Supplementary Appendices

First, some notations are collected for reference in all proofs and lemmas. The letter C is used as a generic positive constant, which need not be the same from line to line. P_n is the empirical probability measure, and $\mathbb{G}_n f = \sqrt{n}\,(P_n - P)\,f$ is the empirical process indexed by f. Define

$$
\begin{aligned}
&\theta = \left(\gamma, \beta'\right)' = \left(\gamma, \beta_1', \beta_2'\right)'.\ \overline{\beta}_0 = \tfrac{\beta_{10}+\beta_{20}}{2} = \beta_{10} - \tfrac{\delta_0}{2} = \beta_{20} + \tfrac{\delta_0}{2}.\\
&\widetilde{\beta}_\ell = \beta_\ell - \beta_{\ell 0} \text{ is the local perturbation of } \beta_\ell \text{ around } \beta_{\ell 0}.\ e_{\ell i} = y - \mathbf{x}'\beta_{\ell 0}.\\
&h = (v, u_1', u_2')' =: (v, u')' \text{ is the local parameter for } \theta.\\
&s\left(w|\theta\right) = \tfrac{1}{2}\left(y - \mathbf{x}'\beta_1 1(q \le \gamma) - \mathbf{x}'\beta_2 1(q > \gamma)\right)^2.\\
&S_n\left(\theta\right) = P_n s\left(\cdot|\theta\right),\ S\left(\theta\right) = Ps\left(\cdot|\theta\right),\ \mathbb{G}_n s\left(\theta\right) = \sqrt{n}\left(S_n\left(\theta\right) - S\left(\theta\right)\right).\\
&T\left(w|\beta_\ell, \beta_{\ell 0}\right) = \left(y - \mathbf{x}'\tfrac{\beta_{\ell 0}+\beta_\ell}{2}\right)\mathbf{x}'\left(\beta_{\ell 0} - \beta_\ell\right) = \left(m_\ell(x,q) - \mathbf{x}'\tfrac{\beta_{\ell 0}+\beta_\ell}{2}\right)\mathbf{x}'\left(\beta_{\ell 0} - \beta_\ell\right) + \left(\beta_{\ell 0} - \beta_\ell\right)'\mathbf{x}\varepsilon_\ell\\
&= -\left(y - \mathbf{x}'\beta_{\ell 0} - \tfrac{\mathbf{x}'\widetilde{\beta}_\ell}{2}\right)\mathbf{x}'\widetilde{\beta}_\ell = -\left(m_\ell(x,q) - \mathbf{x}'\beta_{\ell 0} - \tfrac{\mathbf{x}'\widetilde{\beta}_\ell}{2}\right)\mathbf{x}'\widetilde{\beta}_\ell - \widetilde{\beta}_\ell'\mathbf{x}\varepsilon_\ell,\\
&\overline{z}_1\left(w|\beta_2, \beta_{10}\right) = \left(y - \mathbf{x}'\tfrac{\beta_{10}+\beta_2}{2}\right)\mathbf{x}'\left(\beta_{10} - \beta_2\right) = \left(m_1(x,q) - \mathbf{x}'\tfrac{\beta_{10}+\beta_2}{2}\right)\mathbf{x}'\left(\beta_{10} - \beta_2\right) + \left(\beta_{10} - \beta_2\right)'\mathbf{x}\varepsilon_1\\
&= \left(y - \mathbf{x}'\overline{\beta}_0 - \tfrac{\mathbf{x}'\widetilde{\beta}_2}{2}\right)\mathbf{x}'\left(\delta_0 - \widetilde{\beta}_2\right) = \left(m_1(x,q) - \mathbf{x}'\overline{\beta}_0 - \tfrac{\mathbf{x}'\widetilde{\beta}_2}{2}\right)\mathbf{x}'\left(\delta_0 - \widetilde{\beta}_2\right) + \left(\delta_0 - \widetilde{\beta}_2\right)'\mathbf{x}\varepsilon_1\\
&= \left(y - \mathbf{x}'\beta_{10} - \tfrac{\mathbf{x}'\left(\widetilde{\beta}_2 - \delta_0\right)}{2}\right)\mathbf{x}'\left(\delta_0 - \widetilde{\beta}_2\right) = \left(m_1(x,q) - \mathbf{x}'\beta_{10} - \tfrac{\mathbf{x}'\left(\widetilde{\beta}_2 - \delta_0\right)}{2}\right)\mathbf{x}'\left(\delta_0 - \widetilde{\beta}_2\right) + \left(\delta_0 - \widetilde{\beta}_2\right)'\mathbf{x}\varepsilon_1,\\
&\overline{z}_2\left(w|\beta_1, \beta_{20}\right) = \left(y - \mathbf{x}'\tfrac{\beta_{20}+\beta_1}{2}\right)\mathbf{x}'\left(\beta_{20} - \beta_1\right) = \left(m_2(x,q) - \mathbf{x}'\tfrac{\beta_{20}+\beta_1}{2}\right)\mathbf{x}'\left(\beta_{20} - \beta_1\right) + \left(\beta_{20} - \beta_1\right)'\mathbf{x}\varepsilon_2\\
&= -\left(y - \mathbf{x}'\overline{\beta}_0 - \tfrac{\mathbf{x}'\widetilde{\beta}_1}{2}\right)\mathbf{x}'\left(\delta_0 + \widetilde{\beta}_1\right) = -\left(m_2(x,q) - \mathbf{x}'\overline{\beta}_0 - \tfrac{\mathbf{x}'\widetilde{\beta}_1}{2}\right)\mathbf{x}'\left(\delta_0 + \widetilde{\beta}_1\right) - \left(\delta_0 + \widetilde{\beta}_1\right)'\mathbf{x}\varepsilon_2\\
&= -\left(y - \mathbf{x}'\beta_{20} - \tfrac{\mathbf{x}'\left(\delta_0 + \widetilde{\beta}_1\right)}{2}\right)\mathbf{x}'\left(\delta_0 + \widetilde{\beta}_1\right) = -\left(m_2(x,q) - \mathbf{x}'\beta_{20} - \tfrac{\mathbf{x}'\left(\delta_0 + \widetilde{\beta}_1\right)}{2}\right)\mathbf{x}'\left(\delta_0 + \widetilde{\beta}_1\right) - \left(\delta_0 + \widetilde{\beta}_1\right)'\mathbf{x}\varepsilon_2,\\
&\overline{z}_{1i} = \overline{z}_1\left(w_i|\beta_{20}, \beta_{10}\right) = \left(y_i - \mathbf{x}_i'\overline{\beta}_0\right)\mathbf{x}_i'\delta_0, \overline{z}_{2i} = \overline{z}_2\left(w_i|\beta_{10}, \beta_{20}\right) = -\left(y_i - \mathbf{x}_i'\overline{\beta}_0\right)\mathbf{x}_i'\delta_0.
\end{aligned}
$$

where $\overline{z}_{1i}$ and $\overline{z}_{2i}$ reduce to $\delta_{q0}q_i\left(y_i - \mathbf{x}_i'\overline{\beta}_0\right)$ and $-\delta_{q0}q_i\left(y_i - \mathbf{x}_i'\overline{\beta}_0\right)$ in CTR and reduces to $\left(\varepsilon_{1i} + \frac{\mathbf{x}_i'\delta_0}{2}\right)\mathbf{x}_i'\delta_0$ and $\left(-\varepsilon_{2i} + \frac{\mathbf{x}_i'\delta_0}{2}\right)\mathbf{x}_i'\delta_0$ as in Yu (2014) and YZ in CS models.

The following formulas are used repetitively in the following analysis:

$$
\begin{aligned}
&s\left(w|\theta\right) = \tfrac{1}{2}\left(y - \mathbf{x}'\beta_1 1(q \le \gamma) - \mathbf{x}'\beta_2 1(q > \gamma)\right)^2 = \tfrac{1}{2}\left(y - \mathbf{x}'\beta_1\right)^2 1(q \le \gamma) + \tfrac{1}{2}\left(y - \mathbf{x}'\beta_2\right)^2 1(q > \gamma)\\
&= \tfrac{1}{2}\left(m_1(x,q) + \varepsilon_1 - \mathbf{x}'\beta_1\right)^2 1(q \le \gamma \wedge \gamma_0) + \tfrac{1}{2}\left(m_2(x,q) + \varepsilon_2 - \mathbf{x}'\beta_2\right)^2 1(q > \gamma \vee \gamma_0)\\
&\quad + \tfrac{1}{2}\left(m_1(x,q) + \varepsilon_1 - \mathbf{x}'\beta_2\right)^2 1(\gamma \wedge \gamma_0 < q \le \gamma_0) + \tfrac{1}{2}\left(m_2(x,q) + \varepsilon_2 - \mathbf{x}'\beta_1\right)^2 1(\gamma_0 < q \le \gamma \vee \gamma_0),
\end{aligned}
$$

so

$$
\begin{aligned}
&s\left(w|\theta\right) - s\left(w|\theta_0\right) = \left(y - \mathbf{x}'\tfrac{\beta_{10}+\beta_1}{2}\right)\mathbf{x}'\left(\beta_{10} - \beta_1\right)1(q \le \gamma \wedge \gamma_0) + \left(y - \mathbf{x}'\tfrac{\beta_{20}+\beta_2}{2}\right)\mathbf{x}'\left(\beta_{20} - \beta_2\right)1(q > \gamma \vee \gamma_0)\\
&\quad + \left[\left(y - \mathbf{x}'\tfrac{\beta_{10}+\beta_2}{2}\right)\mathbf{x}'\left(\beta_{10} - \beta_2\right)\right]1(\gamma \wedge \gamma_0 < q \le \gamma_0) + \left[\left(y - \mathbf{x}'\tfrac{\beta_{20}+\beta_1}{2}\right)\mathbf{x}'\left(\beta_{20} - \beta_1\right)\right]1(\gamma_0 < q \le \gamma \vee \gamma_0)\\
&= T\left(w|\beta_1, \beta_{10}\right)1(q \le \gamma \wedge \gamma_0) + T\left(w|\beta_2, \beta_{20}\right)1(q > \gamma \vee \gamma_0)\\
&\quad + \overline{z}_1\left(w_i|\beta_2, \beta_{10}\right)1(\gamma \wedge \gamma_0 < q \le \gamma_0)\ + \overline{z}_2\left(w|\beta_1, \beta_{20}\right)1(\gamma_0 < q \le \gamma \vee \gamma_0)\\
&= T\left(w|\beta_1, \beta_{10}\right)1(q \le \gamma_0) + T\left(w|\beta_2, \beta_{20}\right)1(q > \gamma_0)\\
&\quad + \left(\overline{z}_1\left(w|\beta_2, \beta_{10}\right) - T\left(w|\beta_1, \beta_{10}\right)\right)1(\gamma \wedge \gamma_0 < q \le \gamma_0) + \left(\overline{z}_2\left(w|\beta_1, \beta_{20}\right) - T\left(w|\beta_2, \beta_{20}\right)\right)1(\gamma_0 < q \le \gamma \vee \gamma_0).
\end{aligned}
$$

We use VW for an abbreviation of Van der Vaart and Wellner (1996), KP for Kim and Pollard (1990), and GC for Glivenko-Cantelli. $\rightsquigarrow$ signifies weak convergence over a compact metric space.

Appendix A: Proofs

Proof of Proposition 1. Note that

$$\delta'\mathbb{E}\left[\mathbf{x}\mathbf{x}'|q=\gamma\right]\delta = \left(\delta_\alpha, \delta_x', \delta_q\right)\begin{pmatrix} 1 & \mathbb{E}\left[x'|q=\gamma\right] & \gamma \\ \mathbb{E}\left[x|q=\gamma\right] & \mathbb{E}\left[xx'|q=\gamma\right] & \gamma\mathbb{E}\left[x|q=\gamma\right] \\ \gamma & \gamma\mathbb{E}\left[x'|q=\gamma\right] & \gamma^2 \end{pmatrix}\begin{pmatrix} \delta_\alpha \\ \delta_x \\ \delta_q \end{pmatrix}$$
$$= \delta_\alpha^2 + \delta_x'\mathbb{E}\left[xx'|q=\gamma\right]\delta_x + \gamma^2\delta_q^2 + 2\gamma\delta_q\mathbb{E}\left[x'|q=\gamma\right]\delta_x + 2\mathbb{E}\left[x'|q=\gamma\right]\delta_x\delta_\alpha + 2\gamma\delta_\alpha\delta_q$$
$$= \left(\delta_\alpha, \delta_x'\right)\begin{pmatrix} 1 & \mathbb{E}\left[x'|q=\gamma\right] \\ \mathbb{E}\left[x|q=\gamma\right] & \mathbb{E}\left[xx'|q=\gamma\right] \end{pmatrix}\begin{pmatrix} \delta_\alpha \\ \delta_x \end{pmatrix} + \gamma^2\delta_q^2 + 2\gamma\delta_q\mathbb{E}\left[x'|q=\gamma\right]\delta_x + 2\gamma\delta_\alpha\delta_q$$
$$=: A + \gamma^2\delta_q^2 + 2\gamma\left(\mathbb{E}\left[x'|q=\gamma\right]\delta_x + \delta_\alpha\right)\delta_q.$$

"$\Longleftarrow$" If $\delta_x = 0$, then $\delta'\mathbb{E}\left[\mathbf{x}\mathbf{x}'|q=\gamma\right]\delta = \delta_\alpha^2 + \delta_q^2\gamma^2 + 2\delta_\alpha\delta_q\gamma = \left(\delta_\alpha + \delta_q\gamma\right)^2$, which is equal to zero only if $\delta_\alpha + \delta_q\gamma$. "$\Longrightarrow$" If $\gamma = 0$, then

$$\delta'\mathbb{E}\left[\mathbf{x}\mathbf{x}'|q=\gamma\right]\delta = \left(\delta_\alpha, \delta_x'\right)\begin{pmatrix} 1 & \mathbb{E}\left[x'|q=0\right] \\ \mathbb{E}\left[x|q=0\right] & \mathbb{E}\left[xx'|q=0\right] \end{pmatrix}\begin{pmatrix} \delta_\alpha \\ \delta_x \end{pmatrix},$$

which is equal to zero only if $\left(\delta_\alpha, \delta_x'\right)' = \mathbf{0}$ when $\begin{pmatrix} 1 & \mathbb{E}\left[x'|q=0\right] \\ \mathbb{E}\left[x|q=0\right] & \mathbb{E}\left[xx'|q=0\right] \end{pmatrix} = \mathbb{E}\left[\begin{pmatrix} 1 \\ x \end{pmatrix}(1, x')\,|q=0\right] > 0$ (equivalently, $Var\left(x|q=0\right) > 0$). In this case, $\left(\delta_\alpha, \delta_x'\right)' = \mathbf{0}$ is equivalent to $\delta_x = \mathbf{0}$ and $\delta_\alpha + \delta_q\gamma = 0$. When $\gamma \neq 0$, is there any other case such that $\delta'\mathbb{E}\left[\mathbf{x}\mathbf{x}'|q=\gamma\right]\delta = 0$ but not the case that $\delta_x = \mathbf{0}$ and $\delta_\alpha + \delta_q\gamma = 0$? If $\gamma > 0$, letting $\delta_q = \frac{-2\gamma\left(\mathbb{E}\left[x'|q=\gamma\right]\delta_x + \delta_\alpha\right) \pm \sqrt{4\gamma^2\left(\mathbb{E}\left[x'|q=\gamma\right]\delta_x + \delta_\alpha\right)^2 - 4\gamma^2 A}}{2\gamma^2} = \frac{-\left(\mathbb{E}\left[x'|q=\gamma\right]\delta_x + \delta_\alpha\right) \pm \sqrt{\left(\mathbb{E}\left[x'|q=\gamma\right]\delta_x + \delta_\alpha\right)^2 - A}}{\gamma}$ can serve the purpose. However, $A - \left(\mathbb{E}\left[x'|q=\gamma\right]\delta_x + \delta_\alpha\right)^2 = \delta_x'\mathbb{E}\left[xx'|q=\gamma\right]\delta_x - \delta_x\mathbb{E}\left[x|q=\gamma\right]\mathbb{E}\left[x'|q=\gamma\right]\delta_x = \delta_x' Var\left(x|q=\gamma\right)\delta_x > 0$ if $\delta_x \neq \mathbf{0}$, so δ_x must be $\mathbf{0}$. If $\delta_x = \mathbf{0}$, we have shown that $\delta_\alpha + \delta_q\gamma$ must be zero to have $\delta'\mathbb{E}\left[\mathbf{x}\mathbf{x}'|q=\gamma\right]\delta = 0$. The case for $\gamma < 0$ can be similarly analyzed. In summary, if $Var\left(x|q=\gamma\right) > 0$, then the only case where $\delta'\mathbb{E}\left[\mathbf{x}\mathbf{x}'|q=\gamma\right]\delta = 0$ but $\delta \neq \mathbf{0}$ is $\delta_x = 0$ and $\delta_\alpha + \delta_q\gamma = 0$. ■

Proof of Proposition 2. Making the change-of-variables $v = \left(\frac{\varpi_-}{\mu_-^2}\right)^{\frac{1}{2\tau-1}} r$, noting the distributional equality $B_\ell(a^2 r) = aB_\ell(r)$, and setting $\omega = \frac{\varpi_-}{\mu_-^2}$, we have

$$\arg\max_v \begin{cases} -\frac{1}{2}\mu_-\left|v\right|^\tau + \sqrt{\varpi_-}B_1(-v), & \text{if } v \le 0, \\ -\frac{1}{2}\mu_+ v^\tau + \sqrt{\varpi_+}B_2(v), & \text{if } v > 0, \end{cases}$$
$$= \left(\frac{\varpi_-}{\mu_-^2}\right)^{\frac{1}{2\tau-1}} \arg\max_r \begin{cases} -\frac{1}{2}\mu_-\left(\frac{\varpi_-}{\mu_-^2}\right)^{\frac{\tau}{2\tau-1}}\left|r\right|^\tau + \sqrt{\varpi_-}B_1\left(-\left(\frac{\varpi_-}{\mu_-^2}\right)^{\frac{1}{2\tau-1}} r\right), & \text{if } r \le 0, \\ -\frac{1}{2}\mu_+\left(\frac{\varpi_-}{4\lambda_-^2}\right)^{\frac{\tau}{2\tau-1}} r^\tau + \sqrt{\varpi_+}B_2\left(\left(\frac{\varpi_-}{\mu_-^2}\right)^{\frac{1}{2\tau-1}} r\right), & \text{if } r > 0, \end{cases}$$
$$= \omega^{\frac{1}{2\tau-1}} \arg\max_r \begin{cases} -\frac{1}{2}\left(\frac{\varpi_-^\tau}{\mu_-}\right)^{\frac{1}{2\tau-1}}\left|r\right|^\tau + \left(\frac{\varpi_-^\tau}{\mu_-}\right)^{\frac{1}{2\tau-1}} B_1(-r), & \text{if } r \le 0, \\ -\frac{1}{2}\frac{\mu_+}{\mu_-}\left(\frac{\varpi_-^\tau}{\mu_-}\right)^{\frac{1}{2\tau-1}} r^\tau + \left(\frac{\varpi_+}{\varpi_-}\right)^{1/2}\left(\frac{\varpi_-^\tau}{\mu_-}\right)^{\frac{1}{2\tau-1}} B_2(r), & \text{if } r > 0, \end{cases}$$
$$= \omega^{\frac{1}{2\tau-1}} \arg\max_r \begin{cases} -\frac{1}{2}\left|r\right|^\tau + B_1(-r), & \text{if } r \le 0, \\ -\frac{1}{2}\varphi r^\tau + \sqrt{\phi}B_2(r), & \text{if } r > 0, \end{cases}$$

where $\varphi = \mu_+/\mu_-$ and $\phi = \varpi_0^+/\varpi_0^-$. From the analysis above,

$$\max_\nu \begin{cases} -\frac{1}{2}\mu_-\left|v\right|^\tau + \sqrt{\varpi_-}B_1(-v), & \text{if } v \le 0, \\ -\frac{1}{2}\mu_+ v^\tau + \sqrt{\varpi_+}B_2(v), & \text{if } v > 0. \end{cases} = \left(\frac{\varpi_-^\tau}{\mu_-}\right)^{\frac{1}{2\tau-1}} \max_r \begin{cases} -\frac{1}{2}\left|r\right|^\tau + B_1(-r), & \text{if } r \le 0, \\ -\frac{1}{2}\varphi r^\tau + \sqrt{\phi}B_2(r), & \text{if } r > 0. \end{cases}$$

In $\xi(\varphi,\phi;1)$, $\xi_1 = \max_{r\leq 0}\left\{B_1(-r)-\frac{|r|}{2}\right\}$, $\xi_2(\varphi,\phi) = \max_{r>0}\left\{\sqrt{\phi}B_2(r)-\varphi\frac{|r|}{2}\right\}$, and ξ_1 and ξ_2 are independent. From Bhattacharya and Brockwell (1976), ξ_1 follows the standard exponential function, and $\xi_2(\varphi,\phi)$ follows an exponential distribution with mean ϕ/φ. It follows that

$$P(\xi(\varphi,\phi)\leq x) = P(\xi_1\leq x,\xi_2(\varphi,\phi)\leq x) = P(\xi_1\leq x)P(\xi_2(\varphi,\phi)\leq x) = (1-e^{-x})(1-e^{-x\varphi/\phi}).$$

■

To derive the asymptotic distribution of an estimator, we will follow the standard procedure, (a) consistency, (b) convergence rate, and (c) derive the weak limit. Consistency is proved in Lemma 1 except the shrinking threshold effect case in Section 4.2 whose consistency is proved in Lemma 3.

Proof of Theorem 1. First, $\widehat{\theta} = \arg\min_\theta S_n(\theta)$ implies

$$\begin{aligned}\widehat{h}_n \quad : &= \left(n(\widehat{\gamma}-\gamma_0),\sqrt{n}\left(\widehat{\beta}-\beta_0\right)\right)\\ &= \arg\min_{(v,u)} nP_n\left(s\left(\cdot\left|\gamma_0+\frac{v}{n},\beta_0+\frac{u}{n^{1/2}}\right.\right)-s\left(\cdot\left|\gamma_0,\beta_0\right.\right)\right) = \arg\min_h\left\{\mathbb{M}_n(h)+o_p(1)\right\}.\end{aligned}$$

where from Lemma 8,

$$\mathbb{M}_n(h) = \frac{1}{2}u_1'M_0u_1+\frac{1}{2}u_2'\overline{M}_0u_2-W_n(u)+D_n(v),$$

and $(W_n(u),D_n(v))$ is defined there. Now, a modified version of the argmax continuous mapping theorem (Theorem 3.2.2 of VW) is used to derive the asymptotic distribution.

(i) $D_n(v)-W_n(u)\rightsquigarrow D(v)-W(u)$, where $D(v)$ is defined in the main text and $W(u) = u_1'W_1+u_2'W_2$ with W_ℓ defined in (9). This is proved in Lemma 12.

(ii) $n(\widehat{\gamma}-\gamma_0) = O_p(1)$ and $\sqrt{n}\left(\widehat{\beta}-\beta_0\right) = O_p(1)$. This is proved in Lemma 2.

(iii) $\arg\min_v D(v) = O_p(1)$ and $\arg\min_{u_1,u_2}\left\{\frac{1}{2}u_1'M_0u_1+\frac{1}{2}u_2'\overline{M}_0u_2-u_1'W_1-u_2'W_2\right\} = O_p(1)$. The former is shown in Appendix D of Yu (2012) and the minimizer in the latter statement is equal to $\left(M_0^{-1}W_1,\overline{M}_0^{-1}W_2\right) = O_p(1)$.

(iv) $\arg\min_v D(v)$ is unique and $\arg\min_{u_1,u_2}\left\{\frac{1}{2}u_1'M_0u_1+\frac{1}{2}u_2'\overline{M}_0u_2-u_1'W_1-u_2'W_2\right\}$ is unique. The former is guaranteed by Assumption (x) and the latter is obvious.

The asymptotic independence between $\widehat{\gamma},\widehat{\beta}_1$ and $\widehat{\beta}_2$ is implied by the independence between $D(v)$, W_1 and W_2 which is shown in Lemma 12. ■

Proof of Theorem 2. First, $\widehat{\theta} = \arg\min_\theta S_n(\theta)$ implies

$$\begin{aligned}\widehat{h}_n \quad : &= \left(a_n(\widehat{\gamma}-\gamma_0),\sqrt{n}\left(\widehat{\beta}-\beta_0\right)\right)\\ &= \arg\min_{(v,u)} nP_n\left(s\left(\cdot\left|\gamma_0+\frac{v}{a_n},\beta_0+\frac{u}{n^{1/2}}\right.\right)-s\left(\cdot\left|\gamma_0,\beta_0\right.\right)\right) = \arg\min_h\left\{\mathbb{M}_n(h)+o_p(1)\right\},\end{aligned}$$

where from Lemma 9,

$$\mathbb{M}_n(h) = \frac{1}{2}u_1'M_0u_1+\frac{1}{2}u_2'\overline{M}_0u_2-W_n(u)+C_n(v),$$

and $(W_n(u),C_n(v))$ is defined there. Now, we apply Theorem 2.7 of KP to derive the asymptotic distribution.

(i) $\mathbb{M}_n(h)\rightsquigarrow\mathbb{M}(h) = \frac{1}{2}u_1'M_0u_1+\frac{1}{2}u_2'\overline{M}_0u_2-W(u)+C(v)\in\mathbf{C}_{\min}\left(\mathbb{R}^{2d+3}\right)$, where $C(v)$ is defined in the main text, and $W(u) = u'W$ is the same as in the proof of Theorem 1, $\mathbf{C}_{\min}\left(\mathbb{R}^{2d+3}\right)$ is defined as the subset of continuous functions $x(\cdot) \in \mathbf{B}_{\text{loc}}\left(\mathbb{R}^{2d+3}\right)$ for which (i) $x(t) \to \infty$ as $\|t\| \to \infty$ and (ii) $x(t)$ achieves its minimum at a unique point in $\mathbb{R}^{2d+3}$, and $\mathbf{B}_{\text{loc}}(\mathbb{R}^{2d+3})$ is the space of all locally bounded real functions on $\mathbb{R}^{2d+3}$, endowed with the uniform metric on compacta. The weak convergence is proved in Lemma 13. We now check $\mathbb{M}(h) \in \mathbf{C}_{\min}\left(\mathbb{R}^{2d+3}\right)$. Because u and v are separable in $\mathbb{M}(h)$, we can check $\mathbb{M}_1(u) := \frac{1}{2}u_1' M_0 u_1 + \frac{1}{2}u_2'\overline{M}_0 u_2 - u'W \in \mathbf{C}_{\min}\left(\mathbb{R}^{2d+2}\right)$ and $\mathbb{M}_2(v) := C(v) \in \mathbf{C}_{\min}(\mathbb{R})$ separately. First, $\mathbb{M}_1(u) \in \mathbf{C}_{\min}\left(\mathbb{R}^{2d+2}\right)$ because it is continuous, has a unique explicit minimizer and $\lim_{\|u\|\to\infty} \mathbb{M}_1(u) = \infty$ with probability one given that for each value of W, $\mathbb{M}_1(u)$ is a quadratic function in u. Second, $\mathbb{M}_2(v) \in \mathbf{C}_{\min}(\mathbb{R})$ because it is continuous, has a unique minimum (see Lemma 2.6 of KP), and $\lim_{|v|\to\infty} \mathbb{M}_2(v) = \infty$ almost surely which follows since $\lim_{|v|\to\infty} B_\ell(v)/|v| = 0$ almost surely by virtue of the law of the iterated logarithm for Brownian motion.

(ii) $a_n(\widehat{\gamma} - \gamma_0) = O_p(1)$ and $\sqrt{n}\left(\widehat{\beta} - \beta_0\right) = O_p(1)$. This is shown in Lemma 4.

Now, by Proposition 2(i),

$$a_n(\widehat{\gamma} - \gamma_0) \xrightarrow{d} \arg\max_v\{-C(v)\} \overset{d}{=} \arg\max_v \begin{cases} -\frac{1}{2}\mu_-|v| + \sqrt{\varpi_-}B_1(-v), & \text{if } v \le 0, \\ -\frac{1}{2}\mu_+|v| + \sqrt{\varpi_+}B_2(v), & \text{if } v > 0, \end{cases} = \omega\zeta(\varphi, \phi; 1)$$

where $\omega = \varpi_-/\mu_-^2$, $\varphi = \mu_+/\mu_-$ and $\phi = \varpi_0^+/\varpi_0^-$. The asymptotic independence between $\widehat{\gamma}, \widehat{\beta}_1$ and $\widehat{\beta}_2$ is implied by the independence between $C(v)$, W_1 and W_2 which is shown in Lemma 13. ■

Proof of Corollary 1. By the CMT and Proposition 2(ii),

$$\begin{aligned} n\left(S_n(\gamma_0) - S_n(\widehat{\gamma})\right) &= n\left[\left(S_n\left(\gamma_0, \widehat{\beta}(\gamma_0)\right) - S_n(\gamma_0, \beta_0)\right) - \left(S_n\left(\widehat{\gamma}, \widehat{\beta}\right) - S_n(\gamma_0, \beta_0)\right)\right] \\ &\xrightarrow{d} \min_u\left\{\frac{1}{2}u'S_{\beta\beta}u - u'W\right\} - \min_{u,v}\left\{\frac{1}{2}u'S_{\beta\beta}u - u'W + C(v)\right\} \\ &= \max_v\ \{-C(v)\} \overset{d}{=} \max_v \begin{cases} -\frac{1}{2}\mu_-|v| + \sqrt{\varpi_-}B_1(-v), & \text{if } v \le 0, \\ -\frac{1}{2}\mu_+|v| + \sqrt{\varpi_+}B_2(v), & \text{if } v > 0, \end{cases} \\ &= \eta^2\xi(\varphi, \phi; 1), \end{aligned}$$

where $\eta^2 = \varpi_-/\mu_-$, and the distribution of $\xi(\varphi, \phi; 1)$ is derived in Proposition 2(iii). The required result follows by Slutsky's theorem. ■

Proof of Theorem 3. We apply Theorem 2.7 of KP to find the asymptotic distribution of $n^{1/3}\left(\widehat{\theta} - \theta_0\right)$. First, $\widehat{\theta} = \arg\min_\theta S_n(\theta)$ implies

$$n^{1/3}\left(\widehat{\theta} - \theta_0\right) = \arg\min_h n^{2/3}\left(S_n\left(\theta_0 + h/n^{1/3}\right) - S_n(\theta_0)\right) = \arg\min_h\{\mathbb{M}_n(h) + o_p(1)\}$$

from Lemma 10, where

$$\mathbb{M}_n(h) := \frac{1}{2}h'S_{\theta\theta}^- h1(v \le 0) + \frac{1}{2}h'S_{\theta\theta}^+ h1(v > 0) + \Xi_n(v),$$

and $\Xi_n(v)$ is defined in Lemma 10.

(i) $\mathbb{M}_n(h) \rightsquigarrow \mathbb{M}(h) := \frac{1}{2}h'S_{\theta\theta}^- h1(v \le 0) + \frac{1}{2}h'S_{\theta\theta}^+ h1(v > 0) + \Xi(v) \in \mathbf{C}_{\min}\left(\mathbb{R}^{2d+3}\right)$. The weak convergence is proved in Lemma 14. We now check $\mathbb{M}(h) \in \mathbf{C}_{\min}\left(\mathbb{R}^{2d+3}\right)$. It is not hard to check $\mathbb{M}(h)$ is continuous, has a unique minimum (see Lemma 2.6 of KP), and $\lim_{\|h\|\to\infty} \mathbb{M}(h) = \infty$ almost surely (see Lemma 2.5 of KP given that the covariance kernel of $\mathbb{M}(h)$ satisfies the rescaling property (2.4) of KP).

(ii) $n^{1/3}\left(\widehat{\theta}-\theta_0\right)=O_p(1)$. This is proved in Lemma 5.

