,\gamma_0),g(z,\gamma_0)\big\}\Big]\boldsymbol{\psi}'(u,z)dudz$$
$$\xrightarrow{d} N(0,\sigma_v^2).$$

This, together with (S2.12), shows that the asymptotic variance of $\sum_{j=1}^d\int_{-1}^1\widehat{\beta}_j(u)\psi_j(u)du+\int_{-1}^1\widehat{g}(z)\psi_g(z)dz$ is the same as that of $\int_{[u,z]\in[-1,1]^2}(\boldsymbol{\beta}'(u,z,\widetilde{\gamma}),g(z,\widetilde{\gamma}))\psi'(u,z)dudz$. As explained in Bickel et al. (1998) , $\sum_{j=1}^d\int_{-1}^1\widehat{\beta}_j(u)\psi_j(u)du+\int_{-1}^1\widehat{g}(z)\psi_g(z)dz$ is asymptotically efficient for the estimation of $\sum_{j=1}^d\int_{-1}^1\beta_j\psi_j(u)du+\int_{-1}^1 g(z)\psi_g(z)dz$. Thus we complete the proof of Theorem 3.

References

Bickel, P. J., C. A. Klaassen, P. J. Bickel, Y. Ritov, J. Klaassen, J. A. Wellner, and Y. Ritov (1998). *Efficient and adaptive estimation for semiparametric models.* New York: Springer.

Chen, K., S. Guo, L. Sun, and J.-L. Wang (2010). Global partial likelihood for nonparametric proportional hazards models. *J. Amer. Statist. Assoc.* **105**, 750–760.

Pinelis, I. F. and A. I. Sakhanenko (1985). Remarks on inequalities for the probabilities of large deviations. *Theory. Probab. its. Appl.* **30**, 143–148.

Pollard, D. (1984). *Convergence of Stochastic Processes.* New York: Springer Verlag.

Center of Statistical Research and School of Statistics, Southwestern University of Finance and Economics, Chengdu, Sichuan, China.

E-mail: linhz@swufe.edu.cn

Center of Statistical Research and School of Statistics, Southwestern University of Finance and Economics, Chengdu, Sichuan, China.

E-mail: 117020208008@smail.swufe.edu.cn

School of Mathematics and Statistics, Yangtze Normal University, Chongqing, China.

E-mail:lhq213@126.com

Center of Statistical Research and School of Statistics, Southwestern University of Finance and Economics, Chengdu, Sichuan, China.

E-mail: 344848859@qq.com

Department of Biostatistics, University of Michigan, Ann Arbor, MI 48109, USA

E-mail: yili@med.umich.edu