Finally, we show that $\arg\min_h \mathbb{M}(h)$ takes the simplified form in the theorem. We first concentrate on v. Given v, it is not hard to see that

$$\begin{aligned}\widehat{u}(v) &= \arg\min_u \mathbb{M}(v,u) = -S_{\beta\beta}^{-1}\left[S_{\beta\gamma}^{-}v1(v\le 0)+S_{\beta\gamma}^{+}v1(v>0)\right]\\ &= -\begin{pmatrix} M_0^{-1}\left[S_{\beta_1\gamma}^{-}1(v\le 0)+S_{\beta_1\gamma}^{+}1(v>0)\right]v \\ \overline{M}_0^{-1}\left[S_{\beta_2\gamma}^{-}1(v\le 0)+S_{\beta_2\gamma}^{+}1(v>0)\right]v\end{pmatrix}.\end{aligned}$$

Plugging $\widehat{u}(v)$ in $\mathbb{M}(h)$ we have

$$\begin{aligned}\min_u \mathbb{M}(h) &= \frac{v^2}{2}\left[\left(2\lambda^- - S_{\gamma\beta}^{-}S_{\beta\beta}^{-1}S_{\beta\gamma}^{-}\right)1(v\le 0)+\left(2\lambda^+ - S_{\gamma\beta}^{+}S_{\beta\beta}^{-1}S_{\beta\gamma}^{+}\right)1(v>0)\right]+\Xi(v)\\ &= \begin{cases}\frac{1}{2}\left(2\lambda^- - S_{\gamma\beta_1}^{-}M_0^{-1}S_{\beta_1\gamma}^{-}-S_{\gamma\beta_2}^{-}\overline{M}_0^{-1}S_{\beta_2\gamma}^{-}\right)v^2+\sqrt{f_0\omega_0^{-}}B_1(-v), & \text{if } v\le 0,\\ \frac{1}{2}\left(2\lambda^+ - S_{\gamma\beta_1}^{+}M_0^{-1}S_{\beta_1\gamma}^{+}-S_{\gamma\beta_2}^{+}\overline{M}_0^{-1}S_{\beta_2\gamma}^{+}\right)v^2+\sqrt{f_0\omega_0^{+}}B_2(v), & \text{if } v>0,\end{cases}\\ &=: \begin{cases}\frac{1}{2}\mu_- v^2+\sqrt{\varpi_-}B_1(-v), & \text{if } v\le 0,\\ \frac{1}{2}\mu_+ v^2+\sqrt{\varpi_+}B_2(v), & \text{if } v>0,\end{cases} =: \mathbb{M}(v),\end{aligned}$$

where note that $S_{\gamma\beta}^{\pm}S_{\beta\beta}^{-1}S_{\beta\gamma}^{\pm}=S_{\gamma\beta_1}^{\pm}M_0^{-1}S_{\beta_1\gamma}^{\pm}+S_{\gamma\beta_2}^{\pm}\overline{M}_0^{-1}S_{\beta_2\gamma}^{\pm}$, and $\mu_\pm>0$ by $S_{\theta\theta}^{\pm}>0$. By Proposition 2(i), we have

$$n^{1/3}(\widehat{\gamma}-\gamma_0)\xrightarrow{d}\arg\max_v\{-\mathbb{M}(v)\}\overset{d}{=}\arg\max_v\begin{cases}-\frac{1}{2}\mu_- v^2+\sqrt{\varpi_-}B_1(-v), & \text{if } v\le 0,\\ -\frac{1}{2}\mu_+ v^2+\sqrt{\varpi_+}B_2(v), & \text{if } v>0,\end{cases} = \omega^{\frac{1}{3}}\zeta(\varphi,\phi;2),$$

where $\omega=\varpi_-/\mu_-^2$, $\varphi=\mu_+/\mu_-$ and $\phi=\varpi_+/\varpi_-=\omega_0^+/\omega_0^-$. ■

Proof of Corollary 2. Note that

$$\begin{aligned}& n^{2/3}(S_n(\gamma_0)-S_n(\widehat{\gamma}))-n^{2/3}\left(S_n(\gamma_0,\beta_0)-S_n(\widehat{\gamma},\widehat{\beta})\right)\\ =\;& n^{2/3}\left(S_n(\gamma_0,\widehat{\beta}(\gamma_0))-S_n(\gamma_0,\beta_0)\right)\\ =\;& -n^{2/3}\left(\widehat{\beta}(\gamma_0)-\beta_0\right)'\frac{1}{n}\text{diag}\left\{X_{\le\gamma_0}'X_{\le\gamma_0},X_{>\gamma_0}'X_{>\gamma_0}\right\}\left(\widehat{\beta}(\gamma_0)-\beta_0\right)\\ =\;& n^{2/3}O_p\left(n^{-1/2}\right)O(1)O_p\left(n^{-1/2}\right)=o_p(1),\end{aligned}$$

where $\widehat{\beta}(\gamma_0)-\beta_0=O_p\left(n^{-1/2}\right)$, so

$$n^{2/3}(S_n(\gamma_0)-S_n(\widehat{\gamma}))=n^{2/3}\left(S_n(\gamma_0)-S_n(\widehat{\gamma},\widehat{\beta})\right)\xrightarrow{d}-\min_h\mathbb{M}(h)=-\min_v\mathbb{M}(v)=\max_v\{-\mathbb{M}(v)\}.$$

Now, by Proposition 2(ii),

$$\max_v\{-\mathbb{M}(v)\}\overset{d}{=}\max_v\begin{cases}-\frac{1}{2}\mu_- v^2+\sqrt{\varpi_-}B_1(-v), & \text{if } v\le 0,\\ -\frac{1}{2}\mu_+ v^2+\sqrt{\varpi_+}B_2(v), & \text{if } v>0,\end{cases} = \eta^{2/3}\xi(\varphi,\phi;2),$$

where $\eta^2=\varpi_-^2/\mu_-$. The required result follows by Slutsky's theorem. ■

Proof of Theorem 4. We first apply Theorem 2.7 of KP to find the asymptotic distribution of $\left(\rho_n(\widehat{\gamma}-\gamma_0),\kappa_n\left(\widehat{\beta}-\beta_0\right)\right)$ and then refine the asymptotic distribution of $\widehat{\beta}$ to $\min\left(n^{1/2},\rho_n\right)\left(\widehat{\beta}-\beta_0\right)$.

First, $\widehat{\theta} = \arg\min_\theta S_n(\theta)$ implies

$$\left(\rho_n(\widehat{\gamma}-\gamma_0), \kappa_n\left(\widehat{\beta}-\beta_0\right)\right) = \arg\min_h \sqrt{n\rho_n}\left(S_n(\gamma_0 + v/\rho_n, \beta_0 + u/\kappa_n) - S_n(\theta_0)\right) = \arg\min_h \{\mathbb{M}_n(h) + o_p(1)\}$$

from Lemma 10, where

$$\mathbb{M}_n(h) := \frac{1}{2}u_1' M_0 u_1 + \frac{1}{2}u_2' \overline{M}_0 u_2 + \lambda_- |v|^\alpha 1(v \le 0) + \lambda_+ v^\alpha 1(v > 0) + \Xi_n(v),$$

and $\Xi_n(v)$ is defined in Lemma 10.

(i) $\mathbb{M}_n(h) \rightsquigarrow \mathbb{M}(h) := \frac{1}{2}u_1' M_0 u_1 + \frac{1}{2}u_2' \overline{M}_0 u_2 + \lambda_- |v|^\alpha 1(v \le 0) + \lambda_+ v^\alpha 1(v > 0) + \Xi(v) =: \mathbb{M}_1(u) + \mathbb{M}_2(v) \in \mathbf{C}_{\min}\left(\mathbb{R}^{2d+3}\right)$. The weak convergence is shown in Lemma 14. We now check $\mathbb{M}(h) \in \mathbf{C}_{\min}\left(\mathbb{R}^{2d+3}\right)$. It is not hard to check $\mathbb{M}(h)$ is continuous, has a unique minimum (see Lemma 2.6 of KP), so it remains to show $\lim_{\|h\|\to\infty} \mathbb{M}(h) = \infty$ almost surely. Because u and v are separable in $\mathbb{M}(h)$, we can check $\lim_{\|u\|\to\infty} \mathbb{M}_1(u) = \infty$ and $\lim_{|v|\to\infty} \mathbb{M}_2(v) = \infty$ separately. First, $\lim_{\|u\|\to\infty} \mathbb{M}_1(u) = \infty$ is obvious since it is a deterministic quadratic function. Second, since

$$\mathbb{M}_2(v) = \left[\lambda_- |v|^\alpha + \sqrt{\varpi_-} B_1(-v)\right] 1(v \le 0) + \left[\lambda_+ v^\alpha + \sqrt{\varpi_+} B_2(v)\right] 1(v > 0),$$

and by the law of the iterated logarithms for Brownian motion, i.e., $B(v) \le \sqrt{2v \log\log v}$ as $|v| \to \infty$, we have the $B(v)$ term is dominated by the $|v|^\alpha$, $\alpha \ge 1$, term, so $\overline{\lim}_{|v|\to\infty} \mathbb{M}_2(v) = \infty$.

(ii) $\left(\rho_n(\widehat{\gamma}-\gamma_0), \kappa_n\left(\widehat{\beta}-\beta_0\right)\right) = O_p(1)$. This is proved in Lemma 5.

Now, $\left(\rho_n(\widehat{\gamma}-\gamma_0), \kappa_n\left(\widehat{\beta}-\beta_0\right)\right) \xrightarrow{d} \arg\min_h \mathbb{M}(h) = (\arg\min_v \mathbb{M}_2(v), \mathbf{0})$; in other words, $\kappa_n\left(\widehat{\beta}-\beta_0\right)$ is degenerate and the convergence rate of $\widehat{\beta}$ is faster. We first show that $\arg\min_v \mathbb{M}_2(v)$ can be simplied to the form as stated in the theorem. By Proposition 2(i), we have

$$\arg\min_v \ \mathbb{M}_2(v) = \arg\max_v\{-\mathbb{M}_2(v)\} \stackrel{d}{=} \arg\max_v \begin{cases} -\lambda_- |v|^\alpha + \sqrt{\varpi_-} B_1(-v), & \text{if } v \le 0, \\ -\lambda_+ v^\alpha + \sqrt{\varpi_+} B_2(v), & \text{if } v > 0, \end{cases} = \omega^{\frac{1}{2\alpha-1}} \zeta(\varphi, \phi; \alpha),$$

where $\omega = \varpi_-/4\lambda_-^2$, $\varphi = \lambda_+/\lambda_-$ and $\phi = \omega_0^+/\omega_0^-$. We next show the asymptotic distribution of $\widehat{\beta}$. Because the randomness in the γ direction dominates, we cannot search over β and γ jointly; rather, we fix $\widehat{\gamma}$ and concentrate on the randomness in the β direction. Note that

$$\begin{aligned}
&\widehat{\beta}-\beta_0 \approx \arg\min_{u\in\mathcal{N}_u} \{S_n(\widehat{\gamma}, \beta_0 + u) - S_n(\gamma_0, \beta_0)\} \\
&= \arg\min_{u\in\mathcal{N}_u} \left\{[S(\widehat{\gamma}, \beta_0 + u) - S(\gamma_0, \beta_0)] + \left[S_n(\widehat{\gamma}, \beta_0 + u) - S(\widehat{\gamma}, \beta_0 + u) - n^{-1/2}\mathbb{G}_n(s(\cdot|\gamma_0, \beta_0))\right]\right\} \\
&= \arg\min_{u\in\mathcal{N}_u} \left\{\tfrac{1}{2}u' S_{\beta\beta} u + \left[(\widehat{\gamma}-\gamma_0) S_{\gamma\beta}^- u + \lambda_- |\widehat{\gamma}-\gamma_0|^\alpha L_-(|\widehat{\gamma}-\gamma_0|)\right] 1(\widehat{\gamma} \le \gamma_0)\right. \\
&\quad + \left[(\widehat{\gamma}-\gamma_0) S_{\gamma\beta}^+ u + \lambda_+ (\widehat{\gamma}-\gamma_0)^\alpha L_+(\widehat{\gamma}-\gamma_0)\right] 1(\widehat{\gamma} > \gamma_0) \\
&\quad \left. + n^{-1/2}\mathbb{G}_n(s(\cdot|\widehat{\gamma}, \beta_0 + u)) - n^{-1/2}\mathbb{G}_n(s(\cdot|\gamma_0, \beta_0)) + o\left(\rho_n^{-\alpha} L\left(\rho_n^{-1}\right)\right) + o\left(\kappa_n^{-2}\right)\right\} \\
&= \arg\min_{u\in\mathcal{N}_u} \left\{\tfrac{1}{2}u' S_{\beta\beta} u + (\widehat{\gamma}-\gamma_0) S_{\gamma\beta}^- u 1(\widehat{\gamma} \le \gamma_0) + (\widehat{\gamma}-\gamma_0) S_{\gamma\beta}^+ u 1(\widehat{\gamma} > \gamma_0)\right. \\
&\quad \left. - n^{-1/2} u' \mathbb{G}_n \begin{pmatrix} \mathbf{x} e_1 1(q \le \gamma_0) \\ \mathbf{x} e_2 1(q > \gamma_0) \end{pmatrix} + o_p\left(n^{-1/2}\right)\right\} \\
&= \begin{pmatrix} M_0^{-1}\left\{-\left[S_{\beta_1\gamma}^- 1(\widehat{\gamma} \le \gamma_0) + S_{\beta_1\gamma}^+ 1(\widehat{\gamma} > \gamma_0)\right](\widehat{\gamma}-\gamma_0) + n^{-1/2}\mathbb{G}_n(\mathbf{x} e_1 1(q \le \widehat{\gamma}))\right\} \\ \overline{M}_0^{-1}\left\{-\left[S_{\beta_2\gamma}^- 1(\widehat{\gamma} \le \gamma_0) + S_{\beta_2\gamma}^+ 1(\widehat{\gamma} > \gamma_0)\right](\widehat{\gamma}-\gamma_0) + n^{-1/2}\mathbb{G}_n(\mathbf{x} e_2 1(q > \widehat{\gamma}))\right\} \end{pmatrix} + o_p\left(n^{-1/2}\right),
\end{aligned}$$

where the minimization operation is over a κ_n^{-1}-neighborhood of 0, $\mathcal{N}_u$, because $\widehat{\beta}$ is κ_n-consistent, $S_n(\gamma_0, \beta_0) = S(\gamma_0, \beta_0) + n^{-1/2}\mathbb{G}_n(s(\cdot|\gamma_0, \beta_0))$ does not involve u,

$$\begin{aligned} & n^{-1/2}\mathbb{G}_n(s(\cdot|\widehat{\gamma}, \beta_0 + u)) - n^{-1/2}\mathbb{G}_n(s(\cdot|\gamma_0, \beta_0)) \\ = \; & n^{-1/2}\mathbb{G}_n(s(\cdot|\gamma_0, \beta_0 + u)) - n^{-1/2}\mathbb{G}_n(s(\cdot|\gamma_0, \beta_0)) + o_p\left(n^{-1/2}\right) \\ = \; & -n^{-1/2}u'\mathbb{G}_n\begin{pmatrix} \mathbf{x}e_1 1(q \le \gamma_0) \\ \mathbf{x}e_2 1(q > \gamma_0) \end{pmatrix} + \frac{n^{-1/2}}{2}\mathbb{G}_n\begin{pmatrix} u_1'\mathbf{x}u_1 1(q \le \gamma_0) \\ u_2'\mathbf{x}u_2 1(q > \gamma_0) \end{pmatrix} + o_p\left(n^{-1/2}\right) \\ = \; & -n^{-1/2}u'\mathbb{G}_n\begin{pmatrix} \mathbf{x}e_1 1(q \le \gamma_0) \\ \mathbf{x}e_2 1(q > \gamma_0) \end{pmatrix} + o_p\left(n^{-1/2}\right) \end{aligned}$$

and $o\left(\rho_n^{-\alpha}L\left(\rho_n^{-1}\right)\right) = o\left(n^{-1/2}\right)$ and $o\left(\kappa_n^{-2}\right) = o\left(n^{-1/2}\right)$. In summary, the convergence rate of $\widehat{\beta} - \beta_0$ is $\min\left(\rho_n, n^{1/2}\right)$. When $1 < \alpha < 1.5$,

$$n^{1/2}\left(\widehat{\beta} - \beta_0\right) = \text{diag}\left\{M_0^{-1}, \overline{M}_0^{-1}\right\}\mathbb{G}_n\begin{pmatrix} \mathbf{x}e_1 1(q \le \gamma_0) \\ \mathbf{x}e_2 1(q > \gamma_0) \end{pmatrix} + o_p(1) \xrightarrow{d} \begin{pmatrix} Z_{\beta_1} \\ Z_{\beta_2} \end{pmatrix},$$

and when $1.5 < \alpha < 2$,

$$\begin{aligned} \rho_n\left(\widehat{\beta} - \beta_0\right) \; & = \; -\text{diag}\left\{M_0^{-1}, \overline{M}_0^{-1}\right\}\left[S_{\beta\gamma}^{-}1(\widehat{\gamma} \le \gamma_0) + S_{\beta\gamma}^{+}1(\widehat{\gamma} > \gamma_0)\right]\rho_n(\widehat{\gamma} - \gamma_0) + o_p(1) \\ & \xrightarrow{d} -\begin{pmatrix} M_0^{-1}\left[S_{\beta_1\gamma}^{-}1(Z_\gamma(\alpha) \le 0) + S_{\beta_1\gamma}^{+}1(Z_\gamma(\alpha) > 0)\right]Z_\gamma(\alpha), \\ \overline{M}_0^{-1}\left[S_{\beta_2\gamma}^{-}1(Z_\gamma(\alpha) \le 0) + S_{\beta_2\gamma}^{+}1(Z_\gamma(\alpha) > 0)\right]Z_\gamma(\alpha), \end{pmatrix} \end{aligned}$$

and when $\alpha = 1.5$, the weak limit of $n^{1/2}\left(\widehat{\beta} - \beta_0\right)$ is the sum of both limits. ■

Proof of Corollary 3. Note that

$$\begin{aligned} & \sqrt{n\rho_n}\left(S_n(\gamma_0) - S_n(\widehat{\gamma})\right) - \sqrt{n\rho_n}\left(S_n(\gamma_0, \beta_0) - S_n(\widehat{\gamma}, \widehat{\beta})\right) = \sqrt{n\rho_n}\left(S_n(\gamma_0, \widehat{\beta}(\gamma_0)) - S_n(\gamma_0, \beta_0)\right) \\ & = -\sqrt{n\rho_n}\left(\widehat{\beta}(\gamma_0) - \beta_0\right)'\frac{1}{n}\text{diag}\left\{X_{\le\gamma_0}'X_{\le\gamma_0}, X_{>\gamma_0}'X_{>\gamma_0}\right\}\left(\widehat{\beta}(\gamma_0) - \beta_0\right) \\ & = \sqrt{n\rho_n}O_p\left(n^{-1/2}\right)O(1)O_p\left(n^{-1/2}\right) = o_p(1), \end{aligned}$$

where $\widehat{\beta}(\gamma_0) - \beta_0 = O_p\left(n^{-1/2}\right)$, and $\rho_n \prec n$ so $\sqrt{n\rho_n} \prec n$. As a result,

$$\sqrt{n\rho_n}\left(S_n(\gamma_0) - S_n(\widehat{\gamma})\right) = \sqrt{n\rho_n}\left(S_n(\beta_0, \gamma_0) - S_n(\widehat{\beta}, \widehat{\gamma})\right) \xrightarrow{d} -\min_h \mathbb{M}(h) = -\min_v \mathbb{M}_2(v) = \max_v\{-\mathbb{M}_2(v)\}.$$

Now, by Proposition 2(ii),

$$\max_v\{-\mathbb{M}_2(v)\} \overset{d}{=} \max_v\begin{cases} -\lambda_-|v|^\alpha + \sqrt{\varpi_-}B_1(-v), & \text{if } v \le 0, \\ -\lambda_+ v^\alpha + \sqrt{\varpi_+}B_2(v), & \text{if } v > 0, \end{cases} = \eta^{\frac{2}{2\alpha-1}}\xi(\varphi, \phi; \alpha),$$

where $\eta^2 = \frac{\varpi_-^\alpha}{2\lambda_-}$. The required result follows by Slutsky's theorem. ■

Proof of Theorem 5. First, $\widehat{\theta} = \arg\min_\theta S_n(\theta)$ implies

$$\widehat{h}_n := \sqrt{n}\left(\widehat{\theta} - \theta_0\right) = \arg\min_{(v,u)} nP_n\left(s\left(\cdot\left|\gamma_0 + \tfrac{v}{n^{1/2}}, \beta_0 + \tfrac{u}{n^{1/2}}\right.\right) - s(\cdot|\gamma_0, \beta_0)\right) = \arg\min_h\{\mathbb{M}_n(h) + o_p(1)\},$$

where from Lemma 11,

$$\mathbb{M}_n(h) = \frac{1}{2}h' S^-_{\theta\theta} h 1(v \le 0) + \frac{1}{2}h' S^+_{\theta\theta} h 1(v > 0) - W_n(u),$$

and $W_n(u)$ is defined there. Now, we apply the argmax continuous mapping theorem (Theorem 3.2.2 of VW) to derive the asymptotic distribution of $\sqrt{n}\left(\widehat{\theta} - \theta_0\right)$.

(i) $W_n(u) \rightsquigarrow u'W$, where W is defined in (9). This is proved in Lemma 12. So

$$\mathbb{M}_n(h) \rightsquigarrow \frac{1}{2}h' S^-_{\theta\theta} h 1(v \le 0) + \frac{1}{2}h' S^+_{\theta\theta} h 1(v > 0) - u'W =: \mathbb{M}(h).$$

(ii) $\sqrt{n}\left(\widehat{\theta} - \theta_0\right) = O_p(1)$. This is proved in Lemma 6.

(iii) That $\arg\min_h \mathbb{M}(h) = O_p(1)$ is obvious.

(iv) That $\arg\min_h \mathbb{M}(h)$ is unique is obvious.

Finally, we show that $\arg\min_h \mathbb{M}(h)$ takes the simplified form in the theorem. We first concentrate on u. Given u, it is not hard to see that

$$\begin{aligned}\widehat{v}(u) = \arg\min_v \ \mathbb{M}(v,u) = &- \left(S^-_{\gamma\gamma}\right)^{-1} S^-_{\gamma\beta} u \cdot 1(u' S^-_{\beta\gamma}\left(S^-_{\gamma\gamma}\right)^{-1} S^-_{\gamma\beta} u \ge u' S^+_{\beta\gamma}\left(S^+_{\gamma\gamma}\right)^{-1} S^+_{\gamma\beta} u) \\ &- \left(S^+_{\gamma\gamma}\right)^{-1} S^+_{\gamma\beta} u 1(u' S^-_{\beta\gamma}\left(S^-_{\gamma\gamma}\right)^{-1} S^-_{\gamma\beta} u < u' S^+_{\beta\gamma}\left(S^+_{\gamma\gamma}\right)^{-1} S^+_{\gamma\beta} u).\end{aligned}$$

Plugging $\widehat{v}(u)$ in $\mathbb{M}(h)$ we have

$$\begin{aligned}\min_v \mathbb{M}(h) &= -u'W + \tfrac{1}{2}u'\left(S_{\beta\beta} - S^-_{\beta\gamma}\left(S^-_{\gamma\gamma}\right)^{-1} S^-_{\gamma\beta}\right) u \wedge \tfrac{1}{2}u'\left(S_{\beta\beta} - S^+_{\beta\gamma}\left(S^+_{\gamma\gamma}\right)^{-1} S^+_{\gamma\beta}\right) u \\ &= \min\left\{-u'W + \tfrac{1}{2}u'\left(S_{\beta\beta} - S^-_{\beta\gamma}\left(S^-_{\gamma\gamma}\right)^{-1} S^-_{\gamma\beta}\right) u, -u'W + \tfrac{1}{2}u'\left(S_{\beta\beta} - S^+_{\beta\gamma}\left(S^+_{\gamma\gamma}\right)^{-1} S^+_{\gamma\beta}\right) u\right\}.\end{aligned}$$

So

$$\widehat{u} = \arg\min_u \left\{\min_v \mathbb{M}(h)\right\} = \begin{cases} \left(S_{\beta\beta} - S^-_{\beta\gamma}\left(S^-_{\gamma\gamma}\right)^{-1} S^-_{\gamma\beta}\right)^{-1} W, \text{ if } W \in R_1, \\ \left(S_{\beta\beta} - S^+_{\beta\gamma}\left(S^+_{\gamma\gamma}\right)^{-1} S^+_{\gamma\beta}\right)^{-1} W, \text{ if } W \in \overline{R}_1, \end{cases}$$

where

$$R_1 = \left\{W \middle| -\tfrac{1}{2}W'\left(S_{\beta\beta} - S^-_{\beta\gamma}\left(S^-_{\gamma\gamma}\right)^{-1} S^-_{\gamma\beta}\right)^{-1} W \le -\tfrac{1}{2}W'\left(S_{\beta\beta} - S^+_{\beta\gamma}\left(S^+_{\gamma\gamma}\right)^{-1} S^+_{\gamma\beta}\right)^{-1} W\right\},$$

and the event in R_1 cannot reduce to $W'\left(S^-_{\beta\gamma}\left(S^-_{\gamma\gamma}\right)^{-1} S^-_{\gamma\beta} - S^+_{\beta\gamma}\left(S^+_{\gamma\gamma}\right)^{-1} S^+_{\gamma\beta}\right) W \ge 0$,[16] and

$$\widehat{v} = \widehat{v}(\widehat{u}) \xrightarrow{d} \begin{cases} -\left(S^-_{\gamma\gamma}\right)^{-1} S^-_{\gamma\beta}\left(S_{\beta\beta} - S^-_{\beta\gamma}\left(S^-_{\gamma\gamma}\right)^{-1} S^-_{\gamma\beta}\right)^{-1} W, \text{ if } W \in R_1 \cap R_2, \\ -\left(S^+_{\gamma\gamma}\right)^{-1} S^+_{\gamma\beta}\left(S_{\beta\beta} - S^-_{\beta\gamma}\left(S^-_{\gamma\gamma}\right)^{-1} S^-_{\gamma\beta}\right)^{-1} W, \text{ if } W \in R_1 \cap \overline{R}_2, \\ -\left(S^-_{\gamma\gamma}\right)^{-1} S^-_{\gamma\beta}\left(S_{\beta\beta} - S^+_{\beta\gamma}\left(S^+_{\gamma\gamma}\right)^{-1} S^+_{\gamma\beta}\right)^{-1} W, \text{ if } W \in \overline{R}_1 \cap R_3, \\ -\left(S^+_{\gamma\gamma}\right)^{-1} S^+_{\gamma\beta}\left(S_{\beta\beta} - S^+_{\beta\gamma}\left(S^+_{\gamma\gamma}\right)^{-1} S^+_{\gamma\beta}\right)^{-1} W, \text{ if } W \in \overline{R}_1 \cap \overline{R}_3, \end{cases}.$$

[16] Note that this is different from the result that $A \le B$ then $A^{-1} \ge B^{-1}$. If $S_{\beta\beta} - S^-_{\beta\gamma}\left(S^-_{\gamma\gamma}\right)^{-1} S^-_{\gamma\beta} \ge S_{\beta\beta} - S^+_{\beta\gamma}\left(S^+_{\gamma\gamma}\right)^{-1} S^+_{\gamma\beta}$ or the converse, then only one of the two cases can happen.

where

$$R_2=\left\{W|W'\left(S_{\beta\beta}-S^-_{\beta\gamma}\left(S^-_{\gamma\gamma}\right)^{-1}S^-_{\gamma\beta}\right)^{-1}\left[S^-_{\beta\gamma}\left(S^-_{\gamma\gamma}\right)^{-1}S^-_{\gamma\beta}-S^+_{\beta\gamma}\left(S^+_{\gamma\gamma}\right)^{-1}S^+_{\gamma\beta}\right]\left(S_{\beta\beta}-S^-_{\beta\gamma}\left(S^-_{\gamma\gamma}\right)^{-1}S^-_{\gamma\beta}\right)^{-1}W\ge 0\right\},$$
$$R_3=\left\{W|W'\left(S_{\beta\beta}-S^+_{\beta\gamma}\left(S^+_{\gamma\gamma}\right)^{-1}S^+_{\gamma\beta}\right)^{-1}\left[S^-_{\beta\gamma}\left(S^-_{\gamma\gamma}\right)^{-1}S^-_{\gamma\beta}-S^+_{\beta\gamma}\left(S^+_{\gamma\gamma}\right)^{-1}S^+_{\gamma\beta}\right]\left(S_{\beta\beta}-S^+_{\beta\gamma}\left(S^+_{\gamma\gamma}\right)^{-1}S^+_{\gamma\beta}\right)^{-1}W\ge 0\right\},$$

and $\overline{R}_2$ and $\overline{R}_3$ are their negations. ■

Proof of Corollary 4. By the CMT,

$$\begin{aligned}
&n\left(S_n\left(\gamma_0\right)-S_n\left(\widehat{\gamma}\right)\right)\\
&=n\left[\left(S_n\left(\gamma_0,\widehat{\beta}\left(\gamma_0\right)\right)-S_n\left(\gamma_0,\beta_0\right)\right)-\left(S_n\left(\widehat{\gamma},\widehat{\beta}\right)-S_n\left(\gamma_0,\beta_0\right)\right)\right]\\
&\xrightarrow{d}\min_u\left\{\tfrac{1}{2}u'S_{\beta\beta}u-u'W\right\}-\min_{u,v}\left\{\tfrac{1}{2}h'S^-_{\theta\theta}h1(v\le 0)+\tfrac{1}{2}h'S^+_{\theta\theta}h1(v>0)-u'W\right\}\\
&=-\tfrac{1}{2}W'S^{-1}_{\beta\beta}W-\min\left\{-\tfrac{1}{2}W'\left(S_{\beta\beta}-S^-_{\beta\gamma}\left(S^-_{\gamma\gamma}\right)^{-1}S^-_{\gamma\beta}\right)^{-1}W,-\tfrac{1}{2}W'\left(S_{\beta\beta}-S^+_{\beta\gamma}\left(S^+_{\gamma\gamma}\right)^{-1}S^+_{\gamma\beta}\right)^{-1}W\right\}\\
&=\tfrac{1}{2}\max\left\{W'\left[\left(S_{\beta\beta}-S^-_{\beta\gamma}\left(S^-_{\gamma\gamma}\right)^{-1}S^-_{\gamma\beta}\right)^{-1}-S^{-1}_{\beta\beta}\right]W,W'\left[\left(S_{\beta\beta}-S^+_{\beta\gamma}\left(S^+_{\gamma\gamma}\right)^{-1}S^+_{\gamma\beta}\right)^{-1}-S^{-1}_{\beta\beta}\right]W\right\}.
\end{aligned}$$

It can be shown that

$$\begin{aligned}
\left(S_{\beta\beta}-S^\pm_{\beta\gamma}\left(S^\pm_{\gamma\gamma}\right)^{-1}S^\pm_{\gamma\beta}\right)^{-1}-S^{-1}_{\beta\beta}&=S^{-1}_{\beta\beta}S^\pm_{\beta\gamma}\left(S^\pm_{\gamma\gamma}\right)^{-1}\left(\left(S^\pm_{\gamma\gamma}\right)^{-1}-\left(S^\pm_{\gamma\gamma}\right)^{-1}S^\pm_{\gamma\beta}S^{-1}_{\beta\beta}S^\pm_{\beta\gamma}\left(S^\pm_{\gamma\gamma}\right)^{-1}\right)^{-1}\left(S^\pm_{\gamma\gamma}\right)^{-1}S^\pm_{\gamma\beta}S^{-1}_{\beta\beta}\\
&=S^{-1}_{\beta\beta}S^\pm_{\beta\gamma}\left(S^\pm_{\gamma\gamma}-S^\pm_{\gamma\beta}S^{-1}_{\beta\beta}S^\pm_{\beta\gamma}\right)^{-1}S^\pm_{\gamma\beta}S^{-1}_{\beta\beta},
\end{aligned}$$

which results in the conclusion in the corollary. ■

Proof of Theorem 6. We apply Theorem 2.7 of KP to find the asymptotic distribution of $\left(n^{1/3}\left(\widehat{\gamma}-\gamma_0\right),\sqrt{n}\left(\widehat{\beta}-\beta_0\right)\right)$. First, $\widehat{\theta}=\arg\min_\theta S_n\left(\theta\right)$ implies

$$\left(n^{1/3}\left(\widehat{\gamma}-\gamma_0\right),n^{1/2}\left(\widehat{\beta}-\beta_0\right)\right)=\arg\min_h n\left(S_n\left(\gamma_0+v/n^{1/3},\beta_0+u/n^{1/2}\right)-S_n\left(\theta_0\right)\right)=\arg\min_h\left\{\mathbb{M}_n\left(h\right)+o_p\left(1\right)\right\}$$

from Lemma 11, where

$$\mathbb{M}_n\left(h\right):=\frac{1}{2}u_1'M_0u_1+\frac{1}{2}u_2'\overline{M}_0u_2+\lambda_-\left|v\right|^3 1(v\le 0)+\lambda_+v^3 1(v>0)-W_n\left(u\right)+\Xi_n\left(v\right),$$

and $\Xi_n\left(v\right)$ is defined in Lemma 11.

(i) $\mathbb{M}_n\left(h\right)\rightsquigarrow\mathbb{M}\left(h\right):=\frac{1}{2}u_1'M_0u_1+\frac{1}{2}u_2'\overline{M}_0u_2+\lambda_-\left|v\right|^3 1(v\le 0)+\lambda_+v^3 1(v>0)-u'W+\Xi\left(v\right)\in\mathbf{C}_{\min}\left(\mathbb{R}^{2d+3}\right)$. The weak convergence is shown in Lemma 15. Because u and v are separable in $\mathbb{M}\left(h\right)$, we can check $\mathbb{M}_1\left(u\right):=\frac{1}{2}u_1'M_0u_1+\frac{1}{2}u_2'\overline{M}_0u_2-u'W\in\mathbf{C}_{\min}\left(\mathbb{R}^{2d+2}\right)$ and $\mathbb{M}_2\left(v\right):=\lambda_-\left|v\right|^3 1(v\le 0)+\lambda_+v^3 1(v>0)+\Xi\left(v\right)\in\mathbf{C}_{\min}\left(\mathbb{R}\right)$ separately. First, $\mathbb{M}_1\left(u\right)\in\mathbf{C}_{\min}\left(\mathbb{R}^{2d+2}\right)$ by the same arguments as in proof of Theorem 2. Second,

$$\mathbb{M}_2\left(v\right)=\left[\lambda_-\left|v\right|^3+\sqrt{\varpi_-}B_1(-v^3)\right]1(v\le 0)+\left[\lambda_+v^3+\sqrt{\varpi_+}B_2(v^3)\right]1(v>0)$$

is continuous, and has a unique minimizer by Lemma 2.6 of KP. Since

$$\lim_{|v|\to\infty}\mathbb{M}_2\left(v\right)=\lim_{|v|^3\to\infty}\mathbb{M}_2\left(v\right)=\lim_{|\nu|\to\infty}\left[\lambda_-\left|\nu\right|+\sqrt{\varpi_-}B_1(-\nu)\right]1(\nu\le 0)+\left[\lambda_+\nu+\sqrt{\varpi_+}B_2(\nu)\right]1(\nu>0),$$

we have $\overline{\lim}_{|v|\to\infty}\mathbb{M}_2\left(v\right)=\infty$ by the same arguments as in proof of Theorem 2.

(ii) $\left(n^{1/3}\left(\widehat{\gamma}-\gamma_0\right), n^{1/2}\left(\widehat{\beta}-\beta_0\right)\right)=O_p(1)$. This is proved in Lemma 6.

Now, we show that $\arg\min_h \mathbb{M}(h)$ can be simplified to the form as stated in the theorem. First, $\arg\min_u \mathbb{M}_1(u)=\left(Z_{\beta_1}, Z_{\beta_2}\right)$ is standard. Second,

$$\arg\min_v \ \mathbb{M}_2(v)=\arg\max_v\{-\mathbb{M}_2(v)\} \stackrel{d}{=} \arg\max_v \begin{cases} -\lambda_- |v|^3+\sqrt{\varpi_-}B_1(-v^3), & \text{if } v\le 0, \\ -\lambda_+ v^3+\sqrt{\varpi_+}B_2(v^3), & \text{if } v>0, \end{cases} = \widehat{\nu}^{1/3},$$

where

$$\widehat{\nu}=\arg\max_\nu \begin{cases} -\frac{1}{2}2\lambda_- |\nu|+\sqrt{\varpi_-}B_1(-\nu), & \text{if } \nu\le 0, \\ -\frac{1}{2}2\lambda_+\nu+\sqrt{\varpi_+}B_2(\nu), & \text{if } \nu>0. \end{cases}$$

From Proposition 2(i), $\widehat{\nu}=\omega\zeta(\varphi,\phi;1)$ with $\omega=\frac{\varpi_-}{(2\lambda_-)^2}$, $\varphi=\frac{\lambda_+}{\lambda_-}$ and $\phi=\frac{\omega_0^+}{\omega_0^-}$. ■

Proof of Corollary 5. By the CMT and Proposition 2(ii),

$$\begin{aligned}
& n\left(S_n(\gamma_0)-S_n(\widehat{\gamma})\right) \\
= \ & n\left[\left(S_n\left(\gamma_0,\widehat{\beta}(\gamma_0)\right)-S_n(\gamma_0,\beta_0)\right)-\left(S_n\left(\widehat{\gamma},\widehat{\beta}\right)-S_n(\gamma_0,\beta_0)\right)\right] \\
& \xrightarrow{d} \min_u\left\{\frac{1}{2}u'S_{\beta\beta}u-u'W\right\}-\min_{u,v}\left\{\frac{1}{2}u'S_{\beta\beta}u-u'W+\lambda_-|v|^3 1(v\le 0)+\lambda_+ v^3 1(v>0)+\Xi(v)\right\} \\
= \ & -\min_v \ \left\{\lambda_-|v|^3 1(v\le 0)+\lambda_+ v^3 1(v>0)+\Xi(v)\right\} \\
& \stackrel{d}{=} \max_v \begin{cases} -\lambda_- |v|^3+\sqrt{\varpi_-}B_1(-v^3), & \text{if } v\le 0, \\ -\lambda_+ v^3+\sqrt{\varpi_+}B_2(v^3), & \text{if } v>0, \end{cases} \\
= \ & \max_r \begin{cases} -\frac{1}{2}2\lambda_- |r|+\sqrt{\varpi_-}B_1(-r), & \text{if } r\le 0, \\ -\frac{1}{2}2\lambda_+ |r|+\sqrt{\varpi_+}B_2(r), & \text{if } r>0, \end{cases} = \eta^2\xi(\varphi,\phi;1)
\end{aligned}$$

where $\eta^2=\frac{\varpi_-}{2\lambda_-}$, and the distribution of $\xi(\varphi,\phi;1)$ is derived in Proposition 2(iii). The required result follows by Slutsky's theorem. ■

Proof of Theorem 7. We apply Theorem 2.7 of KP to find the asymptotic distribution of $\left(n^{1/5}\left(\widehat{\gamma}-\gamma_0\right), n^{2/5}\left(\widehat{\beta}-\beta_0\right)\right)$. First, $\widehat{\theta}=\arg\min_\theta S_n(\theta)$ implies

$$\left(n^{1/5}\left(\widehat{\gamma}-\gamma_0\right), n^{2/5}\left(\widehat{\beta}-\beta_0\right)\right)=\arg\min_h n^{4/5}\left(S_n\left(\gamma_0+v/n^{1/5},\beta_0+u/n^{2/5}\right)-S_n(\theta_0)\right)=\arg\min_h\{\mathbb{M}_n(h)+o_p(1)\}$$

from Lemma 11, where

$$\mathbb{M}_n(h):=\frac{1}{2}\left(u',v^2\right)\mathbb{S}^-\left(u',v^2\right)'1(v\le 0)+\frac{1}{2}\left(u',v^2\right)\mathbb{S}^+\left(u',v^2\right)'1(v>0)+\Xi_n(v),$$

and $\Xi_n(v)$ is defined in Lemma 11.

(i) $\mathbb{M}_n(h)\rightsquigarrow\mathbb{M}(h):=\frac{1}{2}\left(u',v^2\right)\mathbb{S}^-\left(u',v^2\right)'1(v\le 0)+\frac{1}{2}\left(u',v^2\right)\mathbb{S}^+\left(u',v^2\right)'1(v>0)+\Xi(v)\in\mathbf{C}_{\min}\left(\mathbb{R}^{2d+3}\right)$. The weak convergence is shown in Lemma 15. We now check $\mathbb{M}(h)\in\mathbf{C}_{\min}\left(\mathbb{R}^{2d+3}\right)$. It is not hard to check $\mathbb{M}(h)$ is continuous, has a unique minimum (see Lemma 2.6 of KP), so it remains to show $\lim_{\|h\|\to\infty}\ \mathbb{M}(h)=\infty$ almost surely. Note that $\mathbb{M}(h)-\Xi(v)\ge\frac{C}{2}u'S_{\beta\beta}u+C(\lambda_-\wedge\lambda_+)v^4$, and $\Xi(v)=\sqrt{\varpi_-}B_1(-v^3)1(v\le 0)+\sqrt{\varpi_+}B_2(v^3)1(v>0)$. By the law of the iterated logarithms for Brownian motion, $B(v^3)\le\sqrt{2v^3\log\log v^3}$ as $|v|\to\infty$, so $B(v^3)$ is dominated by v^4. As a result, $\lim_{\|h\|\to\infty}\ \mathbb{M}(h)=\infty$.

(ii) $\left(n^{1/5}\left(\widehat{\gamma}-\gamma_0\right), n^{2/5}\left(\widehat{\beta}-\beta_0\right)\right)=O_p(1)$. This is proved in Lemma 6.

Now, we show that $\arg\min_h \mathbb{M}(h)$ can be simplified to the form as stated in the theorem. We first concentrate on v. Given v, it is not hard to see that

$$\widehat{u}(v)=\arg\min_u \ \mathbb{M}(v,u)=-S_{\beta\beta}^{-1}\left(\frac{1}{2}S_{\beta\gamma^2}\right)v^2=-\frac{1}{2}\begin{pmatrix} M_0^{-1}S_{\beta_\ell\gamma^2} \\ \overline{M}_0^{-1}S_{\beta_\ell\gamma^2}\end{pmatrix}v^2.$$

Plugging $\widehat{u}(v)$ in $\mathbb{M}(h)$ we have

$$\begin{aligned}
&\min\nolimits_u \mathbb{M}(h)=\tfrac{v^4}{2}\left[2\lambda^- 1(v\le 0)+2\lambda^+ 1(v>0)-\tfrac{1}{4}S_{\gamma^2\beta}S_{\beta\beta}^{-1}S_{\beta\gamma^2}\right]+\Xi(v)\\
&=\begin{cases} \tfrac{1}{2}\left(2\lambda^- -\tfrac{1}{4}S_{\gamma^2\beta_\ell}M_0^{-1}S_{\beta_\ell\gamma^2}-\tfrac{1}{4}S_{\gamma^2\beta_\ell}\overline{M}_0^{-1}S_{\beta_\ell\gamma^2}\right)v^4+\sqrt{\tfrac{f_0\delta_{q0}^2}{3}\omega_0^-}B_1(-v^3), & \text{if } v\le 0,\\ \tfrac{1}{2}\left(2\lambda^+ -\tfrac{1}{4}S_{\gamma^2\beta_\ell}M_0^{-1}S_{\beta_\ell\gamma^2}-\tfrac{1}{4}S_{\gamma^2\beta_\ell}\overline{M}_0^{-1}S_{\beta_\ell\gamma^2}\right)v^4+\sqrt{\tfrac{f_0\delta_{q0}^2}{3}\omega_0^+}B_2(v^3), & \text{if } v>0,\end{cases}\\
&=:\begin{cases}\tfrac{1}{2}\mu_- v^4+\sqrt{\varpi_-}B_1(-v^3), & \text{if } v\le 0,\\ \tfrac{1}{2}\mu_+ v^4+\sqrt{\varpi_+}B_2(v^3), & \text{if } v>0,\end{cases} =:\mathbb{M}(v),
\end{aligned}$$

where $\mu_\pm>0$ by $\mathbb{S}_{\theta\theta}^{\pm}>0$. So

$$n^{1/5}\left(\widehat{\gamma}-\gamma_0\right)\xrightarrow{d}\arg\min_v \ \mathbb{M}(v)=\arg\max_v\{-\mathbb{M}(v)\}\overset{d}{=}\arg\max_v\begin{cases}-\tfrac{1}{2}\mu_- v^4+\sqrt{\varpi_-}B_1(-v^3), & \text{if } v\le 0,\\ -\tfrac{1}{2}\mu_+ v^4+\sqrt{\varpi_+}B_2(v^3), & \text{if } v>0,\end{cases}=\widehat{\nu}^{1/3},$$

where by Proposition 2(i),

$$\widehat{\nu}=\arg\max_\nu\begin{cases}-\tfrac{1}{2}\mu_-\nu^{4/3}+\sqrt{\varpi_-}B_1(-\nu), & \text{if } \nu\le 0,\\ -\tfrac{1}{2}\mu_+\nu^{4/3}+\sqrt{\varpi_+}B_2(\nu), & \text{if } \nu>0,\end{cases}=\omega^{3/5}\zeta\left(\varphi,\phi;4/3\right),$$

with $\omega=\varpi_-/\mu_-^2$, $\varphi=\mu_+/\mu_-$ and $\phi=\varpi_+/\varpi_-=\omega_0^+/\omega_0^-$. ■

Proof of Corollary 6. Note that

$$\begin{aligned}
&n^{4/5}\left(S_n(\gamma_0)-S_n\left(\widehat{\gamma}\right)\right)-n^{4/5}\left(S_n(\gamma_0,\beta_0)-S_n(\widehat{\gamma},\widehat{\beta})\right)=n^{4/5}\left(S_n(\gamma_0,\widehat{\beta}\left(\gamma_0\right))-S_n\left(\gamma_0,\beta_0\right)\right)\\
&=-n^{4/5}\left(\widehat{\beta}\left(\gamma_0\right)-\beta_0\right)'\tfrac{1}{n}\mathrm{diag}\left\{X_{\le\gamma_0}'X_{\le\gamma_0},X_{>\gamma_0}'X_{>\gamma_0}\right\}\left(\widehat{\beta}\left(\gamma_0\right)-\beta_0\right)\\
&=n^{4/5}O_p\left(n^{-1/2}\right)O\left(1\right)O_p\left(n^{-1/2}\right)=o_p\left(1\right),
\end{aligned}$$

where $\widehat{\beta}\left(\gamma_0\right)-\beta_0=O_p\left(n^{-1/2}\right)$, so

$$n^{4/5}\left(S_n(\gamma_0)-S_n\left(\widehat{\gamma}\right)\right)=n^{4/5}\left(S_n(\theta_0,\gamma_0)-S_n(\widehat{\beta},\widehat{\gamma})\right)\xrightarrow{d}-\min_h\mathbb{M}(h)=-\min_v\mathbb{M}(v)=\max_v\left\{-\mathbb{M}(v)\right\}.$$

Now, by Proposition 2(ii),

$$\begin{aligned}
&\max_v\left\{-\mathbb{M}(v)\right\}\overset{d}{=}\max_v\begin{cases}-\tfrac{1}{2}\mu_- v^4+\sqrt{\varpi_-}B_1(-v^3), & \text{if } v\le 0,\\ -\tfrac{1}{2}\mu_+ v^4+\sqrt{\varpi_+}B_2(v^3), & \text{if } v>0,\end{cases}\\
&=\ \max_\nu\begin{cases}-\tfrac{1}{2}\mu_-\nu^{4/3}+\sqrt{\varpi_-}B_1(-\nu), & \text{if } \nu\le 0,\\ -\tfrac{1}{2}\mu_+\nu^{4/3}+\sqrt{\varpi_+}B_2(\nu), & \text{if } \nu>0,\end{cases}=\eta^{6/5}\xi\left(\varphi,\phi;4/3\right),
\end{aligned}$$

where $\eta^2=\frac{\varpi_-^{4/3}}{\mu_-}$. The required result follows by Slutsky's theorem. ■

Proof of Theorem 8. We apply Theorem 2.7 of KP to find the asymptotic distribution of $\left(\rho_n\left(\widehat{\gamma}-\gamma_0\right), \kappa_n\left(\widehat{\beta}-\beta_0\right)\right)$. First, $\widehat{\theta}=\arg\min_\theta S_n(\theta)$ implies

$$\left(\rho_n\left(\widehat{\gamma}-\gamma_0\right), \kappa_n\left(\widehat{\beta}-\beta_0\right)\right)=\arg\min_h \sqrt{n\rho_n^3}\left(S_n\left(\gamma_0+v/\rho_n, \beta_0+u/\kappa_n\right)-S_n\left(\theta_0\right)\right)=\arg\min_h\left\{\mathbb{M}_n(h)+o_p(1)\right\}$$

from Lemma 11, where

$$\mathbb{M}_n(h):=\frac{1}{2}u_1'M_0u_1+\frac{1}{2}u_2'\overline{M}_0u_2+\lambda_-|v|^\alpha 1(v\leq 0)+\lambda_+v^\alpha 1(v>0)+\Xi_n(v),$$

and $\Xi_n(v)$ is defined in Lemma 11.

(i) $\mathbb{M}_n(h)\rightsquigarrow\mathbb{M}(h):=\frac{1}{2}u_1'M_0u_1+\frac{1}{2}u_2'\overline{M}_0u_2+\lambda_-|v|^\alpha 1(v\leq 0)+\lambda_+v^\alpha 1(v>0)+\Xi(v)=:\mathbb{M}_1(u)+\mathbb{M}_2(v)\in\mathbf{C}_{\min}\left(\mathbb{R}^{2d+3}\right)$. The weak convergence is shown in Lemma 15. We now check $\mathbb{M}(h)\in\mathbf{C}_{\min}\left(\mathbb{R}^{2d+3}\right)$. The proof is similar to that in the proof of Theorem 4; only notice that

$$\mathbb{M}_2(v)=\left[\lambda_-|v|^\alpha+\sqrt{\varpi_-}B_1(-v^3)\right]1(v\leq 0)+\left[\lambda_+v^\alpha+\sqrt{\varpi_+}B_2(v^3)\right]1(v>0),$$

so

$$\lim_{|v|\to\infty}\mathbb{M}_2(v)=\lim_{|v|^3\to\infty}\mathbb{M}_2(v)=\lim_{|\nu|\to\infty}\left[\lambda_-|\nu|^{\alpha/3}+\sqrt{\varpi_-}B_1(-\nu)\right]1(\nu\leq 0)+\left[\lambda_+\nu^{\alpha/3}+\sqrt{\varpi_+}B_2(\nu)\right]1(\nu>0)\to\infty,$$

as $\alpha/3\geq 1$.

(ii) $\left(\rho_n\left(\widehat{\gamma}-\gamma_0\right), \kappa_n\left(\widehat{\beta}-\beta_0\right)\right)=O_p(1)$. This is proved in Lemma 6.

Now, $\left(\rho_n\left(\widehat{\gamma}-\gamma_0\right), \kappa_n\left(\widehat{\beta}-\beta_0\right)\right)\xrightarrow{d}\arg\min_h\mathbb{M}(h)=\left(\arg\min_v\mathbb{M}_2(v),\mathbf{0}\right)$; in other words, $\kappa_n\left(\widehat{\beta}-\beta_0\right)$ is degenerate and the convergence rate of $\widehat{\beta}$ is faster. We first show that $\arg\min_v\mathbb{M}_2(v)$ can be simplified to the form as stated in the theorem. First,

$$\arg\min_v\ \mathbb{M}_2(v)=\arg\max_v\{-\mathbb{M}_2(v)\}\overset{d}{=}\arg\max_v\begin{cases}-\frac{1}{2}2\lambda_-|v|^\alpha+\sqrt{\varpi_-}B_1(-v^3), & \text{if } v\leq 0,\\ -\frac{1}{2}2\lambda_+v^\alpha+\sqrt{\varpi_+}B_2(v^3), & \text{if } v>0,\end{cases}=\widehat{\nu}^{1/3},$$

where

$$\widehat{\nu}=\arg\max_\nu\begin{cases}-\frac{1}{2}2\lambda_-|\nu|^{\alpha/3}+\sqrt{\varpi_-}B_1(-\nu), & \text{if } \nu\leq 0,\\ -\frac{1}{2}2\lambda_+\nu^{\alpha/3}+\sqrt{\varpi_+}B_2(\nu), & \text{if } \nu>0.\end{cases}$$

From Proposition 2(i), $\widehat{\nu}=\omega^{\frac{1}{2\tau-1}}\zeta(\varphi,\phi;\tau)$ with $\tau=\alpha/3$, $\omega=\varpi_-/\left(2\lambda_-\right)^2$, $\varphi=\lambda_+/\lambda_-$, and $\phi=\omega_0^+/\omega_0^-$.

We next show the asymptotic distribution of $\widehat{\beta}$. As in the proof of Theorem 4,

$$\begin{aligned}
&\widehat{\beta}-\beta_0\approx\arg\min_{u\in\mathcal{N}_u}\left\{S_n\left(\widehat{\gamma},\beta_0+u\right)-S_n\left(\gamma_0,\beta_0\right)\right\}\\
&=\arg\min_{u\in\mathcal{N}_u}\left\{\left[S\left(\widehat{\gamma},\beta_0+u\right)-S\left(\gamma_0,\beta_0\right)\right]+\left[S_n\left(\widehat{\gamma},\beta_0+u\right)-S\left(\widehat{\gamma},\beta_0+u\right)-n^{-1/2}\mathbb{G}_n\left(s\left(\cdot|\gamma_0,\beta_0\right)\right)\right]\right\}\\
&=\arg\min_{u\in\mathcal{N}_u}\left\{\tfrac{1}{2}u'S_{\beta\beta}u+\left[\tfrac{1}{2}u'S_{\beta\gamma^2}^-\left(\widehat{\gamma}-\gamma_0\right)^2+\lambda_-\left|\widehat{\gamma}-\gamma_0\right|^\alpha L_-\left(\left|\widehat{\gamma}-\gamma_0\right|\right)\right]1(\widehat{\gamma}\leq\gamma_0)\right.\\
&\quad+\left[\tfrac{1}{2}u'S_{\beta\gamma^2}^+\left(\widehat{\gamma}-\gamma_0\right)^2+\lambda_+\left(\widehat{\gamma}-\gamma_0\right)^\alpha L_+\left(\widehat{\gamma}-\gamma_0\right)\right]1(\widehat{\gamma}>\gamma_0)\\
&\quad\left.+n^{-1/2}\mathbb{G}_n\left(s\left(\cdot|\widehat{\gamma},\beta_0+u\right)\right)-n^{-1/2}\mathbb{G}_n\left(s\left(\cdot|\gamma_0,\beta_0\right)\right)+O\left(\kappa_n^{-2}\rho_n^{-1}\right)+o\left(\rho_n^{-\alpha}L\left(\rho_n^{-1}\right)\right)+o\left(\kappa_n^{-1}\rho_n^{-2}\right)\right\}\\
&=\arg\min_{u\in\mathcal{N}_u}\left\{\tfrac{1}{2}u'S_{\beta\beta}u+\tfrac{1}{2}u'S_{\beta\gamma^2}\left(\widehat{\gamma}-\gamma_0\right)^2-n^{-1/2}u'\mathbb{G}_n\begin{pmatrix}\mathbf{x}e_1 1\left(q\leq\gamma_0\right)\\ \mathbf{x}e_2 1\left(q>\gamma_0\right)\end{pmatrix}+o_p\left(n^{-1/2}\right)\right\}
\end{aligned}$$

$$= \left(\begin{array}{c} M_0^{-1}\left[-\frac{1}{2}S_{\beta_\ell\gamma^2}\left(\widehat{\gamma}-\gamma_0\right)^2+n^{-1/2}\mathbb{G}_n\left(\mathbf{x}e_1 1\left(q\le\gamma_0\right)\right)\right] \\ \overline{M}_0^{-1}\left[-\frac{1}{2}S_{\beta_\ell\gamma^2}\left(\widehat{\gamma}-\gamma_0\right)^2+n^{-1/2}\mathbb{G}_n\left(\mathbf{x}e_2 1\left(q>\gamma_0\right)\right)\right] \end{array} \right) + o_p\left(n^{-1/2}\right),$$

where $\mathcal{N}_u$ is a κ_n^{-1}-neighborhood of 0, $O\left(\kappa_n^{-2}\rho_n^{-1}\right)=o\left(n^{-1/2}\right)$, $o\left(\rho_n^{-\alpha}L\left(\rho_n^{-1}\right)\right)=o\left(n^{-1/2}\right)$, and $o\left(\kappa_n^{-1}\rho_n^{-2}\right)=o\left(n^{-1/2}\right)$.

In summary, the convergence rate of $\widehat{\beta}-\beta_0$ is $\min\left(\rho_n^2,n^{1/2}\right)$. When $3<\alpha<3.5$,

$$n^{1/2}\left(\widehat{\beta}-\beta_0\right)=\text{diag}\left\{M_0^{-1},\overline{M}_0^{-1}\right\}\mathbb{G}_n\left(\begin{array}{c} \mathbf{x}e_1 1\left(q\le\gamma_0\right) \\ \mathbf{x}e_2 1\left(q>\gamma_0\right) \end{array} \right)+o_p\left(1\right)\xrightarrow{d}\left(\begin{array}{c} Z_{\beta_1} \\ Z_{\beta_2} \end{array} \right),$$

and when $3.5<\alpha<4$,

$$\rho_n^2\left(\widehat{\beta}-\beta_0\right)=-\frac{1}{2}\left(\begin{array}{c} M_0^{-1}S_{\beta_\ell\gamma^2} \\ \overline{M}_0^{-1}S_{\beta_\ell\gamma^2} \end{array} \right)\rho_n^2\left(\widehat{\gamma}-\gamma_0\right)^2+o_p\left(1\right)\xrightarrow{d}-\frac{1}{2}\left(\begin{array}{c} M_0^{-1}S_{\beta_\ell\gamma^2} \\ \overline{M}_0^{-1}S_{\beta_\ell\gamma^2} \end{array} \right)Z_\gamma\left(\alpha\right)^2,$$

and when $\alpha=3.5$, the weak limit of $n^{1/2}\left(\widehat{\beta}-\beta_0\right)$ is the sum of both limits. ■

Proof of Corollary 7. Note that

$$\begin{aligned} &\sqrt{n\rho_n^3}\left(S_n(\gamma_0)-S_n\left(\widehat{\gamma}\right)\right)-\sqrt{n\rho_n^3}\left(S_n(\gamma_0,\beta_0)-S_n(\widehat{\gamma},\widehat{\beta})\right)=\sqrt{n\rho_n^3}\left(S_n(\gamma_0,\widehat{\beta}\left(\gamma_0\right))-S_n\left(\gamma_0,\beta_0\right)\right) \\ &=-\sqrt{n\rho_n^3}\left(\widehat{\beta}\left(\gamma_0\right)-\beta_0\right)'\tfrac{1}{n}\text{diag}\left\{X_{\le\gamma_0}'X_{\le\gamma_0},X_{>\gamma_0}'X_{>\gamma_0}\right\}\left(\widehat{\beta}\left(\gamma_0\right)-\beta_0\right) \\ &=\sqrt{n\rho_n^3}O_p\left(n^{-1/2}\right)O\left(1\right)O_p\left(n^{-1/2}\right)=o_p\left(1\right), \end{aligned}$$

where $\widehat{\beta}\left(\gamma_0\right)-\beta_0=O_p\left(n^{-1/2}\right)$, and $\rho_n\prec n^{1/3}$ so $\sqrt{n\rho_n^3}\prec n$. As a result,

$$\sqrt{n\rho_n^3}\left(S_n(\gamma_0)-S_n\left(\widehat{\gamma}\right)\right)=\sqrt{n\rho_n^3}\left(S_n(\gamma_0,\beta_0)-S_n(\widehat{\gamma},\widehat{\beta})\right)\xrightarrow{d}-\min_h\mathbb{M}\left(h\right)=-\min_v\mathbb{M}_2\left(v\right)=\max_v\left\{-\mathbb{M}_2\left(v\right)\right\}.$$

Now, by Proposition 2(ii),

$$\begin{aligned} \max_v\left\{-\mathbb{M}\left(v\right)\right\} &\overset{d}{=}\max_v\begin{cases} -\lambda_-\left|v\right|^\alpha+\sqrt{\varpi_-}B_1(-v^3), & \text{if } v\le 0, \\ -\lambda_+v^\alpha+\sqrt{\varpi_+}B_2(v^3), & \text{if } v>0, \end{cases} \\ &=\max_\nu\begin{cases} -\frac{1}{2}2\lambda_-\left|\nu\right|^{\alpha/3}+\sqrt{\varpi_-}B_1(-\nu), & \text{if } \nu\le 0, \\ -\frac{1}{2}2\lambda_+\nu^{\alpha/3}+\sqrt{\varpi_+}B_2(\nu), & \text{if } \nu>0, \end{cases}=\eta^{\frac{2}{2\tau-1}}\xi\left(\varphi,\phi;\tau\right), \end{aligned}$$

where $\eta^2=\frac{\varpi_-^\tau}{2\lambda_-}$ and $\tau=\frac{\alpha}{3}$. The required result follows by Slutsky's theorem. ■

Proof of Theorem 9. First, $\widehat{\theta}=\arg\min_\theta S_n\left(\theta\right)$ implies

$$\widehat{h}_n:=\left(\rho_n\left(\widehat{\gamma}-\gamma_0\right),\kappa_n\left(\widehat{\beta}-\beta_0\right)\right)=\arg\min_h\mathbb{M}_n\left(h\right)=\arg\min_{(v,u)}nP_n\left(s\left(\cdot\left|\gamma_0+\tfrac{v}{\rho_n},\beta_0+\tfrac{u}{\kappa_n}\right.\right)-s\left(\cdot\left|\gamma_0,\beta_0\right.\right)\right).$$

From Lemma 11,

$$\mathbb{M}_n\left(h\right)=\frac{1}{2}u_1'M_0u_1+\frac{1}{2}u_2'\overline{M}_0u_2+\lambda_-\left|v\right|^\alpha 1(v\le 0)+\lambda_+v^\alpha 1(v>0)-W_n\left(u\right)+o_P\left(1\right),$$

where $W_n\left(u\right)$ is defined in Lemma 8. Now, we apply the argmax continuous mapping theorem (Theorem 3.2.2 of VW) to derive the asymptotic distribution of $\left(\rho_n\left(\widehat{\gamma}-\gamma_0\right),\kappa_n\left(\widehat{\beta}-\beta_0\right)\right)$.

(i) $W_n(u) \rightsquigarrow u'W$, where W is defined in (9). This is proved in Lemma 12. So

$$\mathbb{M}_n(h) \rightsquigarrow \frac{1}{2}u_1' M_0 u_1 + \frac{1}{2}u_2' \overline{M}_0 u_2 + \lambda_- |v|^\alpha 1(v \le 0) + \lambda_+ v^\alpha 1(v > 0) - u'W =: \mathbb{M}(h).$$

(ii) $\left(\rho_n(\widehat{\gamma} - \gamma_0), \kappa_n\left(\widehat{\beta} - \beta_0\right)\right) = O_p(1)$. This is proved in Lemma 6.

(iii) That $\arg\min_h \mathbb{M}(h) = O_p(1)$ is obvious.

(iv) That $\arg\min_h \mathbb{M}(h)$ is unique is obvious.

Now, $\left(\rho_n(\widehat{\gamma} - \gamma_0), \kappa_n\left(\widehat{\beta} - \beta_0\right)\right) \xrightarrow{d} \arg\min_h \mathbb{M}(h) = \left(0, M_0^{-1}W_1, \overline{M}_0^{-1}W_2\right)$; in other words, $\rho_n(\widehat{\gamma} - \gamma_0)$ is degenerate and the convergence rate of $\widehat{\gamma}$ is faster. Lemma 7 shows that $n^{1/2}(\widehat{\gamma} - \gamma_0) = O_p(1)$ when $2 < \alpha \le 2.5$ and $\varrho_n(\widehat{\gamma} - \gamma_0) = O_p(1)$ when $2.5 < \alpha < 3$. We then apply Theorem 2.7 of KP to find the asymptotic distribution of $\min\left(n^{1/2}, \varrho_n\right)(\widehat{\gamma} - \gamma_0)$. First, when $2.5 < \alpha < 3$, the proof is the exactly the same as that of Theorem 8, just noting that $\alpha/3 > 5/6$, so $B(v)/|v|^{\alpha/3} \le \sqrt{2|v|\log\log|v|}/|v|^{5/6} \to 0$ as $|v| \to \infty$. Second, when $2 < \alpha < 2.5$, from Lemma 7,

$$\begin{aligned}\sqrt{n}(\widehat{\gamma} - \gamma_0) = \arg\min_v \frac{n^{\alpha/2}}{L\left(\frac{1}{n^{1/2}}\right)} \Bigg\{ & \left[\left(\widehat{\beta} - \beta_0\right)' S_\beta^-\left(\tfrac{v}{n^{1/2}}\right) + \lambda_- \left|\tfrac{v}{n^{1/2}}\right|^\alpha L_-\left(\left|\tfrac{v}{n^{1/2}}\right|\right)\right] 1(v \le 0) \\ & + \left[\left(\widehat{\beta} - \beta_0\right)' S_\beta^+\left(\tfrac{v}{n^{1/2}}\right) + \lambda_+ \left(\tfrac{v}{n^{1/2}}\right)^\alpha L_+\left(\tfrac{v}{n^{1/2}}\right)\right] 1(v > 0) + o_p(1) \Bigg\},\end{aligned}$$

where $\frac{n^{(\alpha-1)/2}}{L\left(\frac{1}{n^{1/2}}\right)} S_\beta^\pm\left(\frac{v}{n^{1/2}}\right) \to \psi_\pm |v|^{\alpha-1}$ and $\frac{n^{\alpha/2}}{L\left(\frac{1}{n^{1/2}}\right)} \lambda_\pm \left|\frac{v}{n^{1/2}}\right|^\alpha L_\pm\left(\left|\frac{v}{n^{1/2}}\right|\right) \to \lambda_\pm |v|^\alpha$ by Assumption (x)(ab), so by Theorem 2.7 of KP,

$$\begin{aligned}\sqrt{n}(\widehat{\gamma} - \gamma_0) &\xrightarrow{d} \arg\min_v \left\{ \left[Z_\beta' \psi_- |v|^{\alpha-1} + \lambda_- |v|^\alpha\right] 1(v \le 0) + \left[Z_\beta' \psi_+ v^{\alpha-1} + \lambda_+ v^\alpha\right] 1(v \ge 0) \right\} \\ &= \begin{cases} \frac{\alpha-1}{\alpha}\frac{\psi_-'}{\lambda_-} Z_\beta, & \text{if } \psi_-' Z_\beta < 0,\ \psi_+' Z_\beta < 0 \text{ and } \frac{\left|Z_\beta'\psi_-\right|^\alpha}{\lambda_-^{\alpha-1}} \ge \frac{\left|Z_\beta'\psi_+\right|^\alpha}{\lambda_+^{\alpha-1}}, \\ \frac{\alpha-1}{\alpha}\frac{\psi_-'}{\lambda_-} Z_\beta, & \text{if } \psi_-' Z_\beta < 0 \text{ and } \psi_+' Z_\beta \ge 0, \\ -\frac{\alpha-1}{\alpha}\frac{\psi_+'}{\lambda_+} Z_\beta, & \text{if } \psi_-' Z_\beta < 0,\ \psi_+' Z_\beta < 0 \text{ and } \frac{\left|Z_\beta'\psi_-\right|^\alpha}{\lambda_-^{\alpha-1}} < \frac{\left|Z_\beta'\psi_+\right|^\alpha}{\lambda_+^{\alpha-1}}, \\ -\frac{\alpha-1}{\alpha}\frac{\psi_+'}{\lambda_+} Z_\beta, & \text{if } \psi_-' Z_\beta \ge 0 \text{ and } \psi_+' Z_\beta < 0, \\ 0, & \text{if } \psi_-' Z_\beta \ge 0 \text{ and } \psi_+' Z_\beta \ge 0, \end{cases}\end{aligned}$$

where the minimum on $v \le 0$ is $-\frac{(\alpha-1)^{\alpha-1}}{\alpha^\alpha}\frac{\left(-Z_\beta'\psi_-\right)^\alpha}{\lambda_-^{\alpha-1}}$ and on $v \ge 0$ is $-\frac{(\alpha-1)^{\alpha-1}}{\alpha^\alpha}\frac{\left(-Z_\beta'\psi_+\right)^\alpha}{\lambda_+^{\alpha-1}}$ when the minimizer is achieved on $v < 0$ and $v > 0$, respectively. When $\alpha = 2.5$, from Lemma 7,

$$\sqrt{n}(\widehat{\gamma} - \gamma_0) = \arg\min_v \{\mathbb{M}_n(v) + o_p(1)\},$$

where $\mathbb{M}_n(v)$ is defined there. By Lemma 16,

$$\begin{aligned}\mathbb{M}_n(v) \rightsquigarrow \mathbb{M}(v) := & \left[Z_\beta'\psi_- |v|^{3/2} + \lambda_- |v|^{5/2} - Z_{\beta_2}' \Xi_1^-(v) + Z_{\beta_1}' \Xi_1^-(v) + \Xi_2^-(v)\right] 1(v \le 0) \\ & + \left[Z_\beta'\psi_+ v^{3/2} + \lambda_+ v^{5/2} - Z_{\beta_1}' \Xi_1^+(v) + Z_{\beta_2}' \Xi_1^+(v) + \Xi_2^+(v)\right] 1(v \ge 0),\end{aligned}$$

where $\Xi_1^\pm(v)$ and $\Xi_2^\pm(v)$ are defined in Lemma 16. The asymptotic distribution of $\sqrt{n}(\widehat{\gamma} - \gamma_0)$ is then achieved by Theorem 2.7 of KP. ■

Proof of Corollary 8. By the CMT,

$$
\begin{aligned}
& n\left(S_n\left(\gamma_0\right)-S_n(\widehat{\gamma})\right) \\
= \; & n\left[\left(S_n\left(\gamma_0, \widehat{\beta}\left(\gamma_0\right)\right)-S_n\left(\gamma_0, \beta_0\right)\right)-\left(S_n\left(\widehat{\gamma}, \widehat{\beta}\right)-S_n\left(\gamma_0, \beta_0\right)\right)\right] \\
& \xrightarrow{d} \min_u\left\{\frac{1}{2} u' S_{\beta\beta} u-u' W\right\}-\min_{u, v}\left\{\frac{1}{2} u' S_{\beta\beta} u-u' W+\lambda_{-}|v|^\alpha 1(v \leq 0)+\lambda_{+} v^\alpha 1(v>0)\right\} \\
= \; & 0,
\end{aligned}
$$

so we need a larger normalization rate than n. Note that

$$
S_n\left(\gamma_0, \widehat{\beta}\left(\gamma_0\right)\right)-S_n\left(\widehat{\gamma}, \widehat{\beta}\right)=\left[S_n\left(\gamma_0, \widehat{\beta}\left(\gamma_0\right)\right)-S_n\left(\gamma_0, \beta_0\right)\right]-\left[S_n\left(\widehat{\gamma}, \widehat{\beta}\right)-S_n\left(\gamma_0, \beta_0\right)\right],
$$

and we need to analyze these two terms separately. From the proof of Theorem 8,

$$
\begin{aligned}
\widehat{\beta}-\widehat{\beta}\left(\gamma_0\right) & =\left(\widehat{\beta}-\beta_0\right)-\left(\widehat{\beta}\left(\gamma_0\right)-\beta_0\right) \\
& =-S_{\beta\beta}^{-1} S_\beta^{-}\left(\widehat{\gamma}-\gamma_0\right) 1\left(\widehat{\gamma} \leq \gamma_0\right)-S_{\beta\beta}^{-1} S_\beta^{+}\left(\widehat{\gamma}-\gamma_0\right) 1\left(\widehat{\gamma}>\gamma_0\right)+O_p\left(n^{-1}\right) \\
& =O\left(\left(\widehat{\gamma}-\gamma_0\right)^{\alpha-1} L\left(\widehat{\gamma}-\gamma_0\right)\right)=o_p\left(n^{-1/2}\right) \text { but } \succ O_p\left(n^{-1}\right),
\end{aligned}
$$

where note that $u' S_\beta^{\pm}\left(\widehat{\gamma}-\gamma_0\right)$ rather than $\frac{1}{2} u' S_{\beta\gamma^2}\left(\widehat{\gamma}-\gamma_0\right)^2$ is the dominating term now, and the $O_p\left(n^{-1}\right)$ term in the second equality is from a careful analysis of the remainder term that is shown as $o_p\left(n^{-1/2}\right)$ in the proof of Theorem 8, so

$$
S_n\left(\gamma_0, \widehat{\beta}\left(\gamma_0\right)\right)-S_n\left(\gamma_0, \widehat{\beta}\right)=-\frac{1}{2}\left(\widehat{\beta}-\widehat{\beta}\left(\gamma_0\right)\right)' \operatorname{diag}\left(\frac{1}{n} X_{\leq \gamma_0}' X_{\leq \gamma_0}, \frac{1}{n} X_{>\gamma_0}' X_{>\gamma_0}\right)\left(\widehat{\beta}-\widehat{\beta}\left(\gamma_0\right)\right)=O_p\left(\left\|\widehat{\beta}-\widehat{\beta}\left(\gamma_0\right)\right\|^2\right).
$$

Because

$$
\begin{aligned}
S_n\left(\gamma_0, \widehat{\beta}\right)-S_n\left(\gamma_0, \beta_0\right)= & -\left(\widehat{\beta}-\beta_0\right)' \operatorname{diag}\left(\tfrac{1}{n} X_{\leq \gamma_0}' X_{\leq \gamma_0}, \tfrac{1}{n} X_{>\gamma_0}' X_{>\gamma_0}\right)\left(\widehat{\beta}\left(\gamma_0\right)-\beta_0\right) \\
& +\tfrac{1}{2}\left(\widehat{\beta}-\beta_0\right)' \operatorname{diag}\left(\tfrac{1}{n} X_{\leq \gamma_0}' X_{\leq \gamma_0}, \tfrac{1}{n} X_{>\gamma_0}' X_{>\gamma_0}\right)\left(\widehat{\beta}-\beta_0\right) \\
=-\tfrac{1}{2}\left(\widehat{\beta}-\beta_0\right)' S_{\beta\beta}\left(\widehat{\beta}-\beta_0\right) & +\left(\widehat{\beta}-\beta_0\right)' S_{\beta\beta}\left(\widehat{\beta}-\widehat{\beta}\left(\gamma_0\right)\right)+O_p\left(n^{-3/2}\right),
\end{aligned}
$$

we have

$$
\begin{aligned}
& S_n\left(\gamma_0, \widehat{\beta}\left(\gamma_0\right)\right)-S_n\left(\gamma_0, \beta_0\right) \\
& =S_n\left(\gamma_0, \widehat{\beta}\right)-S_n\left(\gamma_0, \beta_0\right)+O_p\left(\left\|\widehat{\beta}-\widehat{\beta}\left(\gamma_0\right)\right\|^2\right) \\
& =-\tfrac{1}{2}\left(\widehat{\beta}-\beta_0\right) S_{\beta\beta}\left(\widehat{\beta}-\beta_0\right)+\left(\widehat{\beta}-\beta_0\right)' S_{\beta\beta}\left(\widehat{\beta}-\widehat{\beta}\left(\gamma_0\right)\right)+O_p\left(n^{-3/2}\right)+O_p\left(\left\|\widehat{\beta}-\widehat{\beta}\left(\gamma_0\right)\right\|^2\right).
\end{aligned}
$$

Next, in $S_n\left(\widehat{\gamma}, \widehat{\beta}\right)-S_n\left(\gamma_0, \beta_0\right)$ of Lemma 7, the term

$$
\begin{aligned}
& \tfrac{1}{2}\left(\widehat{\beta}-\beta_0\right)' S_{\beta\beta}\left(\widehat{\beta}-\beta_0\right)+C\left(\widehat{\beta}-\beta_0\right) \\
& =\tfrac{1}{2}\left(\widehat{\beta}-\beta_0\right)' S_{\beta\beta}\left(\widehat{\beta}-\beta_0\right)-\left(\widehat{\beta}-\beta_0\right) \operatorname{diag}\left(\tfrac{1}{n} X_{\leq \gamma_0}' X_{\leq \gamma_0}, \tfrac{1}{n} X_{>\gamma_0}' X_{>\gamma_0}\right)\left(\widehat{\beta}\left(\gamma_0\right)-\beta_0\right)+O_p\left(n^{-3/2}\right) \\
& =-\tfrac{1}{2}\left(\widehat{\beta}-\beta_0\right)' S_{\beta\beta}\left(\widehat{\beta}-\beta_0\right)+\left(\widehat{\beta}-\beta_0\right)' S_{\beta\beta}\left(\widehat{\beta}-\widehat{\beta}\left(\gamma_0\right)\right)+O_p\left(n^{-3/2}\right).
\end{aligned}
$$

When $2.5 < \alpha < 3$, collecting the main terms (noting that $-\frac{1}{2}\left(\widehat{\beta}-\beta_0\right)' S_{\beta\beta}\left(\widehat{\beta}-\beta_0\right)$ is offsetted) in Lemma 7, we have

$$\sqrt{n\varrho_n^3}\left[S_n\left(\gamma_0,\widehat{\beta}\left(\gamma_0\right)\right)-S_n\left(\widehat{\gamma},\widehat{\beta}\right)\right]\xrightarrow{d}\max_v\left\{\begin{array}{ll}-\lambda_-\left|v\right|^\alpha+\Xi_2^-\left(\left|v\right|\right), & \text{if } v\le 0,\\ -\lambda_+v^\alpha+\Xi_2^+\left(v\right), & \text{if } v>0,\end{array}\right.=\eta^{\frac{6}{2\alpha-3}}\xi\left(\varphi,\phi;\tau\right),$$

where note that $\sqrt{n\varrho_n^3}n^{-1/2}\left\|\widehat{\beta}-\widehat{\beta}\left(\gamma_0\right)\right\|=o_p\left(1\right)$. When $2<\alpha<2.5$,

$$\begin{aligned}&\frac{n^{\alpha/2}}{L\left(\frac{1}{n^{1/2}}\right)}\left\{\left[S_n\left(\gamma_0,\widehat{\beta}\left(\gamma_0\right)\right)-S_n\left(\gamma_0,\beta_0\right)\right]-\left[\frac{1}{2}\left(\widehat{\beta}-\beta_0\right)'S_{\beta\beta}\left(\widehat{\beta}-\beta_0\right)+C\left(\widehat{\beta}-\beta_0\right)\right]\right\}\\ =&\frac{n^{\alpha/2}}{L\left(\frac{1}{n^{1/2}}\right)}\left\{O_p\left(\left\|\widehat{\beta}-\widehat{\beta}\left(\gamma_0\right)\right\|^2\right)+O_p\left(n^{-3/2}\right)\right\}=o_p\left(1\right),\end{aligned}$$

where $o_p\left(1\right)$ is due to $n^{\alpha/2}/L\left(\frac{1}{n^{1/2}}\right)n^{-1/2}\left\|\widehat{\beta}-\widehat{\beta}\left(\gamma_0\right)\right\|=O_p\left(1\right)$ and $O_p\left(n^{-1}\right)\prec\left\|\widehat{\beta}-\widehat{\beta}\left(\gamma_0\right)\right\|\prec O_p\left(n^{-1/2}\right)$. As a result,

$$\begin{aligned}&\frac{n^{\alpha/2}}{L\left(\frac{1}{n^{1/2}}\right)}\left[S_n\left(\gamma_0,\widehat{\beta}\left(\gamma_0\right)\right)-S_n\left(\widehat{\gamma},\widehat{\beta}\right)\right]\\ \xrightarrow{d}&-\left\{\left[Z_\beta'\psi_-\left|Z_\gamma\right|^{\alpha-1}+\lambda_-\left|Z_\gamma\right|^\alpha\right]1(Z_\gamma\le 0)+\left[Z_\beta'\psi_+Z_\gamma^{\alpha-1}+\lambda_+Z_\gamma^\alpha\right]1(Z_\gamma\ge 0)\right\}\\ =&\frac{(\alpha-1)^{\alpha-1}}{\alpha^\alpha}\frac{\left(-Z_\beta'\psi_-\right)^\alpha}{\lambda_-^{\alpha-1}}1(Z_\beta\in R_1)+\frac{(\alpha-1)^{\alpha-1}}{\alpha^\alpha}\frac{\left(-Z_\beta'\psi_+\right)^\alpha}{\lambda_+^{\alpha-1}}1(Z_\beta\in R_2).\end{aligned}$$

When $\alpha=2.5$, all terms remain and we have

$$n^{5/4}\left[S_n\left(\gamma_0,\widehat{\beta}\left(\gamma_0\right)\right)-S_n\left(\widehat{\gamma},\widehat{\beta}\right)\right]\xrightarrow{d}-\mathbb{M}\left(Z_\gamma\right),$$

where $\mathbb{M}\left(\cdot\right)$ is defined in the proof of Theorem 9. ■

Proof of Theorem 10. We concentrate on γ, i.e., we consider $S_n\left(\gamma\right)$. Note that

$$S_n\left(\gamma\right)=\frac{1}{2n}Y'\left(I-P_\gamma\right)Y,$$

where P_γ is the projection matrix on $X_\gamma^*=[X_{\le\gamma},X_{>\gamma}]$. Now, $\arg\min_\gamma S_n\left(\gamma\right)=\arg\max_\gamma S_n^*\left(\gamma\right)$, where $S_n^*\left(\gamma\right)=\frac{1}{n}Y'P_\gamma Y$. We can show by a GC theorem that

$$\begin{aligned}S_n^*\left(\gamma\right)&=\left(\frac{1}{n}Y'X_\gamma^*\right)\left(\frac{1}{n}X_\gamma^{*\prime}X_\gamma^*\right)^{-1}\left(\frac{1}{n}X_\gamma^*Y\right)\\ &\xrightarrow{p}\left(\mathbb{E}\left[y\mathbf{x}_{\le\gamma}'\right],\mathbb{E}\left[y\mathbf{x}_{>\gamma}'\right]\right)S_{\beta\beta}\left(\gamma\right)^{-1}\left(\begin{array}{c}\mathbb{E}\left[\mathbf{x}_{\le\gamma}y\right]\\ \mathbb{E}\left[\mathbf{x}_{>\gamma}y\right]\end{array}\right)\\ &=\mathbb{E}\left[y\mathbf{x}_{\le\gamma}'\right]\mathbb{E}\left[\mathbf{x}\mathbf{x}_{\le\gamma}\right]^{-1}\mathbb{E}\left[\mathbf{x}_{\le\gamma}y\right]+\mathbb{E}\left[y\mathbf{x}_{>\gamma}'\right]\mathbb{E}\left[\mathbf{x}\mathbf{x}_{>\gamma}\right]^{-1}\mathbb{E}\left[\mathbf{x}_{>\gamma}y\right]\\ &=\beta_{1\gamma}'\mathbb{E}\left[\mathbf{x}\mathbf{x}_{\le\gamma}\right]\beta_{1\gamma}+\beta_{2\gamma}'\mathbb{E}\left[\mathbf{x}\mathbf{x}_{>\gamma}\right]\beta_{2\gamma}=:S^*\left(\gamma\right)\end{aligned}$$

uniformly in $\gamma\in\Gamma$, where $S_{\beta\beta}\left(\gamma\right)=\text{diag}\{\mathbb{E}\left[\mathbf{x}\mathbf{x}_{\le\gamma}\right],\mathbb{E}\left[\mathbf{x}\mathbf{x}_{>\gamma}\right]\}$. From Assumption (viii), $\arg\max_\gamma S^*\left(\gamma\right)=\Gamma_o$. Because $S^*\left(\gamma\right)$ is continuous, Γ_o is a compact set. We denote the value of $S^*\left(\gamma\right)$ on Γ_o as S_0^*. Repeating the proof of Theorem 2.1 of Newey and McFadden (1994) and treating Γ_o as a point, we can show $P\left(\widehat{\gamma}\in\Gamma_o^\epsilon\right)\to 1$ for any $\epsilon>0$, where Γ_o^ϵ is the ϵ-enlargement of Γ_o.

To get the asymptotic distribution of $\widehat{\gamma}$ on Γ_o, we subtract S_0^* from $S_n^*(\gamma)$. Note that for $\gamma \in \Gamma_o$,

$$
\begin{aligned}
&\sqrt{n}\left[S_n^*(\gamma) - S_0^*\right] = \sqrt{n}\left[\tfrac{1}{n}Y'P_\gamma Y - \left(\mathbb{E}\left[y\mathbf{x}'_{\le\gamma}\right], \mathbb{E}\left[y\mathbf{x}'_{>\gamma}\right]\right) S_{\beta\beta}(\gamma)^{-1}\begin{pmatrix}\mathbb{E}\left[\mathbf{x}_{\le\gamma}y\right]\\ \mathbb{E}\left[\mathbf{x}_{>\gamma}y\right]\end{pmatrix}\right]\\
&= \left[\tfrac{1}{\sqrt{n}}\left(Y'X_\gamma^* - \left(\mathbb{E}\left[y\mathbf{x}'_{\le\gamma}\right], \mathbb{E}\left[y\mathbf{x}'_{>\gamma}\right]\right)\right)\right]\left(\tfrac{1}{n}X_\gamma^{*\prime}X_\gamma^*\right)^{-1}\left(\tfrac{1}{n}X_\gamma^* Y\right)\\
&\quad + \left(\mathbb{E}\left[y\mathbf{x}'_{\le\gamma}\right], \mathbb{E}\left[y\mathbf{x}'_{>\gamma}\right]\right) S_{\beta\beta}(\gamma)^{-1}\left[\tfrac{1}{\sqrt{n}}\left(X_\gamma^* Y - \begin{pmatrix}\mathbb{E}\left[\mathbf{x}_{\le\gamma}y\right]\\ \mathbb{E}\left[\mathbf{x}_{>\gamma}y\right]\end{pmatrix}\right)\right]\\
&\quad - \left(\mathbb{E}\left[y\mathbf{x}'_{\le\gamma}\right], \mathbb{E}\left[y\mathbf{x}'_{>\gamma}\right]\right)\left(\tfrac{1}{n}X_\gamma^{*\prime}X_\gamma^*\right)^{-1}\left[\tfrac{1}{\sqrt{n}}\left(X_\gamma^{*\prime}X_\gamma^* - S_{\beta\beta}(\gamma)\right)\right] S_{\beta\beta}(\gamma)^{-1}\left(\tfrac{1}{n}X_\gamma^* Y\right)\\
&\rightsquigarrow 2\beta_\gamma'\mathbb{B}_\gamma^{\mathbf{x}y} - \beta_\gamma'\mathbb{B}_\gamma^{\mathbf{xx}}\beta_\gamma,
\end{aligned}
$$

where the second equality is from the fact that for two invertible matrices A and B, $A^{-1} - B^{-1} = -A^{-1}(A-B)B^{-1}$, the weak convergence is by a Donsker's theorem, $\mathbb{B}_\gamma^{\mathbf{x}y} = \begin{pmatrix}\mathbb{B}_{1\gamma}^{\mathbf{x}y}\\ \mathbb{B}_{1\gamma}^{\mathbf{x}y}\end{pmatrix}$, $\mathbb{B}_\gamma^{\mathbf{xx}} = \begin{pmatrix}\mathbb{B}_{1\gamma}^{\mathbf{xx}} & \mathbf{0}\\ \mathbf{0} & \mathbb{B}_{2\gamma}^{\mathbf{xx}}\end{pmatrix}$, $\gamma \in \Gamma_o$, and $\left(\mathbb{B}_{1\gamma}^{\mathbf{x}y\prime}, \mathbb{B}_{2\gamma}^{\mathbf{x}y\prime}, vec\left(\mathbb{B}_{1\gamma}^{\mathbf{xx}}\right)', vec\left(\mathbb{B}_{2\gamma}^{\mathbf{xx}}\right)'\right)'$ is a $2(d+1)(d+2)$-dimensional zero-mean Gaussian process with the covariance kernel specified in the theorem. The asymptotic distribution of $\widehat{\gamma}$ is then achieved by Theorem 3.2.2 of VW. The asymptotic distribution of $\widehat{\beta}$ is completely determined by that of $\widehat{\gamma}$ because the perfect dependency of β_γ on γ.

If $\beta_{10} = \beta_{20} =: \beta_{\ell 0}$ on Γ_o, then because $\widehat{\beta}_{\ell\gamma} - \beta_{\ell 0}$ is $\sqrt{n}$-consistent, we have for any $\gamma_0 \in \Gamma_o$,

$$
\begin{aligned}
&n\left(S_n\left(\gamma, \beta_0 + \tfrac{u}{\sqrt{n}}\right) - S_n(\gamma_0, \beta_0)\right) = n\left(S_n\left(\gamma, \beta_0 + \tfrac{u}{\sqrt{n}}\right) - S_n(\gamma, \beta_0)\right) + n\left(S_n(\gamma, \beta_0) - S_n(\gamma_0, \beta_0)\right)\\
&= \sum_{i=1}^n T\left(w_i|\beta_{10} + \tfrac{u_1}{\sqrt{n}}, \beta_{10}\right)1(q_i \le \gamma) + \sum_{i=1}^n T\left(w_i|\beta_{20} + \tfrac{u_2}{\sqrt{n}}, \beta_{20}\right)1(q_i > \gamma)\\
&\quad + \sum_{i=1}^n \overline{z}_{1i}1(\gamma \wedge \gamma_0 < q \le \gamma_0) + \sum_{i=1}^n \overline{z}_{2i}1(\gamma_0 < q \le \gamma \vee \gamma_0)\\
&= -\sum_{i=1}^n\left(e_i - \tfrac{\mathbf{x}_i'}{2}\tfrac{u_1}{\sqrt{n}}\right)\mathbf{x}_i'\tfrac{u_1}{\sqrt{n}}1(q_i \le \gamma) - \sum_{i=1}^n\left(e_i - \tfrac{\mathbf{x}_i'}{2}\tfrac{u_2}{\sqrt{n}}\right)\mathbf{x}_i'\tfrac{u_2}{\sqrt{n}}1(q_i > \gamma)\\
&= \tfrac{1}{2}u_1'\widehat{M}_\gamma u_1 - u_1'\mathbb{G}_n\left(\mathbf{x}e1(q \le \gamma)\right) + \tfrac{1}{2}u_2'\widehat{\overline{M}}_\gamma u_2 - u_2'\mathbb{G}_n\left(\mathbf{x}e1(q > \gamma)\right),
\end{aligned}
$$

where $e_i = y_i - \mathbf{x}_i'\beta_{\ell 0} = m(x_i, q_i) - \mathbf{x}_i'\beta_{\ell 0} + \varepsilon_i$, $\mathbb{E}\left[\mathbf{x}_i e_i 1(q_i \le \gamma)\right] = \mathbb{E}\left[\mathbf{x}_i y_i 1(q_i \le \gamma)\right] - \mathbb{E}\left[\mathbf{x}_i\mathbf{x}_i' 1(q_i \le \gamma)\right]\beta_{\ell 0} = \mathbf{0}$ for $\gamma \in \Gamma_o$, and similarly $\mathbb{E}\left[\mathbf{x}_i e_i 1(q_i > \gamma)\right] = 0$ for $\gamma \in \Gamma_o$, $\overline{z}_{\ell i} = 0$ because $\delta_0 = \mathbf{0}$, $\widehat{M}_\gamma = n^{-1}\sum_{i=1}^n \mathbf{x}_i\mathbf{x}_i' 1(q_i \le \gamma)$, and $\widehat{\overline{M}}_\gamma = n^{-1}\sum_{i=1}^n \mathbf{x}_i\mathbf{x}_i' 1(q_i > \gamma)$. Concentrating out u, we have

$$
\begin{aligned}
&n\left(S_n\left(\gamma, \beta_0 + \tfrac{\widehat{u}(\gamma)}{\sqrt{n}}\right) - S_n(\gamma, \beta_0)\right)\\
&= -\tfrac{1}{2}\mathbb{G}_n\left(\mathbf{x}'e1(q \le \gamma)\right)\widehat{M}_\gamma^{-1}\mathbb{G}_n\left(\mathbf{x}e1(q \le \gamma)\right) - \tfrac{1}{2}\mathbb{G}_n\left(\mathbf{x}'e1(q > \gamma)\right)\widehat{\overline{M}}_\gamma^{-1}\mathbb{G}_n\left(\mathbf{x}e1(q > \gamma)\right)\\
&\rightsquigarrow -\tfrac{1}{2}\mathbb{B}_{1\gamma}^{\mathbf{x}e\prime}M_\gamma^{-1}\mathbb{B}_{1\gamma}^{\mathbf{x}e} - \tfrac{1}{2}\mathbb{B}_{2\gamma}^{\mathbf{x}e\prime}\overline{M}_\gamma^{-1}\mathbb{B}_{2\gamma}^{\mathbf{x}e},
\end{aligned}
$$

where $\widehat{u}(\gamma) = \left(\widehat{u}_1(\gamma)', \widehat{u}_2(\gamma)'\right)'$ with $\widehat{u}_1(\gamma) = \widehat{M}_\gamma^{-1}\mathbb{G}_n\left(\mathbf{x}e1(q \le \gamma)\right)$ and $\widehat{u}_2(\gamma) = \widehat{\overline{M}}_\gamma^{-1}\mathbb{G}_n\left(\mathbf{x}e1(q > \gamma)\right)$, $\mathbb{B}_{1\gamma}^{\mathbf{x}e}$ is a zero-mean Gaussian process with the covariance kernel equal to $\mathbb{E}\left[\mathbf{xx}'_{\le\gamma_1\wedge\gamma_2}e^2\right]$, and $\mathbb{B}_{2\gamma}^{\mathbf{x}e} = \mathbb{B}_{1\infty}^{\mathbf{x}e} - \mathbb{B}_{1\gamma}^{\mathbf{x}e}$. Note that Assumption III(v-vi) implies $\mathbb{E}\left[\mathbf{xx}'_{\le\gamma_1\wedge\gamma_2}e^2\right] > 0$. Applying Theorem 3.2.2 of VW, we have

$$
\widehat{\gamma} \xrightarrow{d} \arg\max_{\gamma\in\Gamma_o}\left\{\mathbb{B}_{1\gamma}^{\mathbf{x}e\prime}M_\gamma^{-1}\mathbb{B}_{1\gamma}^{\mathbf{x}e} + \mathbb{B}_{2\gamma}^{\mathbf{x}e\prime}\overline{M}_\gamma^{-1}\mathbb{B}_{2\gamma}^{\mathbf{x}e}\right\} =: \arg\max_{\gamma\in\Gamma_o}\widetilde{\Xi}(\gamma) =: \widetilde{Z}_\gamma.
$$

Given the asymptotic distribution of $\widehat{\gamma}$, we can apply a GC theorem and Donsker's theorem to show

$$\sqrt{n}\left(\widehat{\beta}_1 - \beta_{\ell 0}\right) = \widehat{u}_1(\widehat{\gamma}) \xrightarrow{d} M_{\widetilde{Z}_\gamma}^{-1}\mathbb{B}_{1\widetilde{Z}_\gamma}^{\mathbf{x}e} \text{ and } \sqrt{n}\left(\widehat{\beta}_2 - \beta_{\ell 0}\right) = \widehat{u}_2(\widehat{\gamma}) \xrightarrow{d} \overline{M}_{\widetilde{Z}_\gamma}^{-1}\mathbb{B}_{2\widetilde{Z}_\gamma}^{\mathbf{x}e},$$

where $\widetilde{Z}_\gamma$ should be replaced by Z_γ when $\beta_{10} \neq \beta_{20}$. ■

Proof of Corollary 9. Note that

$$\begin{aligned}
LR_n(\gamma_0) &= \sqrt{n}\left(S_n^*(\widehat{\gamma}) - S_n^*(\gamma_0)\right) = \sqrt{n}\left[\left(S_n^*(\widehat{\gamma}) - S_0^*\right) - \left(S_n^*(\gamma_0) - S_0^*\right)\right] \\
&= \sqrt{n}\left[\max_{\gamma\in\Gamma_o}\left(S_n^*(\gamma) - S_0^*\right) - \left(S_n^*(\gamma_0) - S_0^*\right)\right] \text{ with probability approaching 1} \\
&\xrightarrow{d} \max_{\gamma\in\Gamma_o}\left\{2\beta_\gamma'\mathbb{B}_\gamma^{\mathbf{x}y} - \beta_\gamma'\mathbb{B}_\gamma^{\mathbf{xx}}\beta_\gamma\right\} - \left(2\beta_{\gamma_0}'\mathbb{B}_{\gamma_0}^{\mathbf{x}y} - \beta_{\gamma_0}'\mathbb{B}_{\gamma_0}^{\mathbf{xx}}\beta_{\gamma_0}\right),
\end{aligned}$$

where the third equality is because $P(\widehat{\gamma} \in \Gamma_o^\epsilon) \to 1$ for any $\epsilon > 0$. Note also that when $\gamma \notin \Gamma_o$,

$$LR_n(\gamma) = \sqrt{n}\left[\max_{\gamma\in\Gamma_o}\left(S_n^*(\gamma) - S_0^*\right) - \left(S_n^*(\gamma) - S^*(\gamma)\right) - \left(S^*(\gamma) - S_0^*\right)\right] \to \infty$$

by $S^*(\gamma) - S_0^* < 0$.

If $\beta_{10} = \beta_{20} =: \beta_{\ell 0}$ on Γ_o, the normalization rate of LR_n should be n. Specifically, by the CMT,

$$\begin{aligned}
LR_n(\gamma_0) &= 2n\left(S_n\left(\gamma_0, \widehat{\beta}(\gamma_0)\right) - S_n\left(\widehat{\gamma}, \widehat{\beta}\right)\right) \\
&= -2n\left[\left(S_n\left(\widehat{\gamma}, \widehat{\beta}\right) - S_n(\gamma_0, \beta_0)\right) - \left(S_n\left(\gamma_0, \widehat{\beta}(\gamma_0)\right) - S_n(\gamma_0, \beta_0)\right)\right] \\
&\xrightarrow{d} \max_{\gamma\in\Gamma_o} \widetilde{\Xi}(\gamma) - \widetilde{\Xi}(\gamma_0).
\end{aligned}$$

Similarly, when $\gamma \notin \Gamma_o$, $LR_n(\gamma) \to \infty$ because $S_n\left(\gamma, \widehat{\beta}(\gamma)\right) - S_n(\gamma_0, \beta_0)$ will converge to a positive value as assumed in Assumption (viii). ■

Appendix B: Consistency and Convergence Rate

We collect the proofs for consistency and convergence rates in this appendix and intend to give a uniform treatment for all cases. We apply Theorem 2.1 of Newey and McFadden (1994) to prove the consistency; we apply Corollary 3.2.6 of VW to show the convergence rate except the shrinking thresholds case in Section 4.2 where we apply the proof idea of Theorem 3.2.5 in VW. We will detail the proof for I(1) and then adjust it for other cases.

Lemma 1 *Under Assumption I(α), $1 \leq \alpha \leq 2$ or II(α), $2 \leq \alpha \leq 4$, $\widehat{\theta} - \theta_0 = o_p(1)$.*

Proof. We first show the consistency of $\widehat{\gamma}$. It is not hard to show $\sup_{\gamma\in\Gamma}|S_n(\gamma) - S(\gamma)| \xrightarrow{p} 0$ by a GC theorem, where $S(\gamma) = \frac{1}{2}\mathbb{E}\left[\left(y - \mathbf{x}'\beta_{1\gamma}1(q \leq \gamma) - \mathbf{x}'\beta_{2\gamma}1(q > \gamma)\right)^2\right]$. By assumption (vii), β_γ is continuous in γ, and $S(\gamma)$ is continuous in γ. By by assumption (viii), $\arg\min_{\gamma\in\Gamma} S(\gamma)$ is unique. So from Theorem 2.1 of Newey and McFadden (1994), $\widehat{\gamma} \xrightarrow{p} \gamma_0$. Given the consistency of $\widehat{\gamma}$, $\widehat{\beta} \xrightarrow{p} \beta_0$ because $\sup_{\gamma\in\Gamma}\left|\widehat{\beta}(\gamma) - \beta_\gamma\right| \xrightarrow{p} 0$ and β_γ is continuous at γ_0. ■

Lemma 2 *Under Assumption I(1), $n(\widehat{\gamma} - \gamma_0) = O_p(1)$ and $n^{1/2}\left(\widehat{\beta} - \beta_0\right) = O_p(1)$.*

Proof. First, $S(\theta) - S(\theta_0) \geq Cd^2(\theta, \theta_0)$ with $d(\theta, \theta_0) = \|\beta - \beta_0\| + \sqrt{|\gamma - \gamma_0|}$ for θ in a neighborhood of θ_0. Extending the analyses in Section 3, we can show $\Psi_{\pm}(\beta, \gamma)$ in $S(\theta) - S(\theta_0)$ are dominated, so

$$\begin{aligned} S(\theta) - S(\theta_0) &\approx \Phi(\beta_1) + \overline{\Phi}(\beta_2) + \Lambda_+(\gamma) 1(\gamma > \gamma_0) + \Lambda_-(\gamma) 1(\gamma \leq \gamma_0) \\ &\geq C\left(\|\beta_1 - \beta_{10}\|^2 + \|\beta_2 - \beta_{20}\|^2 + |\gamma - \gamma_0|\right). \end{aligned}$$

Second, $\mathbb{E}\left[\sup_{d(\theta,\theta_0)<\eta} |\mathbb{G}_n(s(w|\theta) - s(w|\theta_0))|\right] \leq C\eta$ for any sufficiently small η. It is not hard to see that $\{s(w|\theta) - s(w|\theta_0) : d(\theta, \theta_0) < \eta\}$ is a VC-subgraph class of functions, so from Theorem 2.14.1 of VW,

$$\mathbb{E}\left[\sup_{d(\theta,\theta_0)<\eta} |\mathbb{G}_n(s(w|\theta) - s(w|\theta_0))|\right] \leq C\sqrt{PF^2},$$

where F is the envelope of $\{s(w|\theta) - s(w|\theta_0) : d(\theta, \theta_0) < \eta\}$. We can set F as

$$\begin{aligned} F = &\left(|m_1(x,q) - \mathbf{x}'\beta_{10}| \|\mathbf{x}\| \eta + \tfrac{1}{2}\eta^2 \|\mathbf{x}\|^2 + \eta \|\mathbf{x}\varepsilon_1\|\right) 1\left(q \leq \gamma_0 + \eta^2\right) \\ &+ \left(|m_2(x,q) - \mathbf{x}'\beta_{20}| \|\mathbf{x}\| \eta + \tfrac{1}{2}\eta^2 \|\mathbf{x}\|^2 + \eta \|\mathbf{x}\varepsilon_2\|\right) 1\left(q > \gamma_0 - \eta^2\right) \\ &+ \left(|m_1(x,q) - \mathbf{x}'\beta_{10}| \|\mathbf{x}\| (\|\delta_0\| + \eta) + (\|\delta_0\| + \eta)^2 \|\mathbf{x}\|^2 + (\|\delta_0\| + \eta) \|\mathbf{x}\varepsilon_1\|\right) 1\left(\gamma_0 - \eta^2 < q \leq \gamma_0\right) \\ &+ \left(|m_2(x,q) - \mathbf{x}'\beta_{20}| \|\mathbf{x}\| (\|\delta_0\| + \eta) + (\|\delta_0\| + \eta)^2 \|\mathbf{x}\|^2 + (\|\delta_0\| + \eta) \|\mathbf{x}\varepsilon_2\|\right) 1\left(\gamma_0 < q \leq \gamma_0 + \eta^2\right) \\ =: &F_1(\mathbf{x}, \varepsilon_1|\eta) + F_2(\mathbf{x}, \varepsilon_2|\eta) + F_3(\mathbf{x}, \varepsilon_1|\eta) + F_4(\mathbf{x}, \varepsilon_2|\eta). \end{aligned}$$

where η^2 rather than η appear in the index functions for q since the metric for γ is $\sqrt{|\gamma - \gamma_0|}$. By Assumptions (iv), $\sqrt{PF^2} = \sqrt{C(\eta^2 + \eta^2)} \leq C\eta$ for $\eta < 1$, i.e., the four terms of F are balanced. So $\phi(\eta) = \eta$ in Corollary 3.2.6 of VW and η/η^{ϱ} is decreasing for all $1 < \varrho < 2$. Since $r_n^2 \phi\left(\frac{1}{r_n}\right) = \sqrt{n}$ for $r_n = \sqrt{n}$, $\sqrt{n} d\left(\widehat{\theta} - \theta_0\right) = O_P(1)$. By the definition of d, the result follows. ■

Lemma 3 *Under Assumption I(1)′, $\widehat{\gamma} - \gamma_0 = o_p(1)$ and $\widehat{\beta} - \beta_0 = o_p(\|\delta_n\|)$.*

Proof. First, as in Lemma A.5 of Hansen (2000),

$$\begin{aligned} \|\delta_n\|^{-2}[S_n(\gamma) - S_n(\theta_0)] &= \|\delta_n\|^{-2}\left[\frac{1}{2n} Y'(I - P_\gamma) Y - \frac{1}{2n}\mathbf{e}'\mathbf{e}\right] \\ &= -\frac{1}{2\|\delta_n\|^2} \frac{1}{n}\mathbf{e}' P_\gamma \mathbf{e} + \frac{\delta_n'}{\|\delta_n\|} \frac{1}{n\|\delta_n\|} X_0'(I - P_\gamma)\mathbf{e} + \frac{1}{2}\frac{\delta_n'}{\|\delta_n\|}\frac{1}{n} X_0'(I - P_\gamma) X_0 \frac{\delta_n}{\|\delta_n\|} \end{aligned}$$

where P_γ is the projection matrix on $X_\gamma^* = [X, X_{\leq\gamma}]$ with X being the matrix stacking $\mathbf{x}_i$, $X_0 = X_{\leq\gamma_0}$ and $\mathbf{e}$ is the vector stacking $e_i = e_{1i} 1(q_i \leq \gamma_0) + e_{2i} 1(q_i > \gamma_0)$. Although

$$n^{-1/2} X'\mathbf{e} = n^{-1/2} X_{\leq\gamma_0}' \mathbf{e}_1 + n^{-1/2} X_{>\gamma_0}' \mathbf{e}_2 = O_p(1),$$

when $\gamma < \gamma_0$,

$$\begin{aligned} n^{-1} X_{\leq\gamma}' \mathbf{e} &= n^{-1}\sum_{i=1}^n \mathbf{x}_i \varepsilon_{1i} 1(q_i \leq \gamma) + n^{-1}\sum_{i=1}^n \mathbf{x}_i (m_1(x_i, q_i) - \mathbf{x}_i'\beta_{10}) 1(q_i \leq \gamma) \\ &= O_p\left(n^{-1/2}\right) + \mathbb{E}[\mathbf{x}(m_1(x,q) - \mathbf{x}'\beta_{10}) 1(q \leq \gamma)] \end{aligned}$$

and when $\gamma \geq \gamma_0$

$$
\begin{aligned}
n^{-1}X'_{\leq\gamma}\mathbf{e} &= n^{-1}\sum_{i=1}^{n}\mathbf{x}_i e_{1i}1(q_i\leq\gamma_0)+n^{-1}\sum_{i=1}^{n}\mathbf{x}_i\left(\varepsilon_{2i}+m_2\left(x_i,q_i\right)-\mathbf{x}_i'\beta_{20}\right)1(\gamma_0<q_i\leq\gamma) \\
&= O_p\left(n^{-1/2}\right)+\mathbb{E}\left[\mathbf{x}\left(m_2\left(x,q\right)-\mathbf{x}'\beta_{20}\right)1(\gamma_0<q\leq\gamma)\right]
\end{aligned}
$$

where $\mathbf{e}_\ell$ is the vector stacking $e_{\ell i}$. As a result, uniformly in γ,

$$
\begin{aligned}
-\frac{1}{2\left\|\delta_n\right\|^2}\frac{1}{n}\mathbf{e}'P_\gamma\mathbf{e} &= -\frac{1}{2}\left(\frac{1}{a_n^{1/2}}n^{-1/2}\mathbf{e}'X,\frac{1}{n\left\|\delta_n\right\|}\mathbf{e}'X_{\leq\gamma}\right)\left(\frac{1}{n}X_\gamma^{*\prime}X_\gamma^*\right)^{-1}\begin{pmatrix}\frac{1}{a_n^{1/2}}n^{-1/2}X'\mathbf{e}\\ \frac{1}{n\|\delta_n\|}X_{\leq\gamma}\mathbf{e}\end{pmatrix} \\
&\xrightarrow{p}\begin{cases}-\frac{1}{2}\left(\mathbf{0}',\varsigma_1\left(\gamma\right)'\right)\begin{pmatrix}M & M_\gamma\\ M_\gamma & M_\gamma\end{pmatrix}^{-1}\begin{pmatrix}\mathbf{0}\\ \varsigma_1\left(\gamma\right)\end{pmatrix}, & \text{if } \gamma<\gamma_0,\\ -\frac{1}{2}\left(\mathbf{0}',\Delta\varsigma_2\left(\gamma\right)'\right)\begin{pmatrix}M & M_\gamma\\ M_\gamma & M_\gamma\end{pmatrix}^{-1}\begin{pmatrix}\mathbf{0}\\ \Delta\varsigma_2\left(\gamma\right)\end{pmatrix}, & \text{if } \gamma\geq\gamma_0,\end{cases}
\end{aligned}
$$

where $\Delta\varsigma_2\left(\gamma\right)=\varsigma_2\left(\gamma_0\right)-\varsigma_2\left(\gamma\right)=\lim_{n\to\infty}\mathbb{E}\left[\mathbf{x}\left(\frac{m_2(x,q)-\mathbf{x}'\beta_{20}}{\|\delta_n\|}\right)1(\gamma_0<q\leq\gamma)\right]$ and $\varsigma_\ell\left(\gamma\right)$ is defined in Assumption (viii). Since

$$
\begin{pmatrix}M & M_\gamma\\ M_\gamma & M_\gamma\end{pmatrix}^{-1}=\begin{pmatrix}\overline{M}_\gamma^{-1} & -\overline{M}_\gamma^{-1}\\ -\overline{M}_\gamma^{-1} & M_\gamma^{-1}+\overline{M}_\gamma^{-1}\end{pmatrix},
$$

we have

$$
\left(\mathbf{0}',\varsigma_\ell\left(\gamma\right)'\right)\begin{pmatrix}M & M_\gamma\\ M_\gamma & M_\gamma\end{pmatrix}^{-1}\begin{pmatrix}\mathbf{0}\\ \varsigma_\ell\left(\gamma\right)\end{pmatrix}=\varsigma_\ell\left(\gamma\right)'\left[M_\gamma^{-1}+\overline{M}_\gamma^{-1}\right]\varsigma_\ell\left(\gamma\right).
$$

Similarly, uniformly in γ,

$$
\begin{aligned}
\frac{1}{n\left\|\delta_n\right\|}X_0'\left(I-P_\gamma\right)\mathbf{e} &\xrightarrow{p}\begin{cases}\left(M_0,M_{\gamma\wedge\gamma_0}\right)\begin{pmatrix}M & M_\gamma\\ M_\gamma & M_\gamma\end{pmatrix}^{-1}\begin{pmatrix}\mathbf{0}\\ \varsigma_1\left(\gamma\right)\end{pmatrix}, & \text{if } \gamma<\gamma_0,\\ \left(M_0,M_{\gamma\wedge\gamma_0}\right)\begin{pmatrix}M & M_\gamma\\ M_\gamma & M_\gamma\end{pmatrix}^{-1}\begin{pmatrix}\mathbf{0}\\ \Delta\varsigma_2\left(\gamma\right)\end{pmatrix}, & \text{if } \gamma\geq\gamma_0,\end{cases} \\
&= \begin{cases}\left[M_{\gamma\wedge\gamma_0}M_\gamma^{-1}+\left(M_{\gamma\wedge\gamma_0}-M_0\right)\overline{M}_\gamma^{-1}\right]\varsigma_1\left(\gamma\right), & \text{if } \gamma<\gamma_0,\\ \left[M_{\gamma\wedge\gamma_0}M_\gamma^{-1}+\left(M_{\gamma\wedge\gamma_0}-M_0\right)\overline{M}_\gamma^{-1}\right]\Delta\varsigma_2\left(\gamma\right), & \text{if } \gamma\geq\gamma_0,\end{cases} \\
&= \begin{cases}\left[I+\left(M_\gamma-M_0\right)\overline{M}_\gamma^{-1}\right]\varsigma_1\left(\gamma\right), & \text{if } \gamma<\gamma_0,\\ M_0M_\gamma^{-1}\Delta\varsigma_2\left(\gamma\right), & \text{if } \gamma\geq\gamma_0.\end{cases}
\end{aligned}
$$

Finally, $\frac{1}{n}X_0'\left(I-P_\gamma\right)X_0$ has the probability limit as in Lemma A.5 of Hansen (2000), i.e., $M_0-M_0M_\gamma^{-1}M_0$. If assume $\frac{\delta_n}{\|\delta_n\|}\to c$, then

$$
\begin{aligned}
&\left\|\delta_n\right\|^{-2}\left[S_n\left(\gamma\right)-S_n\left(\theta_0\right)\right]\xrightarrow{p}S\left(\gamma\right)=:\tfrac{1}{2}c'\left(M_0-M_0M_\gamma^{-1}M_0\right)c \\
&\quad+\begin{cases}\left\{c'\left[I+\left(M_\gamma-M_0\right)\overline{M}_\gamma^{-1}\right]-\frac{1}{2}\varsigma_1\left(\gamma\right)'\left[M_\gamma^{-1}+\overline{M}_\gamma^{-1}\right]\right\}\varsigma_1\left(\gamma\right), & \text{if } \gamma<\gamma_0,\\ \left\{c'M_0M_\gamma^{-1}-\frac{1}{2}\Delta\varsigma_2\left(\gamma\right)'\left[M_\gamma^{-1}+\overline{M}_\gamma^{-1}\right]\right\}\Delta\varsigma_2\left(\gamma\right), & \text{if } \gamma\geq\gamma_0.\end{cases}
\end{aligned}
$$

Although the minimizer of $\frac{1}{2}c'\left(M_0-M_0M_\gamma^{-1}M_0\right)c$ is γ_0, it is not obvious that the minimizer of $S\left(\gamma\right)$ is γ_0; that is why it is assumed in Assumption (viii). Note that since $\Delta\varsigma_2\left(\gamma_0\right)=0$, $S\left(\gamma_0\right)=0$.

The above analysis shows $\sup_{\gamma\in\Gamma}|S_n(\gamma)-S(\gamma)|\xrightarrow{p}0$, so Assumption I(1)′(viii) implies $\widehat{\gamma}-\gamma_0=o_p(1)$. Now we show $\widehat{\beta}-\beta_0=o_p(\|\delta_n\|)$. Note that

$$\begin{aligned}
\frac{1}{\|\delta_n\|}\begin{pmatrix}\widehat{\beta}_2-\beta_{20}\\ \widehat{\delta}-\delta_n\end{pmatrix} &= \frac{1}{\|\delta_n\|}\left(X_{\widehat{\gamma}}^{*\prime}X_{\widehat{\gamma}}^{*}\right)^{-1}X_{\widehat{\gamma}}^{*\prime}\left(Y-X_{\widehat{\gamma}}^{*}\left(\beta_{20}',\delta_n'\right)'\right)\\
&= \frac{1}{\|\delta_n\|}\left(\frac{1}{n}X_{\widehat{\gamma}}^{*\prime}X_{\widehat{\gamma}}^{*}\right)^{-1}\frac{1}{n}X_{\widehat{\gamma}}^{*\prime}\left(X_{\gamma_0}^{*}\left(\beta_{20}',\delta_n'\right)'+\mathbf{e}-X_{\widehat{\gamma}}^{*}\left(\beta_{20}',\delta_n'\right)'\right)\\
&= \left(\frac{1}{n}X_{\widehat{\gamma}}^{*\prime}X_{\widehat{\gamma}}^{*}\right)^{-1}\left[\frac{1}{n}X_{\widehat{\gamma}}^{*\prime}\left(X_{\le\gamma_0}-X_{\le\widehat{\gamma}}\right)\frac{\delta_n}{\|\delta_n\|}+\frac{1}{\sqrt{n}\|\delta_n\|}\frac{1}{\sqrt{n}}X_{\widehat{\gamma}}^{*\prime}\mathbf{e}\right].
\end{aligned}$$

Now, by the stochastic equicontinuity of $\gamma\mapsto\mathbb{G}_n(\mathbf{x}e1(q\le\gamma))$, we have that $\mathbb{G}_n(\mathbf{x}e1(q\le\widehat{\gamma}))=\mathbb{G}_n(\mathbf{x}e1(q\le\gamma_0))+o_p(1)$, so

$$\frac{1}{\sqrt{n}}X_{\le\widehat{\gamma}}'\mathbf{e}=\mathbb{E}\left[\mathbf{x}e1(q\le\widehat{\gamma})\right]+\mathbb{G}_n(\mathbf{x}e1(q\le\gamma_0))+o_p(1)=\frac{1}{\sqrt{n}}X_{\le\gamma_0}'\mathbf{e}+o_p(1)$$

which implies

$$\tfrac{1}{\sqrt{n}}X_{\widehat{\gamma}}^{*\prime}\mathbf{e}=\tfrac{1}{\sqrt{n}}\begin{pmatrix}X'\mathbf{e}\\ X_{\le\widehat{\gamma}}'\mathbf{e}\end{pmatrix}=\begin{pmatrix}\tfrac{1}{\sqrt{n}}X_{\le\gamma_0}'\mathbf{e}+\tfrac{1}{\sqrt{n}}X_{>\gamma_0}'\mathbf{e}\\ \tfrac{1}{\sqrt{n}}X_{\le\gamma_0}'\mathbf{e}+o_p(1)\end{pmatrix}=\tfrac{1}{\sqrt{n}}X_{\gamma_0}^{*}\mathbf{e}+o_p(1)=O_p(1),$$

where $\mathbb{E}\left[\mathbf{x}e1(q\le\widehat{\gamma})\right]=o_p(1)$ because $\widehat{\gamma}-\gamma_0=o_p(1)$ and $\mathbb{E}\left[\mathbf{x}e1(q\le\gamma_0)\right]=\mathbf{0}$, and $\frac{1}{\sqrt{n}}X_{\gamma_0}^{*}\mathbf{e}=O_p(1)$ by the CLT. Since $\sqrt{n}\|\delta_n\|\to\infty$, $\frac{1}{\sqrt{n}\|\delta_n\|}\frac{1}{\sqrt{n}}X_{\widehat{\gamma}}^{*}\mathbf{e}=o_p(1)$. By the continuity of M_γ and the consistency of $\widehat{\gamma}$,

$$\frac{1}{n}X_{\widehat{\gamma}}^{*}\left(X_{\le\gamma_0}-X_{\le\widehat{\gamma}}\right)\xrightarrow{p}\begin{pmatrix}M_0\\ M_0\end{pmatrix}-\begin{pmatrix}M_0\\ M_0\end{pmatrix}=\mathbf{0}$$

and

$$\frac{1}{n}X_{\widehat{\gamma}}^{*\prime}X_{\widehat{\gamma}}^{*}\xrightarrow{p}\begin{pmatrix}M & M_0\\ M_0 & M_0\end{pmatrix}.$$

The determinant of the limit matrix is $|M_0|\left|\overline{M}_0\right|>0$ by Assumption MA(v), so by the CMT, $\left(\frac{1}{n}X_{\widehat{\gamma}}^{*\prime}X_{\widehat{\gamma}}^{*}\right)^{-1}\xrightarrow{p}\begin{pmatrix}M & M_0\\ M_0 & M_0\end{pmatrix}^{-1}>0$. In summary, $\frac{1}{\|\delta_n\|}\begin{pmatrix}\widehat{\beta}_2-\beta_{20}\\ \widehat{\delta}-\delta_n\end{pmatrix}=O_p(1)\,o_p(1)=o_p(1)$. ∎

Lemma 4 *Under Assumption I(1)′, $a_n(\widehat{\gamma}-\gamma_0)=O_p(1)$ and $n^{1/2}\left(\widehat{\beta}-\beta_0\right)=O_p(1)$.*

Proof. Since δ_n depends on n, Corollary 3.2.6 of VW cannot be used. Nevertheless, we can apply the proof idea of Theorem 3.2.5 in VW to prove this result. Define $d_n(\theta,\theta_0)=\|\beta-\beta_0\|+\|\delta_n\|\sqrt{\|\gamma-\gamma_0\|}$ for θ in a neighborhood of θ_0, and

$$\begin{aligned}
&Q_n(\theta)=\tfrac{1}{\|\delta_n\|^2}\left(S_n(\theta)-S_n(\theta_0)\right)^{17}\\
&=\tfrac{1}{a_n}\textstyle\sum_{i=1}^{n}T(w_i|\beta_1,\beta_{10})1(q_i\le\gamma_0)+\tfrac{1}{a_n}\sum_{i=1}^{n}T(w_i|\beta_2,\beta_{20})1(q_i>\gamma_0)\\
&\quad+\tfrac{1}{a_n}\textstyle\sum_{i=1}^{n}\left(\overline{z}_1(w_i|\beta_2,\beta_{10})-T(w_i|\beta_1,\beta_{10})\right)1(\gamma\wedge\gamma_0<q_i\le\gamma_0)\\
&\quad+\tfrac{1}{a_n}\textstyle\sum_{i=1}^{n}\left(\overline{z}_2(w_i|\beta_1,\beta_{20})-T(w_i|\beta_2,\beta_{20})\right)1(\gamma_0<q_i\le\gamma\vee\gamma_0)\\
&=:T_1(\theta)+T_2(\theta)+T_3(\theta)+T_4(\theta).
\end{aligned}$$

[17] Note that $Q_n(\theta)\neq\|\delta_n\|^{-2}\left[S_n(\gamma)-S_n(\theta_0)\right]$.

For each n, the parameter space (minus θ_0) can be partitioned into the "shells" $S_{j,n} = \left\{\theta : 2^{j-1} < \sqrt{n} d_n\left(\theta, \theta_0\right) \leq 2^j\right\}$ with j ranging over the integers. Given an integer J,

$$
\begin{aligned}
P\left(d_n\left(\widehat{\theta}, \theta_0\right) > 2^J\right) \leq & \sum_{j \geq J, \|\beta-\beta_0\| < M\|\delta_n\|, \|\gamma-\gamma_0\| < \eta} P\left(\inf_{\theta \in S_{j,n}} Q_n(\theta) \leq 0\right) \\
& + P\left(2\left\|\beta-\beta_0\right\| \geq M\left\|\delta_n\right\|, 2\left\|\gamma-\gamma_0\right\| \geq \eta\right),
\end{aligned} \tag{32}
$$

where M and η are small positive numbers. The second term on the right hand side of (32) converges to zero as $n \to \infty$ for every $\eta > 0$ and $M > 0$ by Lemma 3, so we can concentrate on the first term.

$$
\begin{aligned}
& P\left(\inf_{\theta \in S_{j,n}} Q_n(\theta) \leq 0\right) \leq P\left(\sup_{\theta \in S_{j,n}} \left|Q_n(\theta) - \mathbb{E}\left[Q_n(\theta)\right]\right| \geq \inf_{\theta \in S_{j,n}} \left|\mathbb{E}\left[Q_n(\theta)\right]\right|\right) \\
\leq & \ \mathbb{E}\left[\sup_{\theta \in S_{j,n}} \left|Q_n(\theta) - \mathbb{E}\left[Q_n(\theta)\right]\right|\right] \Big/ \inf_{\theta \in S_{j,n}} \left|\mathbb{E}\left[Q_n(\theta)\right]\right| \\
\leq & \sum_{k=1}^{4} \mathbb{E}\left[\sup_{\theta \in S_{j,n}} \left|T_k\left(\theta\right) - \mathbb{E}\left[T_k\left(\theta\right)\right]\right|\right] \Big/ \inf_{\theta \in S_{j,n}} \left|\mathbb{E}\left[Q_n(\theta)\right]\right|,
\end{aligned}
$$

where the last equality is from Markov's inequality.

From the analysis in Lemma 2, it is not hard to see that

$$
\begin{aligned}
& \inf_{\theta \in S_{j,n}} \left|\mathbb{E}\left[Q_n(\theta)\right]\right| = \inf_{\theta \in S_{j,n}} \left|\sum\nolimits_{k=1}^{4} \mathbb{E}\left[T_k\left(\theta\right)\right]\right| \\
= & \inf_{\theta \in S_{j,n}} C\left|\tfrac{n}{a_n}\left\|\beta-\beta_0\right\|^2 + \tfrac{n}{a_n}\left[\left\|\beta_{10}-\beta_2\right\|^2 + \left\|\delta_n\right\|\left\|\beta_{10}-\beta_2\right\| + \left\|\beta_{20}-\beta_1\right\|^2 + \left\|\delta_n\right\|\left\|\beta_{20}-\beta_1\right\|\right]\left\|\gamma-\gamma_0\right\|\right| \\
= & \inf_{\theta \in S_{j,n}} C\left|\tfrac{n}{a_n}\left\|\beta-\beta_0\right\|^2 + \tfrac{n}{a_n}\left\|\delta_n\right\|^2\left\|\gamma-\gamma_0\right\|\right| = \inf_{\theta \in S_{j,n}} C\tfrac{n}{a_n} d_n\left(\theta, \theta_0\right)^2 \geq C\tfrac{2^{2j-2}}{a_n} = C\tfrac{2^{2j}}{a_n},
\end{aligned}
$$

where the third equality is because $\beta_{10} - \beta_{20} = \delta_n$ and $\left\|\beta_\ell - \beta_{\ell 0}\right\| < M\left\|\delta_n\right\|$ so that $\left\|\beta_1 - \beta_{20}\right\| = O\left(\left\|\delta_n\right\|\right)$ and $\left\|\beta_{20} - \beta_1\right\| = O\left(\left\|\delta_n\right\|\right)$. From Lemma 2, for $k = 1, 2$,

$$
\sum_{k=1}^{2} \mathbb{E}\left[\sup_{\theta \in S_{j,n}} \left|T_k\left(\theta\right) - \mathbb{E}\left[T_k\left(\theta\right)\right]\right|\right] \leq C\frac{\sup_{\theta \in S_{j,n}} \left\|\beta-\beta_0\right\|}{\sqrt{n}\left\|\delta_n\right\|^2}.
$$

As to $T_3\left(\theta\right)$, applying a maximal inequality (e.g., Theorem 2.14.1 of VW) we can show that

$$
\mathbb{E}\left[\sup_{\theta \in S_{j,n}} \left|T_3\left(\theta\right) - \mathbb{E}\left[T_3\left(\theta\right)\right]\right|\right] \leq C\frac{\sup_{\theta \in S_{j,n}} \sqrt{\left\|\beta_{10}-\beta_2\right\|^2}\sqrt{\left|\gamma-\gamma_0\right|}}{\sqrt{n}\left\|\delta_n\right\|^2} = \frac{\sup_{\theta \in S_{j,n}} \left\|\delta_n\right\|\sqrt{\left|\gamma-\gamma_0\right|}}{\sqrt{n}\left\|\delta_n\right\|^2}.
$$

Similarly, $\mathbb{E}\left[\sup_{\theta \in S_{j,n}} \left|T_3\left(\theta\right) - \mathbb{E}\left[T_3\left(\theta\right)\right]\right|\right] \leq C\frac{\sup_{\theta \in S_{j,n}} \left\|\delta_n\right\|\sqrt{\left|\gamma-\gamma_0\right|}}{\sqrt{n}\left\|\delta_n\right\|^2}$. So

$$
\sum_{k=1}^{4} \mathbb{E}\left[\sup_{\theta \in S_{j,n}} \left|T_k\left(\theta\right) - \mathbb{E}\left[T_k\left(\theta\right)\right]\right|\right] \leq C\frac{\sup_{\theta \in S_{j,n}} d_n\left(\theta, \theta_0\right)}{\sqrt{n}\left\|\delta_n\right\|^2} \leq C\frac{2^j/\sqrt{n}}{\sqrt{n}\left\|\delta_n\right\|^2} = C\frac{2^j}{a_n}.
$$

In summary,

$$\sum_{j\geq J,\|\beta-\beta_0\|<M\|\delta_n\|,|\gamma-\gamma_0|<\eta} P\left(\sup_{\theta\in S_{j,n}} Q_n(\theta)\geq 0\right)\leq C\sum_{j\geq J}\left(\frac{2^j}{a_n}\Big/\frac{2^{2j}}{a_n}\right)\leq C\sum_{j\geq J}\frac{1}{2^j},$$

which can be made arbitrarily small by letting J large enough. So $\sqrt{n}d_n\left(\widehat{\theta},\theta_0\right)=O_p(1)$, which implies $a_n(\widehat{\gamma}-\gamma_0)=O_p(1)$, and $\sqrt{n}(\beta-\beta_0)=O_p(1)$. ∎

Lemma 5 *Under Assumption I(α), $1<\alpha\leq 2$, $\rho_n(\widehat{\gamma}-\gamma_0)=O_p(1)$ and $\kappa_n\left(\widehat{\beta}-\beta_0\right)=O_p(1)$, where ρ_n and κ_n are defined in Section 3; especially, when $\alpha=2$, $n^{1/3}\left(\widehat{\theta}-\theta_0\right)=O_p(1)$.*

Proof. When $1<\alpha<2$, define $d(\theta,\theta_0)=\|\beta-\beta_0\|+\sqrt{\Lambda(|\gamma-\gamma_0|)}$ for θ in a neighborhood of θ_0. Given that $\Lambda(\cdot)$ is monotone, we can check $d(\theta,\theta_0)$ is indeed a peudo-metric.[18] Since $\Psi_\pm(\beta,\gamma)$ in $S(\theta)-S(\theta_0)$ are dominated,

$$\begin{aligned} S(\theta)-S(\theta_0) &\approx \Phi(\beta_1)+\overline{\Phi}(\beta_2)+\Lambda_+(\gamma)1(\gamma>\gamma_0)+\Lambda_-(\gamma)1(\gamma\leq\gamma_0)\\ &\geq C\left(\|\beta_1-\beta_{10}\|^2+\|\beta_2-\beta_{20}\|^2+\Lambda(|\gamma-\gamma_0|)\right). \end{aligned}$$

Second, we use the same envelope function F as in the proof of Lemma 2 except that η^2 in the index functions for q are replaced by $\Lambda^{\leftarrow}(\eta^2)$. Now, $\sqrt{PF^2}=\sqrt{C(\eta^2+\Lambda^{\leftarrow}(\eta^2))}\leq C\sqrt{\Lambda^{\leftarrow}(\eta^2)}$ for $\eta<1$, i.e., the last two terms of F dominate. So $\phi(\eta)=\sqrt{\Lambda^{\leftarrow}(\eta^2)}\in RV_{1/\alpha}$ and η/η^{ϱ} is decreasing for all $1/\alpha<\varrho<2$. Since $r_n^2\phi\left(\frac{1}{r_n}\right)=\sqrt{n}$ implies $r_n=\kappa_n$ which is the convergence rate of $\widehat{\beta}$. By the definition of d, the convergence rate of $\widehat{\gamma}$, ρ_n, can be obtained by solving $\sqrt{\Lambda(1/\rho_n)}=1/r_n$.

When $\alpha=2$, $\Lambda(|\gamma-\gamma_0|)=|\gamma-\gamma_0|^2$, so $d(\theta,\theta_0)=\|\theta-\theta_0\|$ is the Euclidean norm. Now, $\Psi_\pm(\beta,\gamma)$ in $S(\theta)-S(\theta_0)$ are not dominated. Anyway, because $S^{\pm}_{\theta\theta}>0$ and $S(\theta)-S(\theta_0)\approx\frac{1}{2}(\theta-\theta_0)'S^{+}_{\theta\theta}(\theta-\theta_0)1(\gamma>\gamma_0)+\frac{1}{2}(\theta-\theta_0)'S^{-}_{\theta\theta}(\theta-\theta_0)1(\gamma\leq\gamma_0)$, we have $S(\theta)-S(\theta_0)\geq Cd^2(\theta,\theta_0)$. Also, $\phi(\eta)=\sqrt{\eta}$, which implies $r_n=n^{1/3}$, i.e., $n^{1/3}d\left(\widehat{\theta}-\theta_0\right)=O_P(1)$. ∎

Lemma 6 *Under Assumption II(α), $2\leq\alpha\leq 4$, $\rho_n(\widehat{\gamma}-\gamma_0)=O_p(1)$ and $\kappa_n\left(\widehat{\beta}-\beta_0\right)=O_p(1)$, where ρ_n and κ_n are defined in Section 3; especially, when $\alpha=2$, $n^{1/2}\left(\widehat{\theta}-\theta_0\right)=O_p(1)$, when $\alpha=3$, $n^{1/3}(\widehat{\gamma}-\gamma_0)=O_p(1)$ and $n^{1/2}\left(\widehat{\beta}-\beta_0\right)=O_p(1)$, and when $\alpha=4$, $n^{1/5}(\widehat{\gamma}-\gamma_0)=O_p(1)$ and $n^{2/5}\left(\widehat{\beta}-\beta_0\right)=O_p(1)$.*

Proof. Define $d(\theta,\theta_0)=\|\beta-\beta_0\|+\sqrt{\Lambda(|\gamma-\gamma_0|)}$ for θ in a neighborhood of θ_0, which reduces to the Euclidean norm when $\alpha=2$, $\|\beta-\beta_0\|+|\gamma-\gamma_0|^{3/2}$ when $\alpha=3$ and $\|\beta-\beta_0\|+|\gamma-\gamma_0|^2$ when $\alpha=4$. When $\alpha=2$, $\Psi_\pm(\beta,\gamma)$ in $S(\theta)-S(\theta_0)$ are not dominated, but because $S^{\pm}_{\theta\theta}>0$, we have $S(\theta)-S(\theta_0)\geq Cd^2(\theta,\theta_0)$. When $2<\alpha<4$, $\Psi_\pm(\beta,\gamma)$ are dominated, and $S(\theta)-S(\theta_0)\geq C\left(\|\beta-\beta_0\|^2+\Lambda(|\gamma-\gamma_0|)\right)$. When $\alpha=4$, $\Psi_\pm(\beta,\gamma)$ in $S(\theta)-S(\theta_0)$ are not dominated, but we can show

$$\begin{aligned} \Psi_-(\beta,\gamma) &\approx \begin{pmatrix} -\frac{1}{2}\mathbb{E}[\mathbf{x}'q\delta_{q0}1(\gamma<q\leq\gamma_0)]\widetilde{\beta}_1\\ -\frac{1}{2}\mathbb{E}[\mathbf{x}'q\delta_{q0}1(\gamma<q\leq\gamma_0)]'\widetilde{\beta}_2 \end{pmatrix}\approx\begin{pmatrix} \frac{\delta_{q0}f_0}{4}\mathbb{E}[\mathbf{x}'|q=\gamma_0]\widetilde{\beta}_1\gamma^2\\ \frac{\delta_{q0}f_0}{4}\mathbb{E}[\mathbf{x}'|q=\gamma_0]\widetilde{\beta}_2\gamma^2 \end{pmatrix}\\ \text{and }\Psi_+(\beta,\gamma) &\approx \begin{pmatrix} \frac{1}{2}\mathbb{E}[\mathbf{x}'q\delta_{q0}1(\gamma_0<q\leq\gamma)]\widetilde{\beta}_1\\ \frac{1}{2}\mathbb{E}[\mathbf{x}'q\delta_{q0}1(\gamma_0<q\leq\gamma)]\widetilde{\beta}_2 \end{pmatrix}\approx\begin{pmatrix} \frac{\delta_{q0}f_0}{4}\mathbb{E}[\mathbf{x}'|q=\gamma_0]\widetilde{\beta}_1\gamma^2\\ \frac{\delta_{q0}f_0}{4}\mathbb{E}[\mathbf{x}'|q=\gamma_0]\widetilde{\beta}_2\gamma^2 \end{pmatrix} \end{aligned}$$

[18] As noted in the footnote on page 289 of VW, $d(\cdot,\cdot)$ need not be a peudo-metric, but because we need the reverse function of $\Lambda(\cdot)$ below, we assume $\Lambda(\cdot)$ to be monotone.

under Assumptions (vii) and (x)(b). As a result, by Taylor expansion,

$$\begin{aligned}
S(\theta)-S(\theta_0) &\approx \tfrac{1}{2}\widetilde{\beta}_1' S_{\beta_1\beta_1}\widetilde{\beta}_1+\tfrac{1}{2}\widetilde{\beta}_2' S_{\beta_2\beta_2}\widetilde{\beta}_2+\left[\tfrac{\delta_{q0}f_0}{4}\mathbb{E}\left[\mathbf{x}'|q=\gamma_0\right]\widetilde{\beta}_1\gamma^2+\tfrac{\delta_{q0}f_0}{4}\mathbb{E}\left[\mathbf{x}'|q=\gamma_0\right]\widetilde{\beta}_2\gamma^2+\lambda_-\left|\gamma\right|^4\right]1(\gamma\le\gamma_0)\\
&\quad+\left[\tfrac{\delta_{q0}f_0}{4}\mathbb{E}\left[\mathbf{x}'|q=\gamma_0\right]\widetilde{\beta}_1\gamma^2+\tfrac{\delta_{q0}f_0}{4}\mathbb{E}\left[\mathbf{x}'|q=\gamma_0\right]\widetilde{\beta}_2\gamma^2+\lambda_+\gamma^4\right]1(\gamma>\gamma_0)\\
&=\tfrac{1}{2}\left(\widetilde{\beta}_1',\widetilde{\beta}_2',\gamma^2\right)\mathbb{S}_-\left(\widetilde{\beta}_1',\widetilde{\beta}_2',\gamma^2\right)'1(\gamma\le\gamma_0)+\tfrac{1}{2}\left(\widetilde{\beta}_1',\widetilde{\beta}_2',\gamma^2\right)\mathbb{S}_+\left(\widetilde{\beta}_1',\widetilde{\beta}_2',\gamma^2\right)'1(\gamma>\gamma_0)\\
&\ge C\left(\left\|\beta_1-\beta_{10}\right\|^2+\left\|\beta_2-\beta_{20}\right\|^2+\left|\gamma-\gamma_0\right|^4\right),
\end{aligned}$$

where the last equality is from Assumption (x)(d).

To bound the modulus of continuity of the empirical process, we use a different envelope function in CTR. Specifically,

$$\begin{aligned}
F&=\left(\left|m_1(x,q)-\mathbf{x}'\beta_{10}\right|\left\|\mathbf{x}\right\|\eta+\tfrac{1}{2}\eta^2\left\|\mathbf{x}\right\|^2+\eta\left\|\mathbf{x}\varepsilon_1\right\|\right)1\left(q\le\gamma_0+\Lambda^{\leftarrow}\left(\eta^2\right)\right)\\
&\quad+\left(\left|m_2(x,q)-\mathbf{x}'\beta_{20}\right|\left\|\mathbf{x}\right\|\eta+\tfrac{1}{2}\eta^2\left\|\mathbf{x}\right\|^2+\eta\left\|\mathbf{x}\varepsilon_2\right\|\right)1\left(q>\gamma_0-\Lambda^{\leftarrow}\left(\eta^2\right)\right)\\
&\quad+F_3\left(\mathbf{x},\varepsilon_1|\eta\right)+F_4\left(\mathbf{x},\varepsilon_2|\eta\right),
\end{aligned}$$

where $F_3\left(\mathbf{x},\varepsilon_1|\eta\right)$ is based on

$$\begin{aligned}
&\sup_{d(\theta,\theta_0)<\eta}\left|\overline{z}_1\left(w|\beta_2,\beta_{10}\right)\right|1\left(\gamma<q\le\gamma_0\right)\\
=&\sup_{d(\theta,\theta_0)<\eta}\left|\left(m_1(x,q)-\mathbf{x}'\overline{\beta}_0-\tfrac{\mathbf{x}'\widetilde{\beta}_2}{2}\right)\mathbf{x}'\left(\delta_0-\widetilde{\beta}_2\right)+\left(\delta_0-\widetilde{\beta}_2\right)'\mathbf{x}\varepsilon_1\right|1\left(\gamma<q\le\gamma_0\right)\\
\le&\sup_{d(\theta,\theta_0)<\eta}\left|\left(m_1(x,q)-\mathbf{x}'\overline{\beta}_0\right)\mathbf{x}1\left(\gamma<q\le\gamma_0\right)\right|\left\|\widetilde{\beta}_2\right\|+\left|\left(m_1(x,q)-\mathbf{x}'\overline{\beta}_0\right)q1\left(\gamma<q\le\gamma_0\right)\right|\left|\delta_{q0}\right|\\
+&\left(\tfrac{\left\|\widetilde{\beta}_2\right\|\left|\delta_{q0}\right|}{2}\left\|\mathbf{x}\right\|\left|q\right|+\left\|\mathbf{x}\right\|^2\tfrac{\left\|\widetilde{\beta}_2\right\|^2}{2}\right)1\left(\gamma<q\le\gamma_0\right)+\left(\delta_{q0}\left|q\right|\left|\varepsilon_1\right|+\left\|\widetilde{\beta}_2\right\|\left\|\mathbf{x}\varepsilon_1\right\|\right)1\left(\gamma<q\le\gamma_0\right)\\
\le&\, C\eta\left|\left(m_1(x,q)-\mathbf{x}'\overline{\beta}_0\right)\right|\left\|\mathbf{x}\right\|1\left(\gamma_0-\Lambda^{\leftarrow}\left(\eta^2\right)<q\le\gamma_0\right)\\
+&C\Lambda^{\leftarrow}\left(\eta^2\right)\left|\left(m_1(x,q)-\mathbf{x}'\overline{\beta}_0\right)\right|1\left(\gamma_0-\Lambda^{\leftarrow}\left(\eta^2\right)<q\le\gamma_0\right)\\
+&C\left(\eta\Lambda^{\leftarrow}\left(\eta^2\right)\left\|\mathbf{x}\right\|+\eta^2\left\|\mathbf{x}\right\|^2\right)1\left(\gamma_0-\Lambda^{\leftarrow}\left(\eta^2\right)<q\le\gamma_0\right)\\
+&C\left(\Lambda^{\leftarrow}\left(\eta^2\right)\left|\varepsilon_1\right|+\eta\left\|\mathbf{x}\varepsilon_1\right\|\right)1\left(\gamma_0-\Lambda^{\leftarrow}\left(\eta^2\right)<q\le\gamma_0\right),
\end{aligned}$$

and similarly for $F_4\left(\mathbf{x},\varepsilon_2|\eta\right)$. Now,

$$\begin{aligned}
PF^2&=O\left(\eta^2+\eta^2\Lambda^{\leftarrow}\left(\eta^2\right)+\Lambda^{\leftarrow}\left(\eta^2\right)^3+\left(\eta^2\Lambda^{\leftarrow}\left(\eta^2\right)^3+\eta^4\Lambda^{\leftarrow}\left(\eta^2\right)\right)+\left(\Lambda^{\leftarrow}\left(\eta^2\right)^3+\eta^2\Lambda^{\leftarrow}\left(\eta^2\right)\right)\right)\\
&=O\left(\eta^2+\Lambda^{\leftarrow}\left(\eta^2\right)^3\right).
\end{aligned}$$

When $\alpha=2$, $\Lambda^{\leftarrow}\left(\eta^2\right)=\eta^2$, so $PF^2=O\left(\eta^2\right)$. When $\alpha=3$, $\Lambda^{\leftarrow}\left(\eta^2\right)=\eta^{2/3}$, $PF^2=O\left(\eta^2\right)$. When $\alpha=4$, $\Lambda^{\leftarrow}\left(\eta^2\right)=\eta^{1/2}$, so $PF^2=O\left(\eta^{3/2}\right)$. For the in-between case, when $2<\alpha<3$, $PF^2=O\left(\eta^2\right)$ and when $3<\alpha<4$, $PF^2=O\left(\Lambda^{\leftarrow}\left(\eta^2\right)^3\right)$. It follows that $\phi(\eta)=\eta$ when $2\le\alpha\le3$ and $\phi(\eta)=\Lambda^{\leftarrow}\left(\eta^2\right)^{3/2}$ when $3<\alpha\le4$, so we can always find a $\varrho<2$ such that η/η^{ϱ} is decreasing. Solving $r_n^2\phi\left(\frac{1}{r_n}\right)=\sqrt{n}$, we have $r_n=\kappa_n$ which is the convergence rate of $\widehat{\beta}$. By the definition of d, the convergence rate of $\widehat{\gamma}$ is ρ_n. Especially, when $\alpha=2,3,4$, $r_n=n^{1/2},n^{1/2},n^{2/5}$, which implies $\rho_n=\Lambda^{-1}\left(1/r_n^2\right)=n^{1/2},n^{1/3},n^{1/5}$, respectively. ■

Lemma 7 *Under Assumption II(α), $n^{1/2}\left(\widehat{\gamma}-\gamma_0\right)=O_p(1)$ when $2<\alpha\le2.5$ and $\varrho_n\left(\widehat{\gamma}-\gamma_0\right)=O_p(1)$ when $2.5<\alpha<3$, where ϱ_n is defined in Theorem 9.*

Proof. Similar as in the proof of Theorems 4 and 8, because the randomness in the β direction dominates, we cannot search over β and γ jointly; rather, we fix $\widehat{\beta}$ and concentrate on the randomness in the γ direction. Specifically,

$$
\begin{aligned}
&\widehat{\gamma}-\gamma_0 \approx \arg\min_{v\in\mathcal{N}_v}\left\{S_n\left(\gamma_0+v,\widehat{\beta}\right)-S_n\left(\gamma_0,\beta_0\right)\right\}\\
&=\arg\min_{v\in\mathcal{N}_v}\left\{\left[S\left(\gamma_0+v,\widehat{\beta}\right)-S\left(\gamma_0,\beta_0\right)\right]+\left[n^{-1/2}\mathbb{G}_n\left(s\left(\cdot\left|\gamma_0+v,\widehat{\beta}\right.\right)\right)-n^{-1/2}\mathbb{G}_n\left(s\left(\cdot\left|\gamma_0,\beta_0\right.\right)\right)\right]\right\}\\
&=\arg\min_{v\in\mathcal{N}_v}\left\{\tfrac{1}{2}\left(\widehat{\beta}-\beta_0\right)'S_{\beta\beta}\left(\widehat{\beta}-\beta_0\right)+\tfrac{1}{2}\left(\widehat{\beta}-\beta_0\right)'S_{\beta^2\gamma}\left(\widehat{\beta}-\beta_0\right)v+\tfrac{1}{2}\left(\widehat{\beta}-\beta_0\right)'S_{\beta\gamma^2}v^2\right.\\
&\quad\left[\left(\widehat{\beta}-\beta_0\right)'S_\beta^-\left(v\right)+\lambda_-\left|v\right|^\alpha L_-\left(\left|v\right|\right)\right]1(v\le 0)+\left[\left(\widehat{\beta}-\beta_0\right)'S_\beta^+\left(v\right)+\lambda_+v^\alpha L_+\left(v\right)\right]1(v>0)\\
&\quad\left.+n^{-1/2}\mathbb{G}_n\left(s\left(\cdot\left|\gamma_0+v,\widehat{\beta}\right.\right)-s\left(\cdot\left|\gamma_0,\beta_0\right.\right)\right)+o\left(\left|v\right|^\alpha L_-\left(\left|v\right|\right)\right)+o_p\left(n^{-1}\left|v\right|\right)+o_p\left(n^{-1/2}v^2\right)\right\}\\
&=\arg\min_{v\in\mathcal{N}_v}\left\{\left[\left[\left(\widehat{\beta}-\beta_0\right)'S_\beta^-\left(v\right)+\lambda_-\left|v\right|^\alpha L_-\left(\left|v\right|\right)\right]1(v\le 0)+\left[\left(\widehat{\beta}-\beta_0\right)'S_\beta^+\left(v\right)+\lambda_+v^\alpha L_+\left(v\right)\right]1(v>0)\right.\right.\\
&\quad n^{-1/2}O_p\left(v^{3/2}\right)+O_p\left(n^{-1/2}\left\|\widehat{\beta}-\beta_0\right\|\left|v\right|^{1/2}\right)\\
&\quad\left.+\left[O_p\left(n^{-3/2}\left|v\right|^{1/2}\right)+O_p\left(n^{-1}v^{3/2}\right)+o_p\left(n^{-1/2}v^{(\alpha-1)}\right)+o\left(\left|v\right|^\alpha L_-\left(\left|v\right|\right)\right)\right]\right\},
\end{aligned}
$$

where $\mathcal{N}_v$ is a ρ_n^{-1}-neighborhood of 0, $S_\beta^\pm\left(v\right)\in RV_{\alpha-1}$, and

$$
\begin{aligned}
&n^{-1/2}\mathbb{G}_n\left(s\left(\cdot\left|\gamma_0+v,\widehat{\beta}\right.\right)-s\left(\cdot\left|\gamma_0,\beta_0\right.\right)\right)\\
&=C\left(\widehat{\beta}-\beta_0\right)+n^{-1/2}\mathbb{G}_n\left[\left(y-\mathbf{x}'\overline{\beta}_0-\mathbf{x}'\tfrac{\widehat{\beta}_2-\beta_{20}}{2}\right)\left(q\delta_{q0}-\mathbf{x}'\left(\widehat{\beta}_2-\beta_{20}\right)\right)1(v<q\le 0)\right]\\
&\quad-n^{-1/2}\mathbb{G}_n\left(-\left(y-\mathbf{x}'\overline{\beta}_0-\tfrac{\mathbf{x}'\delta_0}{2}-\mathbf{x}'\tfrac{\widehat{\beta}_1-\beta_{10}}{2}\right)\mathbf{x}'\left(\widehat{\beta}_1-\beta_{10}\right)1(v<q\le 0)\right)\\
&\quad-n^{-1/2}\mathbb{G}_n\left(\left(y-\mathbf{x}'\overline{\beta}_0-\mathbf{x}'\tfrac{\widehat{\beta}_1-\beta_{10}}{2}\right)\left(q\delta_{q0}+\mathbf{x}'\left(\widehat{\beta}_1-\beta_{10}\right)\right)1(0<q\le v)\right)\\
&\quad-n^{-1/2}\mathbb{G}_n\left(-\left(y-\mathbf{x}'\overline{\beta}_0+\tfrac{\mathbf{x}'\delta_0}{2}-\mathbf{x}'\tfrac{\widehat{\beta}_2-\beta_{20}}{2}\right)\mathbf{x}'\left(\widehat{\beta}_2-\beta_{20}\right)1(0<q\le v)\right)\\
&=C\left(\widehat{\beta}-\beta_0\right)+n^{-1/2}\mathbb{G}_n\left[\left(y-\mathbf{x}'\overline{\beta}_0\right)q\delta_{q0}1(v<q\le 0)-\left(y-\mathbf{x}'\overline{\beta}_0\right)q\delta_{q0}1(0<q\le v)\right]\\
&\quad-n^{-1/2}\left(\widehat{\beta}_2-\beta_{20}\right)'\mathbb{G}_n\left[\left(y-\mathbf{x}'\overline{\beta}_0\right)\mathbf{x}1(v<q\le 0)\right]\\
&\quad+n^{-1/2}\left(\widehat{\beta}_1-\beta_{10}\right)\mathbb{G}_n\left[\left(y-\mathbf{x}'\overline{\beta}_0\right)\mathbf{x}1(v<q\le 0)\right]\\
&\quad-n^{-1/2}\left(\widehat{\beta}_1-\beta_{10}\right)'\mathbb{G}_n\left[\left(y-\mathbf{x}'\overline{\beta}_0\right)\mathbf{x}1(0<q\le v)\right]\\
&\quad+n^{-1/2}\left(\widehat{\beta}_2-\beta_{20}\right)'\mathbb{G}_n\left[\left(y-\mathbf{x}'\overline{\beta}_0\right)\mathbf{x}1(0<q\le v)\right]+O_p\left(n^{-3/2}\left|v\right|^{1/2}\right)+O_p\left(n^{-1}\left|v\right|^{3/2}\right)\\
&=C\left(\widehat{\beta}-\beta_0\right)+n^{-1/2}O_p\left(v^{3/2}\right)+O_p\left(n^{-1/2}\left\|\widehat{\beta}-\beta_0\right\|\left|v\right|^{1/2}\right)+O_p\left(n^{-3/2}\left|v\right|^{1/2}\right)+O_p\left(n^{-1}v^{3/2}\right)
\end{aligned}
$$

with

$$
\begin{aligned}
&C\left(\widehat{\beta}-\beta_0\right)=n^{-1/2}\left(\widehat{\beta}_1-\beta_{10}\right)'\mathbb{G}_n\left(-\left(y-\mathbf{x}'\beta_{10}-\mathbf{x}'\tfrac{\widehat{\beta}_1-\beta_{10}}{2}\right)\mathbf{x}1(q\le\gamma_0)\right)\\
&\quad+n^{-1/2}\left(\widehat{\beta}_2-\beta_{20}\right)'\mathbb{G}_n\left(-\left(y-\mathbf{x}'\beta_{20}-\mathbf{x}'\tfrac{\widehat{\beta}_2-\beta_{20}}{2}\right)\mathbf{x}1(q>\gamma_0)\right)\\
&=n^{-1/2}\left(\widehat{\beta}_1-\beta_{10}\right)'\mathbb{G}_n\left(-\mathbf{x}e_11(q\le\gamma_0)\right)+n^{-1/2}\left(\widehat{\beta}_2-\beta_{20}\right)'\mathbb{G}_n\left(-\mathbf{x}e_21(q>\gamma_0)\right)+O_p\left(n^{-3/2}\right)\\
&=-\left(\widehat{\beta}-\beta_0\right)'\operatorname{diag}\left(\tfrac{1}{n}X'_{\le\gamma_0}X_{\le\gamma_0},\tfrac{1}{n}X'_{>\gamma_0}X_{>\gamma_0}\right)\left(\widehat{\beta}\left(\gamma_0\right)-\beta_0\right)+O_p\left(n^{-3/2}\right).
\end{aligned}
$$

Now, we balance the "deterministic" part $\max\left\{\left\|\widehat{\beta}-\beta_0\right\|\left\|S_\beta\left(v\right)\right\|,\left|v\right|^\alpha L\left(\left|v\right|\right)\right\}$ and the "random" part $\max\left\{n^{-1/2}O_p\left(v^{3/2}\right),O_p\left(n^{-1/2}\left\|\widehat{\beta}-\beta_0\right\|\left|v\right|^{1/2}\right)\right\}$. We consider three areas of $\mathcal{N}_v$. (i) $|v|\prec n^{-1/2}$. But we then need to balance $\left|v\right|^\alpha L\left(\left|v\right|\right)$ and $O_p\left(n^{-1/2}\left\|\widehat{\beta}-\beta_0\right\|\left|v\right|^{1/2}\right)$ to have $|v|\succ n^{-1/2}$, impossible! (ii) $|v|\succ n^{-1/2}$. Then we need to balance $\left|v\right|^\alpha L\left(\left|v\right|\right)$ and $n^{-1/2}O_p\left(v^{3/2}\right)$ to have $v\sim\varrho_n$, where $\varrho_n\succ n^{-1/2}$ only if $\alpha>$

2.5. (iii) $|v| \sim n^{-1/2}$. If $\alpha > 2.5$, the random part dominates, and the minimizer is $O\left(\rho_n^{-1}\right) \succ n^{-1/2}$, impossible! If $\alpha < 2.5$, the deterministic part dominates, and we minimize $\left[\left(\widehat{\beta}-\beta_0\right)' S_\beta^-(v) + \lambda_- |v|^\alpha L_-(|v|)\right] 1(v \leq 0) + \left[\left(\widehat{\beta}-\beta_0\right)' S_\beta^+(v) + \lambda_+ v^\alpha L_+(v)\right] 1(v > 0)$ to have $v \sim n^{-1/2}$. If $\alpha = 2.5$, we need to balance all terms to have $v \sim n^{-1/2}$. The arguments here can be made rigorous using the slicing method as in the proof of Lemma 4, and the details are omitted.

In summary, when $2.5 < \alpha < 3$,

$$\begin{aligned}\varrho_n(\widehat{\gamma}-\gamma_0) = \arg\min\nolimits_v \sqrt{n\varrho_n^3}\Big\{&\lambda_- \left|\tfrac{v}{\varrho_n}\right|^\alpha L_-\left(\left|\tfrac{v}{\varrho_n}\right|\right) 1(v \leq 0) + \lambda_+\left(\tfrac{v}{\varrho_n}\right)^\alpha L_+\left(\tfrac{v}{\varrho_n}\right) 1(v>0)\\ &n^{-1/2}\mathbb{G}_n\left[\left(y-\mathbf{x}'\overline{\beta}_0\right) q\delta_{q0} 1(\tfrac{v}{\varrho_n} < q \leq 0) - \left(y-\mathbf{x}'\overline{\beta}_0\right) q\delta_{q0} 1(0 < q \leq \tfrac{v}{\varrho_n})\right] + o_p(1)\Big\},\end{aligned}$$

which is the same as the case where β_0 is known; when $2 < \alpha < 2.5$,

$$\begin{aligned}\sqrt{n}(\widehat{\gamma}-\gamma_0) = \arg\min\nolimits_v \frac{n^{\alpha/2}}{L\left(\frac{1}{n^{1/2}}\right)}\Big\{&\left[\left(\widehat{\beta}-\beta_0\right)' S_\beta^-\left(\tfrac{v}{n^{1/2}}\right) + \lambda_-\left|\tfrac{v}{n^{1/2}}\right|^\alpha L_-\left(\left|\tfrac{v}{n^{1/2}}\right|\right)\right] 1(v \leq 0)\\ &+\left[\left(\widehat{\beta}-\beta_0\right)' S_\beta^+\left(\tfrac{v}{n^{1/2}}\right) + \lambda_+\left(\tfrac{v}{n^{1/2}}\right)^\alpha L_+\left(\tfrac{v}{n^{1/2}}\right)\right] 1(v>0) + o_p(1)\Big\},\end{aligned}$$

where the randomness of $\widehat{\gamma}$ is driven only by $\widehat{\beta}$; when $\alpha = 2.5$,

$$\sqrt{n}(\widehat{\gamma}-\gamma_0) = \arg\min_v \left\{\mathbb{M}_n(v) + o_p(1)\right\},$$

where

$$\begin{aligned}\mathbb{M}_n(v) = n^{5/4}\Big\{&\left[\left(\widehat{\beta}-\beta_0\right)' S_\beta^-\left(\tfrac{v}{n^{1/2}}\right) + \lambda_-\left|\tfrac{v}{n^{1/2}}\right|^{5/2} L_-\left(\left|\tfrac{v}{n^{1/2}}\right|\right)\right] 1(v \leq 0)\\ &+\left[\left(\widehat{\beta}-\beta_0\right)' S_\beta^+\left(\tfrac{v}{n^{1/2}}\right) + \lambda_+\left(\tfrac{v}{n^{1/2}}\right)^{5/2} L_+\left(\tfrac{v}{n^{1/2}}\right)\right] 1(v>0)\\ &+n^{-1/2}\mathbb{G}_n\left[\left(y-\mathbf{x}'\overline{\beta}_0\right) q\delta_{q0} 1(\tfrac{v}{n^{1/2}} < q \leq 0) - \left(y-\mathbf{x}'\overline{\beta}_0\right) q\delta_{q0} 1(0 < q \leq \tfrac{v}{n^{1/2}})\right]\\ &-n^{-1/2}\left(\widehat{\beta}_2-\beta_{20}\right)'\mathbb{G}_n\left[\left(y-\mathbf{x}'\overline{\beta}_0\right)\mathbf{x}1(\tfrac{v}{n^{1/2}} < q \leq 0)\right]\\ &+n^{-1/2}\left(\widehat{\beta}_1-\beta_{10}\right)\mathbb{G}_n\left[\left(y-\mathbf{x}'\overline{\beta}_0\right)\mathbf{x}1(\tfrac{v}{n^{1/2}} < q \leq 0)\right]\\ &-n^{-1/2}\left(\widehat{\beta}_1-\beta_{10}\right)'\mathbb{G}_n\left[\left(y-\mathbf{x}'\overline{\beta}_0\right)\mathbf{x}1(0 < q \leq \tfrac{v}{n^{1/2}})\right]\\ &+n^{-1/2}\left(\widehat{\beta}_2-\beta_{20}\right)'\mathbb{G}_n\left[\left(y-\mathbf{x}'\overline{\beta}_0\right)\mathbf{x}1(0 < q \leq \tfrac{v}{n^{1/2}})\right]\Big\}.\end{aligned}$$

■

Appendix C: Local Approximation

This appendix intends to show the main terms in the localized objective function. We will detail the proof for I(1) and then adjust it for other cases.

Lemma 8 *Under Assumption I(1), uniformly for h in a compact set,*

$$\begin{aligned}&nP_n\left(s\left(\cdot\left|\gamma_0+\frac{v}{n},\beta_0+\frac{u}{n^{1/2}}\right.\right) - s\left(\cdot\left|\gamma_0,\beta_0\right.\right)\right)\\ =\;& \frac{1}{2}u_1'M_0u_1 + \frac{1}{2}u_2'\overline{M}_0u_2 - W_n(u) + D_n(v) + o_P(1),\end{aligned}$$

where $W_n(u) = W_{1n}(u_1) + W_{2n}(u_2)$ *with*

$$W_{1n}(u_1) = \frac{u_1'}{\sqrt{n}} \sum_{i=1}^{n} \mathbf{x}_i e_{1i} 1(q_i \le \gamma_0) \text{ and } W_{2n}(u_2) = \frac{u_2'}{\sqrt{n}} \sum_{i=1}^{n} \mathbf{x}_i e_{2i} 1(q_i > \gamma_0),$$

and

$$\begin{aligned} D_n(v) &= nP_n\left(s\left(\cdot\middle|\gamma_0 + \frac{v}{n}, \beta_0\right) - s(\cdot|\gamma_0, \beta_0)\right) \\ &= \sum_{i=1}^{n} \bar{z}_{1i} 1\left(\gamma_0 + \frac{v}{n} < q_i \le \gamma_0\right) + \sum_{i=1}^{n} \bar{z}_{2i} 1\left(\gamma_0 < q_i \le \gamma_0 + \frac{v}{n}\right). \end{aligned}$$

Proof. From the decomposition of $s(w|\theta) - s(w|\theta_0)$,

$$\begin{aligned} & nP_n\left(s\left(\cdot\middle|\gamma_0 + \tfrac{v}{n}, \beta_0 + \tfrac{u}{\sqrt{n}}\right) - s(\cdot|\gamma_0, \beta_0)\right) \\ =& \sum_{i=1}^{n} T\left(w_i|\beta_{10} + \tfrac{u_1}{\sqrt{n}}, \beta_{10}\right) 1(q_i \le \gamma_0) + \sum_{i=1}^{n} T\left(w|\beta_{20} + \tfrac{u_2}{\sqrt{n}}, \beta_{20}\right) 1(q > \gamma_0) \\ &+ \sum_{i=1}^{n} \bar{z}_1\left(w_i|\beta_{20} + \tfrac{u_2}{\sqrt{n}}, \beta_{10}\right) 1(\gamma_0 + \tfrac{v}{n} \wedge \gamma_0 < q_i \le \gamma_0) \\ &- \sum_{i=1}^{n} T\left(w_i|\beta_{10} + \tfrac{u_1}{\sqrt{n}}, \beta_{10}\right) 1(\gamma_0 + \tfrac{v}{n} \wedge \gamma_0 < q_i \le \gamma_0) \\ &+ \sum_{i=1}^{n} \bar{z}_2\left(w_i|\beta_{10} + \tfrac{u_1}{\sqrt{n}}, \beta_{20}\right) 1(\gamma_0 < q_i \le \gamma_0 + \tfrac{v}{n} \vee \gamma_0) \\ &- \sum_{i=1}^{n} T\left(w_i|\beta_{20} + \tfrac{u_2}{\sqrt{n}}, \beta_{20}\right) 1(\gamma_0 < q_i \le \gamma_0 + \tfrac{v}{n} \vee \gamma_0) \\ =:& T_1(u_1) + T_2(u_2) + T_3(u_1, v) - T_4(u_2, v) + T_5(u_2, v) - T_6(u_1, v). \end{aligned}$$

Check each term in turn. The analyses for $T_2(u_2), T_5(u_2, v)$ and $T_6(u_1, v)$ are similar to $T_1(u_1), T_3(u_1, v)$ and $T_4(u_2, v)$, so we concentrate the latter three terms below. First,

$$\begin{aligned} T_1(u_1) &= \sum_{i=1}^{n} -\left(e_{1i} - \tfrac{\mathbf{x}_i'}{2}\tfrac{u_1}{\sqrt{n}}\right)\mathbf{x}_i'\tfrac{u_1}{\sqrt{n}} 1(q_i \le \gamma_0) \\ &= \tfrac{u_1' S_{\beta_1\beta_1} u_1}{2} - \tfrac{u_1'}{\sqrt{n}} \sum_{i=1}^{n} \mathbf{x}_i e_{1i} 1(q_i \le \gamma_0) + o_p(1), \end{aligned}$$

where $o_p(1)$ is from the LLN. By a similar analysis, when $v < 0$,

$$\begin{aligned} T_4(u_1, v) &= \tfrac{u_1'\mathbb{E}\left[\mathbf{x}_i\mathbf{x}_i' 1(\gamma_0 + \frac{v}{n} < q_i \le \gamma_0)\right]u_1}{2} - \tfrac{u_1'}{\sqrt{n}} \sum_{i=1}^{n} \mathbf{x}_i e_{1i} 1(\gamma_0 + \tfrac{v}{n} < q_i \le \gamma_0) + o_p(1) \\ &= o_p(1) - u_1'\left[\mathbb{G}_n\left(\mathbf{x}e_1 1(\gamma_0 + \tfrac{v}{n} < q \le \gamma_0)\right) + \sqrt{n}\mathbb{E}\left[\mathbf{x}_i e_{1i} 1(\gamma_0 + \tfrac{v}{n} < q_i \le \gamma_0)\right]\right] + o_p(1) \\ &= o_p(1), \end{aligned}$$

where the $o_p(1)$ in the first equality is from a GC theorem, and the first $o_p(1)$ in the second equality is because

$$\mathbb{E}\left[\mathbf{x}_i\mathbf{x}_i' 1(\gamma_0 + \frac{v}{n} < q_i \le \gamma_0)\right] = \frac{|v|}{n} f(\bar{\gamma})\, \mathbb{E}\left[\mathbf{x}_i\mathbf{x}_i'|q = \bar{\gamma}\right],$$

for some $\bar{\gamma}$ between $\gamma_0 + \frac{v}{n}$ and γ_0, which is which is $o(1)$ by Assumptions (iv)(b) and (vii), and $o_p(1)$ in the third equality is because by the stochastic equicontinuity of $\gamma \mapsto \mathbb{G}_n(\mathbf{x}e_1 1(\gamma < q \le \gamma_0))$,

$$\mathbb{G}_n\left(\mathbf{x}e_1 1(\gamma_0 + \tfrac{v}{n} < q \le \gamma_0)\right) = \mathbb{G}_n(\mathbf{x}e1(\gamma_0 < q \le \gamma_0)) + o_p(1) = o_p(1)$$

and

$$\sqrt{n}\mathbb{E}\left[\mathbf{x}_i e_{1i} 1(\gamma_0 + \frac{v}{n} \wedge \gamma_0 < q_i \leq \gamma_0)\right] = \frac{|v|}{\sqrt{n}} f\left(\overline{\gamma}\right) \mathbb{E}\left[\mathbf{x}_i e_{1i} | q = \overline{\gamma}\right],$$

which is $o(1)$ by Assumptions (iv)(b) and (vii). Finally, when $v < 0$,

$$\begin{aligned}
T_3(u_2, v) &= \sum_{i=1}^n \overline{z}_1 \left(w_i | \beta_{20} + \tfrac{u_2}{\sqrt{n}}, \beta_{10}\right) 1(\gamma_0 + \tfrac{v}{n} < q_i \leq \gamma_0) \\
&= \sum_{i=1}^n \left(m_1(x_i, q_i) - \mathbf{x}_i' \tfrac{\beta_{10} + \left(\beta_{20} + u_2/\sqrt{n}\right)}{2}\right) \mathbf{x}_i' \left(\beta_{10} - \left(\beta_{20} + \tfrac{u_2}{\sqrt{n}}\right)\right) 1(\gamma_0 + \tfrac{v}{n} < q_i \leq \gamma_0) \\
&\quad + \sum_{i=1}^n \left(\delta_0 - \tfrac{u_2}{\sqrt{n}}\right)' \mathbf{x}_i \varepsilon_{1i} 1(\gamma_0 + \tfrac{v}{n} < q_i \leq \gamma_0) \\
&= \sum_{i=1}^n \left(m_1(x_i, q_i) - \mathbf{x}_i' \overline{\beta}_0\right) \mathbf{x}_i' \delta_0 1(\gamma_0 + \tfrac{v}{n} < q_i \leq \gamma_0) \\
&\quad - \tfrac{1}{\sqrt{n}} \sum_{i=1}^n \left(m_1(x_i, q_i) - \mathbf{x}_i' \overline{\beta}_0\right) \mathbf{x}_i' u_2 1(\gamma_0 + \tfrac{v}{n} < q_i \leq \gamma_0) \\
&\quad - \tfrac{1}{\sqrt{n}} \sum_{i=1}^n \tfrac{\delta_0' \mathbf{x}_i}{2} \mathbf{x}_i' u_2 1(\gamma_0 + \tfrac{v}{n} < q_i \leq \gamma_0) \\
&\quad + u_2' \left[\tfrac{1}{n} \sum_{i=1}^n \tfrac{\mathbf{x}_i \mathbf{x}_i'}{2} 1(\gamma_0 + \tfrac{v}{n} < q_i \leq \gamma_0)\right] u_2 \\
&\quad + \sum_{i=1}^n \delta_0' \mathbf{x}_i \varepsilon_{1i} 1(\gamma_0 + \tfrac{v}{n} < q_i \leq \gamma_0) - \tfrac{u_2'}{\sqrt{n}} \sum_{i=1}^n \mathbf{x}_i \varepsilon_{1i} 1(\gamma_0 + \tfrac{v}{n} < q_i \leq \gamma_0) \\
&= \sum_{i=1}^n \left(m_1(x_i, q_i) - \mathbf{x}_i' \overline{\beta}_0\right) \mathbf{x}' \delta_0 1(\gamma_0 + \tfrac{v}{n} < q_i \leq \gamma_0) + o_p(1) \\
&\quad + \sum_{i=1}^n \delta_0' \mathbf{x}_i \varepsilon_{1i} 1(\gamma_0 + \tfrac{v}{n} < q_i \leq \gamma_0) + o_p(1) \\
&= \sum_{i=1}^n \overline{z}_{1i} 1\left(\gamma_0 + \tfrac{v}{n} < q_i \leq \gamma_0\right) + o_p(1).
\end{aligned}$$

The $o_p(1)$ in the fourth equality need careful analysis. The second term

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n \left(m_1(x_i, q_i) - \mathbf{x}_i' \overline{\beta}_0\right) \mathbf{x}_i' u_2 1(\gamma_0 + \frac{v}{n} < q_i \leq \gamma_0) = o_p(1)$$

by the stochastic equicontinuity of $\gamma \mapsto \mathbb{G}_n\left(\left(m_1(x, q) - \mathbf{x}' \overline{\beta}_0\right) \mathbf{x}' u_2 1(\gamma < q \leq \gamma_0)\right)$ and

$$\sqrt{n}\mathbb{E}\left[\left(m_1(x_i, q_i) - \mathbf{x}_i' \overline{\beta}_0\right) \mathbf{x}_i' u_2 1(\gamma_0 + \frac{v}{n} < q_i \leq \gamma_0)\right] = \frac{|v|}{\sqrt{n}} f\left(\overline{\gamma}\right) \mathbb{E}\left[\left(m_1(x_i, q_i) - \mathbf{x}_i' \overline{\beta}_0\right) \mathbf{x}_i' u_2 | q = \overline{\gamma}\right],$$

which is $o(1)$ by Assumptions (iv)(b) and (vii), the third term can be similarly analyzed, the fourth term

$$\begin{aligned}
& u_2' \left[\tfrac{1}{n} \sum_{i=1}^n \tfrac{\mathbf{x}_i \mathbf{x}_i'}{2} 1(\gamma_0 + \tfrac{v}{n} < q_i \leq \gamma_0)\right] u_2 \\
&= u_2' \mathbb{E}\left[\tfrac{\mathbf{x}_i \mathbf{x}_i'}{2} 1(\gamma_0 + \tfrac{v}{n} < q_i \leq \gamma_0)\right] u_2 + o_p(1) \\
&= O\left(n^{-1}\right) + o_p(1) = o_p(1),
\end{aligned}$$

and the last term is $o_p(1)$ is by the stochastic equicontinuity of $\gamma \mapsto \mathbb{G}_n\left(\mathbf{x}\varepsilon_1 1(\gamma < q \leq \gamma_0)\right)$ and $\mathbb{E}\left[\mathbf{x}\varepsilon_1 1(\gamma < q \leq \gamma_0)\right] = 0$. ■

Lemma 9 *Under Assumption I(1)′, uniformly for h in a compact set,*

$$
\begin{aligned}
& nP_n\left(s\left(\cdot\left|\gamma_0+\frac{v}{a_n},\beta_0+\frac{u}{\sqrt{n}}\right.\right)-s\left(\cdot\left|\gamma_0,\beta_0\right.\right)\right) \\
= \; & \frac{1}{2}u_1'M_0u_1+\frac{1}{2}u_2'\overline{M}_0u_2-W_n\left(u\right)+C_n\left(v\right)+o_P\left(1\right),
\end{aligned}
$$

where $W_n(u)$ is defined in Lemma 8, and

$$
\begin{aligned}
C_n\left(v\right) &= nP_n\left(s\left(\cdot\left|\gamma_0+\frac{v}{n},\beta_0\right.\right)-s\left(\cdot\left|\gamma_0,\beta_0\right.\right)\right) \\
&= \sum_{i=1}^{n}\left(m_1(x_i,q_i)-\mathbf{x}_i'\beta_{10}+\frac{1}{2}\mathbf{x}_i'\delta_n+\varepsilon_{1i}\right)\mathbf{x}_i'\delta_n 1\left(\gamma_0+\frac{v}{a_n}<q_i\le\gamma_0\right) \\
&\quad -\sum_{i=1}^{n}\left(m_2(x_i,q_i)-\mathbf{x}_i'\beta_{20}-\frac{1}{2}\mathbf{x}_i'\delta_n+\varepsilon_{2i}\right)\mathbf{x}_i'\delta_n 1\left(\gamma_0<q_i\le\gamma_0+\frac{v}{a_n}\right).
\end{aligned}
$$

Proof. The proof is almost the same as Lemma 8, just replacing δ_0 by δ_n and n by a_n everywhere. Different from $D_n(v)$, the component

$$
\begin{array}{c}
\sum_{i=1}^{n}\left(m_1(x_i,q_i)-\mathbf{x}_i'\beta_{10}+\frac{1}{2}\mathbf{x}_i'\delta_n\right)\mathbf{x}_i'\delta_n 1\left(\gamma_0+\frac{v}{a_n}<q_i\le\gamma_0\right) \\
-\sum_{i=1}^{n}\left(m_2(x_i,q_i)-\mathbf{x}_i'\beta_{20}-\frac{1}{2}\mathbf{x}_i'\delta_n\right)\mathbf{x}_i'\delta_n 1\left(\gamma_0<q_i\le\gamma_0+\frac{v}{a_n}\right)
\end{array}
$$

in $C_n(v)$ will converge to a constant because $a_n \prec n$ so that averaging is involved. Take the second term as an example. Its variance is

$$
\begin{aligned}
& nVar\left(\left(m_2(x_i,q_i)-\mathbf{x}_i'\beta_{20}-\frac{1}{2}\mathbf{x}_i'\delta_n\right)\mathbf{x}_i'\delta_n 1\left(\gamma_0<q_i\le\gamma_0+\frac{v}{a_n}\right)\right) \\
= \; & n\left(\begin{array}{c}\mathbb{E}\left[\left(m_2(x_i,q_i)-\mathbf{x}_i'\beta_{20}-\frac{1}{2}\mathbf{x}_i'\delta_n\right)^2\left(\mathbf{x}_i'\delta_n\right)^2 1\left(\gamma_0<q_i\le\gamma_0+\frac{v}{a_n}\right)\right] \\ -\mathbb{E}\left[\left(m_2(x_i,q_i)-\mathbf{x}_i'\beta_{20}-\frac{1}{2}\mathbf{x}_i'\delta_n\right)\left(\mathbf{x}_i'\delta_n\right)1\left(\gamma_0<q_i\le\gamma_0+\frac{v}{a_n}\right)\right]^2\end{array}\right) \\
= \; & n\left(O\left(\mathbb{E}\left[\left\{\left(m_2(x_i,q_i)-\mathbf{x}_i'\beta_{20}\right)^4+\left(\mathbf{x}_i'\delta_n\right)^4\right\}1\left(\gamma_0<q_i\le\gamma_0+\frac{v}{a_n}\right)\right]\right)\right) \\
& +n\left(O\left(\mathbb{E}\left[\left\{\left(m_2(x_i,q_i)-\mathbf{x}_i'\beta_{20}\right)\left(\mathbf{x}_i'\delta_n\right)+\left(\mathbf{x}_i'\delta_n\right)^2\right\}1\left(\gamma_0<q_i\le\gamma_0+\frac{v}{a_n}\right)\right]^2\right)\right) \\
= \; & n\left(O\left(\frac{\|\delta_n\|^4}{a_n}\right)+O\left(\left(\frac{\|\delta_n\|^2}{a_n}\right)^2\right)\right)=O\left(\|\delta_n\|^2+n^{-1}\right)=o\left(1\right),
\end{aligned}
$$

uniformly in $v\in[0,\overline{v}]$ for any $\overline{v}>0$, where the second equality uses Cauchy-Schwarz inequality and $ab\le a^2+b^2$ for nonnegative numbers a and b, and the third equality is by Assumption (x)(a,b). As a result,

$$
\begin{aligned}
\sup_{v\in[0,\overline{v}]}&\left|\sum_{i=1}^{n}\left(m_2(x_i,q_i)-\mathbf{x}_i'\beta_{20}-\frac{1}{2}\mathbf{x}_i'\delta_n\right)\mathbf{x}_i'\delta_n 1\left(\gamma_0<q_i\le\gamma_0+\frac{v}{a_n}\right)\right. \\
&\left.-\frac{n}{a_n}\int_0^v\left[\left(-\|\delta_n\|\dot{\varsigma}_2\left(\gamma_0+\frac{v}{a_n}\right)'\delta_n-\frac{1}{2}\delta_n'D_{\gamma_0+\frac{\nu}{a_n}}\delta_n\right)f\left(\gamma_0+\frac{\nu}{a_n}\right)\right]d\nu\right|\xrightarrow{P}0,
\end{aligned}
$$

where by the continuity of $\dot{\varsigma}_2(\gamma)$ and D_γ in γ at γ_0, we have

$$\frac{n}{a_n}\int_0^v \left[\left(-\left\|\delta_n\right\|\dot{\varsigma}_2\left(\gamma_0+\frac{\nu}{a_n}\right)'\delta_n-\frac{1}{2}\delta_n' D_{\gamma_0+\frac{\nu}{a_n}}\delta_n\right)f\left(\gamma_0+\frac{\nu}{a_n}\right)\right]d\nu \to \left(-c'\dot{\varsigma}_{20}-\frac{1}{2}c'D_0c\right)f_0v$$

uniformly in $v\in[0,\overline{v}]$. Similarly,

$$\sum\nolimits_{i=1}^n \left(m_1(x_i,q_i)-\mathbf{x}_i'\beta_{10}+\frac{1}{2}\mathbf{x}_i'\delta_n\right)\mathbf{x}_i'\delta_n 1\left(\gamma_0+\frac{v}{a_n}<q_i\le\gamma_0\right)\xrightarrow{p}\left(c'\dot{\varsigma}_{10}+\frac{1}{2}c'D_0c\right)f_0\left|v\right|$$

uniformly in $v\in[\underline{v},0]$ for any $\underline{v}<0$. In summary,

$$C_n(v)=\begin{cases}\frac{1}{2}f_0\left(c'D_0c+2c'\dot{\varsigma}_{10}\right)|v|+\sum\limits_{i=1}^n\delta_n'\mathbf{x}_i\varepsilon_{1i}1(\gamma_0+\frac{v}{a_n}<q_i\le\gamma_0), & \text{if } v\le 0,\\ \frac{1}{2}f_0\left(c'D_0c+2c'\dot{\varsigma}_{20}\right)v-\sum\limits_{i=1}^n\delta_n'\mathbf{x}_i\varepsilon_{2i}1(\gamma_0<q_i\le\gamma_0+\frac{v}{a_n}), & \text{if } v>0,\end{cases}+o_p(1).$$

■

Lemma 10 *Under Assumption I(α), $1<\alpha<2$, uniformly for h in a compact set,*

$$\begin{aligned}&\sqrt{n\rho_n}P_n\left(s\left(\cdot\left|\gamma_0+\frac{v}{\rho_n},\beta_0+\frac{u}{\kappa_n}\right.\right)-s\left(\cdot\left|\gamma_0,\beta_0\right.\right)\right)\\ =\;&\frac{1}{2}u_1'M_0u_1+\frac{1}{2}u_2'\overline{M}_0u_2+\lambda_-\left|v\right|^\alpha 1(v\le 0)+\lambda_+v^\alpha 1(v>0)+\Xi_n(v)+o_P(1),\end{aligned}$$

where

$$\Xi_n(v)=\begin{cases}\sqrt{\frac{\rho_n}{n}}\sum\limits_{i=1}^n\left[\overline{z}_{1i}1(\gamma_0+\frac{v}{\rho_n}<q_i\le\gamma_0)-\Lambda_-\left(\frac{v}{\rho_n}\right)\right], & \text{if } v\le 0,\\ \sqrt{\frac{\rho_n}{n}}\sum\limits_{i=1}^n\left[\overline{z}_{2i}1(\gamma_0<q_i\le\gamma_0+\frac{v}{\rho_n})-\Lambda_+\left(\frac{v}{\rho_n}\right)\right], & \text{if } v>0.\end{cases}$$

Under Under Assumption I(2), uniformly for h in a compact set,

$$\begin{aligned}&n^{2/3}P_n\left(s\left(\cdot\left|\gamma_0+\frac{v}{n^{1/3}},\beta_0+\frac{u}{n^{1/3}}\right.\right)-s\left(\cdot\left|\gamma_0,\beta_0\right.\right)\right)\\ =\;&\frac{1}{2}h'S^-_{\theta\theta}h1(v\le 0)+\frac{1}{2}h'S^+_{\theta\theta}h1(v>0)+\Xi_n(v)+o_P(1).\end{aligned}$$

where $\Xi_n(v)$ is the same as above except replacing ρ_n by $n^{1/3}$.

Proof. First, check the cases with $1<\alpha<2$. Because $\kappa_n=(n\rho_n)^{1/4}$ and $n^{1/3}\prec\rho_n\prec n$, we have

$$\frac{\sqrt{\rho_n}}{\kappa_n}=\sqrt[4]{\frac{\rho_n}{n}}=o(1)\text{ and }\frac{\sqrt{\rho_n}}{\kappa_n}\frac{\sqrt{n}}{\rho_n}=\sqrt[4]{\frac{n}{\rho_n^3}}=o(1).$$

Now,

$$\begin{aligned}T_1(u_1)&=\frac{\sqrt{n\rho_n}}{n}\sum_{i=1}^n T\left(w_i|\beta_{10}+\frac{u_1}{\kappa_n},\beta_{10}\right)1(q_i\le\gamma_0)\\ &=\frac{\sqrt{n\rho_n}}{n}\sum_{i=1}^n-\left(e_{1i}-\frac{\mathbf{x}_i'}{2}\frac{u_1}{\kappa_n}\right)\mathbf{x}_i'\frac{u_1}{\kappa_n}1(q_i\le\gamma_0).\end{aligned}$$

and

$$\frac{\sqrt{n\rho_n}}{n}\frac{u_1'}{2}\left(\frac{1}{\kappa_n^2}\sum_{i=1}^n \mathbf{x}_i\mathbf{x}_i'1(q_i\le\gamma_0)\right)u_1 = \frac{u_1'}{2}\left(\frac{1}{n}\sum_{i=1}^n \mathbf{x}_i\mathbf{x}_i'1(q_i\le\gamma_0)\right)u_1 = \frac{1}{2}u_1'M_0u_1+o_p(1),$$

$$\frac{\sqrt{n\rho_n}}{n}\frac{u_1'}{\kappa_n}\sum_{i=1}^n \mathbf{x}_ie_{1i}1(q_i\le\gamma_0) = \frac{\sqrt{\rho_n}}{\kappa_n}\frac{u_1'}{\sqrt{n}}\sum_{i=1}^n \mathbf{x}_ie_{1i}1(q_i\le\gamma_0) = o_p(1).$$

In other words, the random part in $T_1(u_1)$ disappears. Next,

$$T_4(u_1,v) = \frac{u_1'\mathbb{E}\left[\mathbf{x}_i\mathbf{x}_i'1(\gamma_0+\frac{v}{\rho_n}<q_i\le\gamma_0)\right]u_1}{2} - \frac{\sqrt{\rho_n}}{\kappa_n}\frac{u_1'}{\sqrt{n}}\sum_{i=1}^n \mathbf{x}_ie_{1i}1(\gamma_0+\tfrac{v}{\rho_n}<q_i\le\gamma_0)+o_p(1)=o_p(1),$$

where

$$\mathbb{E}\left[\mathbf{x}_i\mathbf{x}_i'1(\gamma_0+\tfrac{v}{\rho_n}<q_i\le\gamma_0)\right] = \tfrac{|v|}{\rho_n}f(\overline{\gamma})\,\mathbb{E}\left[\mathbf{x}_i\mathbf{x}_i'|q=\overline{\gamma}\right]=o(1),$$

and

$$\begin{aligned}
&\tfrac{\sqrt{\rho_n}}{\kappa_n}\tfrac{u_1'}{\sqrt{n}}\sum_{i=1}^n \mathbf{x}_ie_{1i}1(\gamma_0+\tfrac{v}{\rho_n}<q_i\le\gamma_0)\\
&\approx \tfrac{\sqrt{\rho_n}}{\kappa_n}\mathbb{G}_n\left(u_1'\mathbf{x}e_11(\gamma_0+\tfrac{v}{\rho_n}<q\le\gamma_0)\right)+\sqrt{n\rho_n}\Psi_-\left(\beta_{10}+\tfrac{u_1}{\kappa_n},\gamma_0+\tfrac{v}{\rho_n}\right)\\
&= \tfrac{\sqrt{\rho_n}}{\kappa_n}o_p(1)+\tfrac{\sqrt{n\rho_n}}{\kappa_n\rho_n}u_1'S^-_{\beta_1\gamma}v\\
&= o_p(1)+o(1)=o_p(1).
\end{aligned}$$

with $\approx$ in the first equality because a smaller term $-\frac{u_1'\mathbb{E}\left[\mathbf{x}_i\mathbf{x}_i'1(\gamma_0+\frac{v}{\rho_n}<q_i\le\gamma_0)\right]u_1}{2}$ is excluded from $\sqrt{n\rho_n}\Psi_-\left(\beta_1+\frac{u_1}{\kappa_n},\gamma+\frac{v}{\rho_n}\right)$. In other words, $T_4(u_1,v)$ can be neglected. Finally,

$$\begin{aligned}
&T_3(u_2,v) = \tfrac{\sqrt{n\rho_n}}{n}\sum_{i=1}^n \overline{z}_1\left(w_i|\beta_{20}+\tfrac{u_2}{\kappa_n},\beta_{10}\right)1(\gamma_0+\tfrac{v}{\rho_n}<q_i\le\gamma_0)\\
&= \tfrac{\sqrt{n\rho_n}}{n}\sum_{i=1}^n\left(m_1(x_i,q_i)-\mathbf{x}_i'\tfrac{\beta_{10}+(\beta_{20}+u_2/\kappa_n)}{2}\right)\mathbf{x}_i'\left(\beta_{10}-\left(\beta_{20}+\tfrac{u_2}{\kappa_n}\right)\right)1(\gamma_0+\tfrac{v}{\rho_n}<q_i\le\gamma_0)\\
&\quad+\tfrac{\sqrt{n\rho_n}}{n}\sum_{i=1}^n\left(\delta_0-\tfrac{u_2}{\kappa_n}\right)'\mathbf{x}_i\varepsilon_{1i}1(\gamma_0+\tfrac{v}{\rho_n}<q_i\le\gamma_0)\\
&= \tfrac{\sqrt{n\rho_n}}{n}\sum_{i=1}^n\left(m_1(x_i,q_i)-\mathbf{x}_i'\overline{\beta}_0\right)\mathbf{x}_i'\delta_01(\gamma_0+\tfrac{v}{\rho_n}<q_i\le\gamma_0)\\
&-\tfrac{\sqrt{n\rho_n}}{n\kappa_n}\sum_{i=1}^n\left(m_1(x_i,q_i)-\mathbf{x}_i'\overline{\beta}_0\right)\mathbf{x}_i'u_21(\gamma_0+\tfrac{v}{\rho_n}<q_i\le\gamma_0)\\
&-\tfrac{\sqrt{n\rho_n}}{n\kappa_n}\sum_{i=1}^n\tfrac{\delta_0'\mathbf{x}_i}{2}\mathbf{x}_i'u_21(\gamma_0+\tfrac{v}{\rho_n}<q_i\le\gamma_0)\\
&+\tfrac{\sqrt{n\rho_n}}{n\kappa_n^2}\sum_{i=1}^n\tfrac{u_2'\mathbf{x}_i\mathbf{x}_i'u_2}{2}1(\gamma_0+\tfrac{v}{\rho_n}<q_i\le\gamma_0)\\
&+\tfrac{\sqrt{n\rho_n}}{n}\sum_{i=1}^n\delta_0'\mathbf{x}_i\varepsilon_{1i}1(\gamma_0+\tfrac{v}{\rho_n}<q_i\le\gamma_0)-\tfrac{\sqrt{n\rho_n}}{n\kappa_n}\sum_{i=1}^n u_2'\mathbf{x}_i\varepsilon_{1i}1(\gamma_0+\tfrac{v}{\rho_n}<q_i\le\gamma_0)\\
&= \sqrt{n\rho_n}\Lambda_-\left(\tfrac{v}{\rho_n}\right)\ +\sqrt{n\rho_n}\Psi_-\left(\beta_{20}+\tfrac{u_2}{\kappa_n},\gamma_0+\tfrac{v}{\rho_n}\right)\\
&+\tfrac{\sqrt{n\rho_n}}{n}\sum_{i=1}^n\left[\left(m_1(x_i,q_i)-\mathbf{x}_i'\overline{\beta}_0\right)\mathbf{x}_i'\delta_01(\gamma_0+\tfrac{v}{\rho_n}<q_i\le\gamma_0)-\Lambda_-\left(\tfrac{v}{\rho_n}\right)\right]\\
&+\tfrac{\sqrt{n\rho_n}}{n}\sum_{i=1}^n\delta_0'\mathbf{x}_i\varepsilon_{1i}1(\gamma_0+\tfrac{v}{\rho_n}<q_i\le\gamma_0)+o_p(1)\\
&= \lambda_-|v|^\alpha+\tfrac{\sqrt{n\rho_n}}{n}\sum_{i=1}^n\left[\overline{z}_{1i}1(\gamma_0+\tfrac{v}{\rho_n}<q_i\le\gamma_0)-\Lambda_-\left(\tfrac{v}{\rho_n}\right)\right]+o_p(1).
\end{aligned}$$

The second to last equality need careful analysis. The variance of the first term in the third equality is

$$
\begin{aligned}
&\rho_n Var\left(\left(m_1(x_i,q_i)-\mathbf{x}_i'\overline{\beta}_0\right)\mathbf{x}_i'\delta_0 1(\gamma_0+\tfrac{v}{\rho_n}<q_i\le\gamma_0)\right)\\
&=\rho_n\left[\mathbb{E}\left[\left(m_1(x_i,q_i)-\mathbf{x}_i'\overline{\beta}_0\right)^2\left(\mathbf{x}_i'\delta_0\right)^2 1(\gamma_0+\tfrac{v}{\rho_n}<q_i\le\gamma_0)\right]-\mathbb{E}\left[\left(m_1(x_i,q_i)-\mathbf{x}_i'\overline{\beta}_0\right)\mathbf{x}_i'\delta_0 1(\gamma_0+\tfrac{v}{\rho_n}<q_i\le\gamma_0)\right]^2\right]\\
&=O\left(\rho_n\rho_n^{-1}\right)-O\left(\rho_n\rho_n^{-2\alpha}L\left(1/\rho_n\right)\right)=O\left(1\right).
\end{aligned}
$$

[19]

where because $\mathbb{E}\left[\left(m_1(x_i,q_i)-\mathbf{x}_i'\overline{\beta}_0\right)^2\left(\mathbf{x}_i'\delta_0\right)^2 1(\gamma<q_i\le\gamma_0)\right]$ need not be $RV_{2\alpha-1}$, the second moment is not $o(1)$ and the variance is not $o(1)$. If $\mathbb{E}\left[\left(m_1(x_i,q_i)-\mathbf{x}_i'\overline{\beta}_0\right)^2\left(\mathbf{x}_i'\delta_0\right)^2 1(\gamma<q_i\le\gamma_0)\right]\in RV_{2\alpha-1}$, the second moment is $\rho_n\rho_n^{-(2\alpha-1)}L\left(1/\rho_n\right)=o(1)$, i.e., the variance is $o(1)$. In general, this term remains random and its variance is the same as its second moment. The analysis here also shows that when $\alpha=1$, $\mathbb{E}\left[\left(m_1(x_i,q_i)-\mathbf{x}_i'\overline{\beta}_0\right)^2\left(\mathbf{x}_i'\delta_0\right)^2 1(\gamma<q_i\le\gamma_0)\right]\in RV_1$, the variance of the first term is $O(1)$ in whatever sense; as a result, it remains random in $D_n(v)$ of Lemma 8. Similarly, the variance of the second term in the third equality is

$$
\begin{aligned}
&\tfrac{\rho_n}{\kappa_n^2}\left(\mathbb{E}\left[\left(m_1(x_i,q_i)-\mathbf{x}_i'\overline{\beta}_0\right)^2\left(\mathbf{x}_i'u_2\right)^2 1(\gamma_0+\tfrac{v}{\rho_n}<q_i\le\gamma_0)\right]-\mathbb{E}\left[\left(m_1(x_i,q_i)-\mathbf{x}_i'\overline{\beta}_0\right)\left(\mathbf{x}_i'u_2\right)1(\gamma_0+\tfrac{v}{\rho_n}<q_i\le\gamma_0)\right]^2\right)\\
&=O\left(\tfrac{\rho_n}{\kappa_n^2}\left(\tfrac{1}{\rho_n}-\tfrac{1}{\rho_n^2}\right)\right)=o\left(1\right),
\end{aligned}
$$

the third term can be similarly analyzed, and the fourth term is

$$
\tfrac{\sqrt{n\rho_n}}{n\kappa_n^2}\sum_{i=1}^n\tfrac{u_2'\mathbf{x}_i\mathbf{x}_i'u_2}{2}1(\gamma_0+\tfrac{v}{\rho_n}<q_i\le\gamma_0)=u_2'\left[\tfrac{1}{n}\sum_{i=1}^n\tfrac{\mathbf{x}_i\mathbf{x}_i'}{2}1(\gamma_0+\tfrac{v}{\rho_n}<q_i\le\gamma_0)\right]u_2.
$$

which is similar to the counterpart in I(1) so is $o_p(1)$. The last term is

$$
\frac{\sqrt{n\rho_n}}{n\kappa_n}\sum_{i=1}^n u_2'\mathbf{x}_i\varepsilon_{1i}1(\gamma_0+\frac{v}{\rho_n}<q_i\le\gamma_0)=\frac{\sqrt{\rho_n}}{\kappa_n}u_2'\mathbb{G}_n\left(\mathbf{x}\varepsilon 1(\gamma_0+\frac{v}{\rho_n}<q\le\gamma_0)\right)=o_p(1)
$$

since $\sqrt{\rho_n}/\kappa_n=(\rho_n/n)^{1/4}=o(1)$. Because $\sqrt{n\rho_n}\Psi_-\left(\beta_{20}+\frac{u_2}{\kappa_n},\gamma_0+\frac{v}{\rho_n}\right)=O\left(\frac{\sqrt{n\rho_n}}{\kappa_n\rho_n}\right)=o(1)$, it can be neglected. In summary, the deterministic part is $\frac{1}{2}u_1'M_0u_1+\frac{1}{2}u_2'\overline{M}_0u_2+\lambda_-\left|v\right|^{\alpha}$, and the random part is

$$
\begin{aligned}
&\tfrac{\sqrt{n\rho_n}}{n}\sum_{i=1}^n\left[\left(m_1(x_i,q_i)-\mathbf{x}_i'\overline{\beta}_0\right)\mathbf{x}_i'\delta_0 1(\gamma_0+\tfrac{v}{\rho_n}<q_i\le\gamma_0)-\Lambda_-\left(\tfrac{v}{\rho_n}\right)\right]+\tfrac{\sqrt{n\rho_n}}{n}\sum_{i=1}^n\delta_0'\mathbf{x}_i\varepsilon_{1i}1(\gamma_0+\tfrac{v}{\rho_n}<q_i\le\gamma_0)\\
&=\tfrac{\sqrt{n\rho_n}}{n}\sum_{i=1}^n\left[\left(y_i-\mathbf{x}_i'\overline{\beta}_0\right)\mathbf{x}_i'\delta_0 1(\gamma_0+\tfrac{v}{\rho_n}<q_i\le\gamma_0)-\Lambda_-\left(\tfrac{v}{\rho_n}\right)\right]=\tfrac{\sqrt{n\rho_n}}{n}\sum_{i=1}^n\left[\overline{z}_{1i}1(\gamma_0+\tfrac{v}{\rho_n}<q_i\le\gamma_0)-\Lambda_-\left(\tfrac{v}{\rho_n}\right)\right].
\end{aligned}
$$

Second, check the case with $\alpha=2$, where $\rho_n=\kappa_n=n^{1/3}$. We only mark the differences from the analyses above. Now, $T_4(u_1,v)$ cannot be neglectable because $\frac{\sqrt{\rho_n}}{\kappa_n}\frac{\sqrt{n}}{\rho_n}=1$. As a result,

$$
T_4(u_1,v)=n^{2/3}\Psi_-\left(\beta_{10}+\frac{u_1}{n^{1/3}},\gamma_0+\frac{v}{n^{1/3}}\right)+o_p(1)=u_1'S_{\beta_1\gamma}^-v.
$$

The second and third terms in $T_3(u_2,v)$ are not neglectable either. It turns out that

$$
n^{2/3}\Psi_-\left(\beta_{20}+\frac{u_2}{n^{1/3}},\gamma_0+\frac{v}{n^{1/3}}\right)=u_2'S_{\beta_2\gamma}^-v.
$$

[19] Note that the second moment is not $\rho_nE\left[\left(m_1(x_i,q_i)-\mathbf{x}_i'\overline{\beta}_0\right)^2\left(\mathbf{x}_i'\delta_0\right)^2 1(\gamma_0+\frac{v}{\rho_n}<q_i\le\gamma_0)\right]$ because the cross terms do not disappear. Otherwise, the second moment is zero and the variance is zero, but the mean is not zero!

In summary, the deterministic part is

$$\frac{1}{2}u_1'M_0u_1+\frac{1}{2}u_2'\overline{M}_0u_2+u_1'S^-_{\beta_1\gamma}v+u_2'S^-_{\beta_2\gamma}v+\lambda_-v^2=\frac{1}{2}h'S^-_{\theta\theta}h.$$

■

Lemma 11 *Under Assumption II(2), uniformly for h in a compact set,*

$$\begin{aligned}&nP_n\left(s\left(\cdot\left|\gamma_0+\frac{v}{n^{1/2}},\beta_0+\frac{u}{n^{1/2}}\right.\right)-s\left(\cdot\left|\gamma_0,\beta_0\right.\right)\right)\\=&\ \frac{1}{2}h'S^-_{\theta\theta}h1(v\le 0)+\frac{1}{2}h'S^+_{\theta\theta}h1(v>0)-W_n(u)+o_P(1);\end{aligned}$$

under Assumption II(α), $2<\alpha<3$, uniformly for h in a compact set,

$$\begin{aligned}&nP_n\left(s\left(\cdot\left|\gamma_0+\frac{v}{\rho_n},\beta_0+\frac{u}{n^{1/2}}\right.\right)-s\left(\cdot\left|\gamma_0,\beta_0\right.\right)\right)\\=&\ \frac{1}{2}u_1'M_0u_1+\frac{1}{2}u_2'\overline{M}_0u_2+\lambda_-\left|v\right|^{\alpha}1(v\le 0)+\lambda_+v^{\alpha}1(v>0)-W_n(u)+o_P(1);\end{aligned}$$

and under under Assumption II(3), uniformly for h in a compact set,

$$\begin{aligned}&nP_n\left(s\left(\cdot\left|\gamma_0+\frac{v}{n^{1/3}},\beta_0+\frac{u}{n^{1/2}}\right.\right)-s\left(\cdot\left|\gamma_0,\beta_0\right.\right)\right)\\=&\ \frac{1}{2}u_1'M_0u_1+\frac{1}{2}u_2'\overline{M}_0u_2+\lambda_-\left|v\right|^{3}1(v\le 0)+\lambda_+v^{3}1(v>0)-W_n(u)+\Xi_n(v)+o_P(1),\end{aligned}$$

where $W_n(u)$ is defined in Lemma 8, and

$$\Xi_n(v)=\begin{cases}\sum_{i=1}^n\left[\overline{z}_{1i}1(\gamma_0+\frac{v}{n^{1/3}}<q_i\le\gamma_0)-\Lambda_-\left(\frac{v}{n^{1/3}}\right)\right] & \text{if } v\le 0,\\ \sum_{i=1}^n\left[\overline{z}_{2i}1(\gamma_0<q_i\le\gamma_0+\frac{v}{n^{1/3}})-\Lambda_+\left(\frac{v}{n^{1/3}}\right)\right] & \text{if } v>0.\end{cases}$$

Under Assumption II(α), $3<\alpha<4$, uniformly for h in a compact set,

$$\begin{aligned}&\sqrt{n\rho_n^3}P_n\left(s\left(\cdot\left|\gamma_0+\frac{v}{\rho_n},\beta_0+\frac{u}{\kappa_n}\right.\right)-s\left(\cdot\left|\gamma_0,\beta_0\right.\right)\right)\\=&\ \frac{1}{2}u_1'M_0u_1+\frac{1}{2}u_2'\overline{M}_0u_2+\lambda_-\left|v\right|^{\alpha}1(v\le 0)+\lambda_+v^{\alpha}1(v>0)+\Xi_n(v)+o_P(1),\end{aligned}$$

where

$$\Xi_n(v)=\begin{cases}\frac{\sqrt{\rho_n^3}}{\sqrt{n}}\sum_{i=1}^n\left[\overline{z}_{1i}1(\gamma_0+\frac{v}{\rho_n}<q_i\le\gamma_0)-\Lambda_-\left(\frac{v}{\rho_n}\right)\right] & \text{if } v\le 0,\\ \frac{\sqrt{\rho_n^3}}{\sqrt{n}}\sum_{i=1}^n\left[\overline{z}_{2i}1(\gamma_0<q_i\le\gamma_0+\frac{v}{\rho_n})-\Lambda_+\left(\frac{v}{\rho_n}\right)\right] & \text{if } v>0;\end{cases}$$

and under Assumption II(4), uniformly for h in a compact set,

$$\begin{aligned}&n^{4/5}P_n\left(s\left(\cdot\left|\gamma_0+\frac{v}{n^{1/5}},\beta_0+\frac{u}{n^{2/5}}\right.\right)-s\left(\cdot\left|\gamma_0,\beta_0\right.\right)\right)\\=&\ \frac{1}{2}\left(u',v^2\right)\mathbb{S}^-\left(u',v^2\right)'1(v\le 0)+\frac{1}{2}\left(u',v^2\right)\mathbb{S}^+\left(u',v^2\right)'1(v>0)+\Xi_n(v)+o_P(1),\end{aligned}$$

where $\Xi_n(v)$ is the same as above except replacing ρ_n by $n^{1/5}$ (i.e., $\sqrt{\rho_n^3}/\sqrt{n}$ is replaced by $n^{-1/5}$).

Proof. From the analyses in the previous two lemmas, we can see that for $T_1(u_1)$, we need pay attention to the scaled $\Phi\left(\beta_{10}+\frac{u_1}{\kappa_n}\right)$ and the random part, for $T_4(u_2,v)$, only the scaled $\Psi_-\left(\beta_{10}+\frac{u_1}{\kappa_n},\gamma_0+\frac{v}{\rho_n}\right)$, and for $T_3(u_2,v)$, the scaled $\Psi_-\left(\beta_{20}+\frac{u_2}{\kappa_n},\gamma_0+\frac{v}{\rho_n}\right)+\Lambda_-\left(\gamma_0+\frac{v}{\rho_n}\right)$ and the random part. The scaled $\Psi_-\left(\beta_{10}+\frac{u_1}{\kappa_n},\gamma_0+\frac{v}{\rho_n}\right)$ and $\Psi_-\left(\beta_{20}+\frac{u_2}{\kappa_n},\gamma_0+\frac{v}{\rho_n}\right)$ may be neglectable, but the scaled $\Phi\left(\beta_{10}+\frac{u_1}{\kappa_n}\right)$ and $\Lambda_-\left(\gamma_0+\frac{v}{\rho_n}\right)$ cannot be neglected. The random part in either $T_1(u_1)$ or $T_3(u_2,v)$ may be neglectable.

When $2\le\alpha\le3$, $T_1(u_1)$ is the same as in Lemma 8. For $T_4(u_1,v)$, we can show the randomness disappears as in Lemma 8. Now,

$$n\Psi_-\left(\beta_{10}+\frac{u_1}{\sqrt{n}},\gamma_0+\frac{v}{\rho_n}\right)\to u_1'S^-_{\beta_1\gamma}v$$

when $\alpha=2$. When $2<\alpha\le3$, by Taylor expansion, it is

$$O\left(S^-_{\beta_1\gamma}\frac{n}{\sqrt{n}\rho_n}+S^-_{\beta_1\beta_1\gamma}\frac{n}{n\rho_n}+\frac{n}{\sqrt{n}}S^-_{\beta_1}\left(\gamma_0+\frac{v}{\rho_n}\right)\right)=O\left(\sqrt{n}\left(\frac{1}{\rho_n}\right)^2+\sqrt{n}\left(\frac{1}{\rho_n}\right)^{\alpha-1}L\left(\frac{1}{\rho_n}\right)\right)=o(1),$$

where note that $S^-_{\beta_1\gamma}=\mathbf{0}$. In $T_3(u_2,v)$, we can show the variances of various terms shrink to zero under Assumption (x)(b) as in Lemma 10. As a result, we need only check

$$n\Psi_-\left(\beta_{20}+\frac{u_2}{\sqrt{n}},\gamma_0+\frac{v}{\rho_n}\right)\to u_2'S^-_{\beta_2\gamma}v$$

when $\alpha=2$ and is neglectable when $2<\alpha\le3$, and

$$n\Lambda_-\left(\gamma_0+\frac{v}{\rho_n}\right)\to\lambda_-\left|v\right|^\alpha$$

by the definition of ρ_n. The variance of the random terms

$$\sum_{i=1}^n\left(m_1(x_i,q_i)-\mathbf{x}_i'\overline{\beta}_0\right)\mathbf{x}_i'\delta_0 1(\gamma_0+\frac{v}{\rho_n}<q_i\le\gamma_0)=\sum_{i=1}^n\left(m_1(x_i,q_i)-\mathbf{x}_i'\overline{\beta}_0\right)q_i\delta_{q0}1(\gamma_0+\frac{v}{\rho_n}<q_i\le\gamma_0)$$

and

$$\sum_{i=1}^n\delta_0'\mathbf{x}_i\varepsilon_{1i}1(\gamma_0+\frac{v}{\rho_n}<q_i\le\gamma_0)=\sum_{i=1}^n\delta_{q0}q_i\varepsilon_{1i}1(\gamma_0+\frac{v}{\rho_n}<q_i\le\gamma_0)$$

is $O\left(\frac{n}{\rho_n^3}\right)$ which is neglectable unless $\alpha=3$. In summary, the deterministic part converges to $\frac{1}{2}h'S^-_{\theta\theta}h$ when $\alpha=2$ and $\frac{1}{2}u_1'M_0u_1+\frac{1}{2}u_2'\overline{M}_0u_2+\lambda_-\left|v\right|^\alpha$ when $2<\alpha\le3$, and the random part converges to $-W_n(u)$ when $2\le\alpha<3$ and $-W_n(u)+\Xi_n(v)$ when $\alpha=3$.

When $3<\alpha\le4$, $\kappa_n=\left(n\rho_n^3\right)^{1/4}$ and $n^{1/5}\preceq\rho_n\prec n^{1/3}$, so $\frac{\sqrt{\rho_n^3}}{\kappa_n}=\frac{\rho_n^{3/4}}{n^{1/4}}=o(1)$. Now,

$$\begin{aligned}T_1(u_1)&=\frac{\sqrt{n\rho_n^3}}{n}\sum_{i=1}^n T\left(w_i|\beta_{10}+\frac{u_1}{\kappa_n},\beta_{10}\right)1(q_i\le\gamma_0)\\&=\frac{\sqrt{n\rho_n^3}}{n}\sum_{i=1}^n-\left(e_{1i}-\frac{\mathbf{x}_i'}{2}\frac{u_1}{\kappa_n}\right)\mathbf{x}_i'\frac{u_1}{\kappa_n}1(q_i\le\gamma_0),\end{aligned}$$

and

$$\begin{aligned}
\frac{\sqrt{n\rho_n^3}}{n}\frac{u_1'}{2}\left(\frac{1}{\kappa_n^2}\sum_{i=1}^{n}\mathbf{x}_i\mathbf{x}_i'1(q_i\le\gamma_0)\right)u_1 &= \frac{u_1'}{2}\left(\frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_i\mathbf{x}_i'1(q_i\le\gamma_0)\right)u_1=\frac{1}{2}u_1'M_0u_1+o_p(1),\\
\frac{\sqrt{n\rho_n^3}}{n}\frac{u_1'}{\kappa_n}\sum_{i=1}^{n}\mathbf{x}_ie_{1i}1(q_i\le\gamma_0) &= \frac{\sqrt{\rho_n^3}}{\kappa_n}\frac{u_1'}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{x}_ie_{1i}1(q_i\le\gamma_0)=o_p(1).
\end{aligned}$$

In other words, the random part in $T_1(u_1)$ disappears. In $T_4(u_1,v)$, by Taylor expansion,

$$\begin{aligned}
\sqrt{n\rho_n^3}\Psi_-\left(\beta_{10}+\frac{u_1}{\kappa_n},\gamma_0+\frac{v}{\rho_n}\right) &= O\left(S_{\beta_1\gamma}^-\frac{\sqrt{n\rho_n^3}}{\kappa_n\rho_n}+S_{\beta_1\beta_1\gamma}^-\frac{\sqrt{n\rho_n^3}}{\kappa_n^2\rho_n}+\frac{\sqrt{n\rho_n^3}}{\kappa_n}S_{\beta_1}^-\left(\gamma_0+\frac{v}{\rho_n}\right)\right)\\
&= O\left(\frac{\sqrt{n\rho_n}}{\kappa_n^2}+\frac{\sqrt{n\rho_n^3}}{\kappa_n}S_{\beta_1\gamma\gamma}^-\left(\frac{1}{\rho_n}\right)^2+\frac{\sqrt{\rho_n^3}}{\kappa_n}\sqrt{n}\left(\frac{1}{\rho_n}\right)^{\alpha-1}L\left(\frac{1}{\rho_n}\right)\right)\\
&= O\left(\frac{\sqrt{n\rho_n^3}}{\kappa_n}S_{\beta_1\gamma\gamma}^-\left(\frac{1}{\rho_n}\right)^2\right)=O\left(\frac{\sqrt{n}}{\sqrt{\rho_n}\kappa_n}\right)
\end{aligned}$$

which is $o(1)$ unless $\alpha=4$ where $\sqrt{n\rho_n^3}\Psi_-\left(\beta_{10}+\frac{u_1}{\kappa_n},\gamma_0+\frac{v}{\rho_n}\right)$ converges to $\frac{1}{2}u_2'S_{\beta_1\gamma^2}^-v^2$. In $T_3(u_2,v)$,

$$\begin{aligned}
\sqrt{n\rho_n^3}\Psi_-\left(\beta_{20}+\frac{u_2}{\kappa_n},\gamma_0+\frac{v}{\rho_n}\right) &= O\left(S_{\beta_2\gamma}^-\frac{\sqrt{n\rho_n^3}}{\kappa_n\rho_n}+S_{\beta_2\beta_2\gamma}^-\frac{\sqrt{n\rho_n^3}}{\kappa_n^2\rho_n}+\frac{\sqrt{n\rho_n^3}}{\kappa_n}S_{\beta_2}^-\left(\gamma_0+\frac{v}{\rho_n}\right)\right)\\
&= O\left(\frac{\sqrt{n\rho_n^3}}{\kappa_n}S_{\beta_2\gamma\gamma}^-\left(\frac{1}{\rho_n}\right)^2\right)=O\left(\frac{\sqrt{n}}{\sqrt{\rho_n}\kappa_n}\right),
\end{aligned}$$

which is $o(1)$ unless $\alpha=4$ where $\sqrt{n\rho_n^3}\Psi_-\left(\beta_{20}+\frac{u_2}{\kappa_n},\gamma_0+\frac{v}{\rho_n}\right)$ converges to $\frac{1}{2}u_2'S_{\beta_2\gamma^2}^-v^2$, and

$$\sqrt{n\rho_n^3}\Lambda_-\left(\gamma_0+\frac{v}{\rho_n}\right)\to\lambda_-|v|^\alpha.$$

As to the random part in $T_3(u_2,v)$,

$$\frac{\sqrt{n\rho_n^3}}{n}\sum_{i=1}^{n}\delta_{q0}q_i\left(y_i-\mathbf{x}_i'\overline{\beta}_0\right)1(\gamma_0+\frac{v}{\rho_n}<q_i\le\gamma_0),$$

whose variance is $O(1)$ so is not neglectable. In summary, the deterministic part converges to $\frac{1}{2}\left(u',v^2\right)\mathbb{S}^-\left(u',v^2\right)'$ when $\alpha=4$ and $\frac{1}{2}u_1'M_0u_1+\frac{1}{2}u_2'\overline{M}_0u_2+\lambda_-|v|^\alpha$ when $3<\alpha<4$, and the random part converges to $\Xi_n(v)$ for any $3<\alpha\le4$. ■

Appendix D: Weak Convergence

This appendix intends to show the weak limit of the localized objective function after neglecting the small terms as done in Appendix C. We apply Theorem 2.11.22 of VW to prove all lemmas. All proofs include two parts: (i) the finite-dimensional limit distributions of the process are the same as specified in the lemmas; (ii) the process is stochastically equicontinuous or tight. For the second part, we need to check the condition (2.11.21) and the uniform-entropy condition in Theorem 2.11.22 of VW.

Lemma 12 *Under Assumption I(1),*

$$\left(W_n\left(u\right), D_n\left(v\right)\right) \rightsquigarrow \left(W\left(u\right), D\left(v\right)\right),$$

where

$$W\left(u\right) = u_1' W_1 + u_2' W_2 = u' W$$

with W defined in (9), and $D\left(v\right)$ defined in the main text. Furthermore, W_1, W_2 and $D\left(v\right)$ are independent of each other.

Proof. Because the proof is almost the same as that of Theorem 1 in Yu (2014), we omit the details. ■

Lemma 13 *Under Assumption I(1)′,*

$$\left(W_n\left(u\right), C_n\left(v\right)\right) \rightsquigarrow \left(W\left(u\right), C\left(v\right)\right),$$

where $W\left(u\right) = u'W$, and $C\left(v\right)$ defined in the main text. Furthermore, W_1, W_2 and $C\left(v\right)$ are independent of each other.

Proof. The $W_n(u)$ part is standard, so we concentrate on the $C_n\left(v\right)$ part. From the proof of Lemma 9, we need only further concentrate on the random part of $C_n\left(v\right)$.

Recall that the random part of $C_n\left(v\right)$ is

$$\mathbb{G}_n\left(T_{3n}\left(v\right) + T_{5n}\left(v\right)\right)$$

where

$$\begin{aligned} T_{3n}\left(v\right) &= \sqrt{n}\delta_n' \mathbf{x}_i \varepsilon_{1i} 1(\gamma_0 + \frac{v}{a_n} < q \leq \gamma_0), \\ T_{5n}\left(v\right) &= -\sqrt{n}\delta_n' \mathbf{x}_i \varepsilon_{2i} 1\left(\gamma_0 < q \leq \gamma_0 + \frac{v}{a_n}\right). \end{aligned}$$

First, for $v_1 < 0$ and $v_2 > 0$, define

$$S_{1i} = \delta_n' \mathbf{x}_i \varepsilon_{1i} \Delta_i^-\left(v_1\right), S_{2i} = \delta_n' \mathbf{x}_i \varepsilon_{2i} \Delta_i^+\left(v_2\right), S_{3i} = \frac{1}{\sqrt{n}}\left(e_{1i}\mathbf{x}_i' 1\left(q_i \leq \gamma_0\right), e_{2i}\mathbf{x}_i' 1\left(q_i > \gamma_0\right)\right)' =: \frac{1}{\sqrt{n}} s_{3i},$$

where $\Delta_i^-\left(v_1\right) = 1\left(\gamma_0 + \frac{v_1}{a_n} < q_i \leq \gamma_0\right)$, $\Delta_i^+\left(v_2\right) = 1\left(\gamma_0 < q_i \leq \gamma_0 + \frac{v_2}{a_n}\right)$, and S_{3i} is the asymptotic random component in $\widehat{\beta}$. By Taylor expansion around $t := (t_1, t_2, t_3')' = \mathbf{0}$,

$$\begin{aligned} & \mathbb{E}\left[\exp\left\{\sqrt{-1}\left[t_1 S_{1i} + t_2 S_{2i} + t_3' S_{3i}\right]\right\}\right] \\ = \; & 1 - \frac{1}{2} t_1^2 \delta_n' \mathbb{E}\left[\mathbf{x}_i \mathbf{x}_i' \varepsilon_{1i}^2 \Delta_i^-\left(v_1\right)\right] \delta_n - \frac{1}{2} t_2^2 \delta_n' \mathbb{E}\left[\mathbf{x}_i \mathbf{x}_i' \varepsilon_{2i}^2 \Delta_i^+\left(v_2\right)\right] \delta_n - \frac{1}{n}\frac{1}{2} t_3' \mathbb{E}\left[s_{3i} s_{3i}'\right] t_3' + o\left(\frac{1}{n}\right) \\ = \; & 1 - \frac{t_1^2}{2n} c' \mathbb{E}\left[\mathbf{x}_i \mathbf{x}_i' \varepsilon_{1i}^2 | q_i = \gamma_0 -\right] c f_0 \left|v_1\right| - \frac{t_2^2}{2n} c' \mathbb{E}\left[\mathbf{x}_i \mathbf{x}_i' \varepsilon_{2i}^2 | q_i = \gamma_0 +\right] c f_0 v_2 - \frac{1}{2n} t_3' \mathbb{E}\left[s_{3i} s_{3i}'\right] t_3' + o\left(\frac{1}{n}\right), \end{aligned}$$

where the cross term is $O\left(n^{-1/2}\left\|\delta_n\right\| a_n^{-1}\right) = o\left(1/n\right)$, so that

$$\begin{aligned} & \mathbb{E}\left[\exp\left\{\sqrt{-1}\left[t_1 \sum_{i=1}^n S_{1i} + t_2 \sum_{i=1}^n S_{2i} + t_3' \sum_{i=1}^n S_{3i}\right]\right\}\right] = \prod_{i=1}^n \mathbb{E}\left[\exp\left\{\sqrt{-1}\left[t_1 S_{1i} + t_2 S_{2i} + t_3' S_{3i}\right]\right\}\right] \\ = \; & \exp\left\{-\tfrac{t_1^2}{2} c' \mathbb{E}\left[\mathbf{x}_i \mathbf{x}_i' \varepsilon_{1i}^2 | q_i = \gamma_0 -\right] c f_0 \left|v_1\right| - \tfrac{t_2^2}{2} c' \mathbb{E}\left[\mathbf{x}_i \mathbf{x}_i' \varepsilon_{2i}^2 | q_i = \gamma_0 +\right] c f_0 v_2 - \tfrac{1}{2} t_3' \mathbb{E}\left[s_{3i} s_{3i}'\right] t_3'\right\} + o(1). \end{aligned}$$

As a result, $\delta_n' \sum_{i=1}^n \mathbf{x}_i \varepsilon_{1i} \Delta_i^- (v_1)$, $\delta_n' \sum_{i=1}^n \mathbf{x}_i \varepsilon_{2i} \Delta_i^+ (v_2)$ and $\widehat{\beta}$ are asymptotically independent and the have finite-dimensional limit distributions as specified in the lemma. Second, the uniform-entropy condition holds since $\{T_{3n}(v) : -\infty < -K \le v \le 0\}$ is VC-subgraph for each n and the VC-index bounded by some constant independent of n (see, e.g., Example 2.11.24 of VW), so it remains to show condition (2.11.21):

$$\text{(i) } P^* F_n^2 = O(1), \text{ (ii) } P^* F_n^2 1\left(F_n > \eta\sqrt{n}\right) \to 0, \forall \eta > 0,$$

and

$$\text{(iii)} \sup_{|v_1 - v_2| < \eta_n} P\left(T_{3n}(v_1) - T_{3n}(v_2)\right)^2 \to 0, \forall \eta_n \downarrow 0,$$

where P^* is the outer probability, F_n is the envelope function of $\{T_{3n}(v) : -\infty < -K \le v \le 0\}$ and can be taken as

$$F_n = \sqrt{n} \|\delta_n\| \|\mathbf{x}\| |\varepsilon_1| 1(\gamma_0 - \frac{K}{a_n} < q \le \gamma_0).$$

(i)

$$\begin{aligned} P^* F_n^2 &\le n \int_{\gamma_0 - K/a_n}^{\gamma_0} \mathbb{E}\left[\|\delta_n\|^2 \|\mathbf{x}\|^2 \varepsilon_1^2 | q\right] f(q)\, dq \\ &\le C \frac{n \|\delta_n\|^2}{a_n} \sup_{\gamma \in \mathcal{N}} \left\{ f(\gamma) \mathbb{E}\left[\|\mathbf{x}\|^2 \varepsilon_1^2 | q = \gamma\right]\right\} = O(1) \end{aligned}$$

by Assumptions (iv)(b) and (vii). (ii)

$$\begin{aligned} & P^* F_n^2 1\left(F_n > \eta\sqrt{n}\right) \\ \le\ & n\mathbb{E}\left[\|\delta_n\|^2 \|\mathbf{x}\|^2 \varepsilon_1^2 1\left(\gamma_0 - K/a_n < q \le \gamma_0\right) 1\left(\|\mathbf{x}\| |\varepsilon_1| > \frac{\eta}{\|\delta_n\|}\right)\right] \\ \le\ & Cn \|\delta_n\|^2 \mathbb{E}\left[\left(\|\mathbf{x}\| |\varepsilon_1|\right)^{2+\epsilon} 1\left(\gamma_0 - K/a_n < q \le \gamma_0\right)\right] / \left(\frac{\eta}{\|\delta_n\|}\right)^{\epsilon} \\ \le\ & C \frac{n \|\delta_n\|^2}{a_n} \sup_{\gamma \in \mathcal{N}} \left\{ f(\gamma) \mathbb{E}\left[\left(\|\mathbf{x}\| |\varepsilon_1|\right)^{2+\epsilon} | q = \gamma\right]\right\} / \left(\frac{\eta}{\|\delta_n\|}\right)^{\epsilon} \\ \to\ & 0, \end{aligned}$$

where the convergence is from Assumptions (iv)(c) and (vii). (iii) Suppose $v_1 < v_2 < 0$,

$$\begin{aligned} & \sup_{|v_1 - v_2| < \eta_n} \mathbb{E}\left(T_{3n}(v_1) - T_{3n}(v_2)\right)^2 \\ =\ & \sup_{|v_1 - v_2| < \eta_n} n\mathbb{E}\left[\|\delta_n\|^2 \|\mathbf{x}\|^2 \varepsilon_1^2 1\left(\gamma_0 + \frac{v_1}{a_n} < q \le \gamma_0 + \frac{v_2}{a_n}\right)\right] \\ \le\ & C \sup_{|v_1 - v_2| < \eta_n} \left\{ |v_1 - v_2| \frac{n \|\delta_n\|^2}{a_n} \sup_{\gamma \in \mathcal{N}} \left\{ f(\gamma) \mathbb{E}\left[\|\mathbf{x}\|^2 \varepsilon_1^2 | q = \gamma\right]\right\}\right\} \\ \to\ & 0, \forall \eta_n \downarrow 0. \end{aligned}$$

■

Lemma 14 *Under Assumption I(α), $1 < \alpha \le 2$,*

$$\Xi_n(v) \rightsquigarrow \Xi(v),$$

where

$$\Xi(v)=\begin{cases}\sqrt{f_0\omega_0^-}B_1(-v)=:\sqrt{\varpi_-}B_1(-v), & \text{if } v\le 0,\\ \sqrt{f_0\omega_0^+}B_2(v)=:\sqrt{\varpi_+}B_1(-v), & \text{if } v>0.\end{cases}$$

Proof. Recall that

$$\Xi_n(v)=\mathbb{G}_n\left(T_{3n}(v)+T_{5n}(v)\right),$$

where

$$\begin{aligned}T_{3n}(v) &= \sqrt{\rho_n}\delta_0'\mathbf{x}\left(y-\mathbf{x}'\overline{\beta}_0\right)1(\gamma_0+\frac{v}{\rho_n}<q\le\gamma_0),\\ T_{5n}(v) &= -\sqrt{\rho_n}\delta_0'\mathbf{x}\left(y-\mathbf{x}'\overline{\beta}_0\right)1\left(\gamma_0<q\le\gamma_0+\frac{v}{\rho_n}\right).\end{aligned}$$

Take $T_{3n}(v)$ as an example for analysis. First, check the covariance kernel of $T_{3n}(v)$. For $v_1,v_2<0$,

$$\begin{aligned}&\mathbb{E}\left[\sqrt{\rho_n}\left(\left(y-\mathbf{x}'\overline{\beta}_0\right)\mathbf{x}'\delta_0 1\left(\gamma_0+\tfrac{v_1}{\rho_n}<q\le\gamma_0\right)-\Lambda_-\left(\tfrac{v_1}{\rho_n}\right)\right)\sqrt{\rho_n}\left(\left(y-\mathbf{x}'\overline{\beta}_0\right)\mathbf{x}'\delta_0 1\left(\gamma_0+\tfrac{v_2}{\rho_n}<q\le\gamma_0\right)-\Lambda_-\left(\tfrac{v_2}{\rho_n}\right)\right)\right]\\ &=\rho_n\mathbb{E}\left[\left(y-\mathbf{x}'\overline{\beta}_0\right)^2(\mathbf{x}'\delta_0)^2 1\left(\gamma_0+\tfrac{v_1\vee v_2}{\rho_n}<q\le\gamma_0\right)\right]-\sqrt{\rho_n}\Lambda_-\left(\tfrac{v_1}{\rho_n}\right)\sqrt{\rho_n}\Lambda_-\left(\tfrac{v_2}{\rho_n}\right)\\ &\to|v_1\vee v_2|\,f_0\mathbb{E}\left[\left(y-\mathbf{x}'\overline{\beta}_0\right)^2(\mathbf{x}'\delta_0)^2|q=\gamma_0-\right],\end{aligned}$$

which is the covariance kernel of $\Xi(v)$ as $v\le 0$. Second, the uniform-entropy condition holds by similar arguments as in the proof of Lemma 13, and we now show condition (2.11.21), where F_n, the envelope function of $\{T_{3n}(v):-\infty<-K\le v\le 0\}$, can be taken as

$$F_n=\sqrt{\rho_n}\left|\left(y-\mathbf{x}'\overline{\beta}_0\right)\mathbf{x}'\delta_0\right|1\left(\gamma_0-K/\rho_n<q\le\gamma_0\right).$$

(i) $P^*F_n^2\le K\frac{1}{K/\rho_n}\int_{\gamma_0-K/\rho_n}^{\gamma_0}\mathbb{E}\left[\left(\varepsilon_1+m_1(x,q)-\mathbf{x}'\overline{\beta}_0\right)^2\|\mathbf{x}\|^2\|\delta_0\|^2|q\right]f(q)\,dq=O(1)$ by Assumption (iv)(b) and (vii). (ii)

$$\begin{aligned}&P^*F_n^2 1\left(F_n>\eta\sqrt{n}\right)\\ \le\;&\rho_n\mathbb{E}\left[\left|\left(y-\mathbf{x}'\overline{\beta}_0\right)\mathbf{x}'\delta_0\right|^2 1\left(\gamma_0-K/\rho_n<q\le\gamma_0\right)1\left(\left|\left(y-\mathbf{x}'\overline{\beta}_0\right)\mathbf{x}'\delta_0\right|>\eta\sqrt{n/\rho_n}\right)\right]\\ \le\;&\rho_n\mathbb{E}\left[\left|\left(y-\mathbf{x}'\overline{\beta}_0\right)\mathbf{x}'\delta_0\right|^{2+\epsilon}1\left(\gamma_0-K/\rho_n<q\le\gamma_0\right)\right]/\left(\eta\sqrt{n/\rho_n}\right)^{\epsilon}\\ \le\;&K\rho_n/\rho_n\sup_{\gamma\in\mathcal{N}}\left\{f(\gamma)\mathbb{E}\left[\left|\left(y-\mathbf{x}'\overline{\beta}_0\right)\mathbf{x}'\delta_0\right|^{2+\epsilon}|q=\gamma\right]\right\}/\left(\eta\sqrt{n/\rho_n}\right)^{\epsilon}\\ \to\;&0,\end{aligned}$$

where the convergence is from Assumptions (iv)(c) and (vii). (iii) Suppose $v_1<v_2<0$,

$$\begin{aligned}&\sup_{|v_1-v_2|<\eta_n}\mathbb{E}\left(T_{3n}(v_1)-T_{3n}(v_2)\right)^2\\ =\;&\sup_{|v_1-v_2|<\eta_n}\mathbb{E}\left[\rho_n\left|\left(y-\mathbf{x}'\overline{\beta}_0\right)\mathbf{x}'\delta_0\right|^2 1\left(\gamma_0+\frac{v_1}{\rho_n}<q\le\gamma_0+\frac{v_2}{\rho_n}\right)\right]\\ \le\;&\sup_{|v_1-v_2|<\eta_n}\left\{|v_1-v_2|\,\rho_n/\rho_n\sup_{\gamma\in\mathcal{N}}\left\{f(\gamma)\mathbb{E}\left[\left|\left(y-\mathbf{x}'\overline{\beta}_0\right)\mathbf{x}'\delta_0\right|^2|q=\gamma\right]\right\}\right\}\\ \to\;&0,\ \forall\eta_n\downarrow 0.\end{aligned}$$

■

Lemma 15 *Under Assumption II(α), $3 \le \alpha \le 4$,*

$$\Xi_n(v) \rightsquigarrow \Xi(v),$$

where

$$\Xi(v) = \begin{cases} \sqrt{\frac{f_0 \delta_{q0}^2}{3}\omega_0^-} B_1(-v^3) =: \sqrt{\varpi_-} B_1(-v^3) & \text{if } v \le 0, \\ \sqrt{\frac{f_0 \delta_{q0}^2}{3}\omega_0^+} B_2(v^3) =: \sqrt{\varpi_+} B_1(-v^3), & \text{if } v > 0. \end{cases}$$

Proof. Only when $\alpha = 3$ both $W_n(u)$ and $\Xi_n(v)$ appear. We can show as in Lemma 13 that they are asymptotically independent. Also, the weak limit of $W_n(u)$ is easy to derive, so we concentrate on $\Xi_n(v)$ in this lemma.

Recall that

$$\Xi_n(v) = \mathbb{G}_n\left(T_{3n}(v) + T_{5n}(v)\right)$$

where

$$\begin{aligned} T_{3n}(v) &= \sqrt{\rho_n^3}\left(y - \mathbf{x}'\overline{\beta}_0\right) q\delta_{q0} 1(\gamma_0 + \frac{v}{\rho_n} < q \le \gamma_0), \\ T_{5n}(v) &= -\sqrt{\rho_n^3}\delta_0'\mathbf{x}\left(y - \mathbf{x}'\overline{\beta}_0\right) q\delta_{q0} 1\left(\gamma_0 < q \le \gamma_0 + \frac{v}{\rho_n}\right). \end{aligned}$$

Take $T_{3n}(v)$ as an example for analysis. First, check the covariance kernel of $T_{3n}(v)$. For $v_1, v_2 < 0$,

$$\begin{aligned} &\mathbb{E}\left[\sqrt{\rho_n^3}\left(\left(y_i - \mathbf{x}'\overline{\beta}_0\right) q\delta_{q0} 1\left(\gamma_0 + \tfrac{v_1}{\rho_n} < q \le \gamma_0\right) - \Lambda_-\left(\tfrac{v_1}{\rho_n}\right)\right)\sqrt{\rho_n^3}\left(\left(y_i - \mathbf{x}'\overline{\beta}_0\right) q\delta_{q0} 1\left(\gamma_0 + \tfrac{v_2}{\rho_n} < q \le \gamma_0\right) - \Lambda_+\left(\tfrac{v_2}{\rho_n}\right)\right)\right] \\ &= \rho_n^3 \mathbb{E}\left[\left(y_i - \mathbf{x}'\overline{\beta}_0\right)^2 (q\delta_{q0})^2 1\left(\gamma_0 + \tfrac{v_1 \vee v_2}{\rho_n} < q \le \gamma_0\right)\right] - \sqrt{\rho_n^3}\Lambda_-\left(\tfrac{v_1}{\rho_n}\right)\sqrt{\rho_n^3}\Lambda_-\left(\tfrac{v_2}{\rho_n}\right) \\ &\to |v_1 \vee v_2|^3 \tfrac{\delta_{q0}^2}{3} f_0 \mathbb{E}\left[\left(y_i - \mathbf{x}'\overline{\beta}_0\right)^2 | q = \gamma_0 -\right], \end{aligned}$$

which is the covariance kernel of $\Xi(v)$ as $v \le 0$. As the the tightness of $\Xi_n(v)$, the proof is almost identical as that of Lemma 14, so omitted. Note that the proof here can be applied to $2.5 < \alpha < 3$ with ϱ_n replacing ρ_n. ■

Lemma 16 *Under Assumption II(2.5),*

$$\begin{pmatrix} \Xi_{1n}(v) \\ \Xi_{2n}(v) \end{pmatrix} \rightsquigarrow \begin{pmatrix} \Xi_1(v) \\ \Xi_2(v) \end{pmatrix} =: \begin{pmatrix} \Xi_1^-(v) 1(v \le 0) + \Xi_1^+(v) 1(v > 0) \\ \Xi_2^-(v) 1(v \le 0) + \Xi_2^+(v) 1(v > 0) \end{pmatrix},$$

where

$$\begin{aligned} \Xi_{1n}(v) &= \begin{cases} n^{1/4}\mathbb{G}_n\left(\left(y - \mathbf{x}'\overline{\beta}_0\right)\mathbf{x} 1(\gamma_0 + \frac{v}{n^{1/2}} < q \le \gamma_0)\right) & \text{if } v \le 0, \\ n^{1/4}\mathbb{G}_n\left(\left(y - \mathbf{x}'\overline{\beta}_0\right)\mathbf{x} 1(\gamma_0 < q \le \gamma_0 + \frac{v}{n^{1/2}})\right) & \text{if } v > 0, \end{cases} \\ \Xi_{2n}(v) &= \begin{cases} n^{3/4}\mathbb{G}_n\left(\left(y - \mathbf{x}'\overline{\beta}_0\right) q\delta_{q0} 1(\gamma_0 + \frac{v}{n^{1/2}} < q \le \gamma_0)\right), & \text{if } v \le 0, \\ -n^{3/4}\mathbb{G}_n\left(\left(y - \mathbf{x}'\overline{\beta}_0\right) q\delta_{q0} 1(\gamma_0 < q \le \gamma_0 + \frac{v}{n^{1/2}})\right), & \text{if } v > 0, \end{cases} \end{aligned}$$

and $\begin{pmatrix} \Xi_1(v) \\ \Xi_2(v) \end{pmatrix}$ *is a $(d+2)$-dimensional Brownian motion with the covariance kernel on $(-\infty, 0]$ is*

$$f_0 \begin{pmatrix} \mathbb{E}\left[\left(y - \mathbf{x}'\overline{\beta}_0\right)^2 \mathbf{x}\mathbf{x}' | q = \gamma_0 -\right] |v_1 \vee v_2| & -\frac{\delta_{q0}}{2}\mathbb{E}\left[\left(y - \mathbf{x}'\overline{\beta}_0\right)^2 \mathbf{x} | q = \gamma_0 -\right] |v_1 \vee v_2|^2 \\ -\frac{\delta_{q0}}{2}\mathbb{E}\left[\left(y - \mathbf{x}'\overline{\beta}_0\right)^2 \mathbf{x}' | q = \gamma_0 -\right] |v_1 \vee v_2|^2 & \frac{\delta_{q0}^2}{3}\mathbb{E}\left[\left(y - \mathbf{x}'\overline{\beta}_0\right)^2 | q = \gamma_0 -\right] |v_1 \vee v_2|^3 \end{pmatrix},$$

and on $[0,\infty)$ *is*

$$f_0\left(\begin{array}{cc}\mathbb{E}\left[\left(y-\mathbf{x}'\overline{\beta}_0\right)^2\mathbf{x}\mathbf{x}'|q=\gamma_0+\right](v_1\wedge v_2) & -\frac{\delta_{q0}}{2}\mathbb{E}\left[\left(y-\mathbf{x}'\overline{\beta}_0\right)^2\mathbf{x}|q=\gamma_0+\right](v_1\wedge v_2)^2\\ -\frac{\delta_{q0}}{2}\mathbb{E}\left[\left(y-\mathbf{x}'\overline{\beta}_0\right)^2\mathbf{x}'|q=\gamma_0+\right](v_1\wedge v_2)^2 & \frac{\delta_{q0}^2}{3}\mathbb{E}\left[\left(y-\mathbf{x}'\overline{\beta}_0\right)^2|q=\gamma_0+\right](v_1\wedge v_2)^3\end{array}\right).$$

Proof. The proof is almost the same as that of Lemma 15, so we only mention the differences. First, $\Xi_{2n}(v)$ is the same as $\Xi_n(v)$ except $\rho_n=n^{1/2}$, so $\Xi_{2n}(v)\rightsquigarrow\Xi_2(v)$ with $\Xi_2(v)$ being the same as $\Xi(v)$ in Lemma 15. For $\Xi_{1n}(v)$, its covariance kernel for $v_1,v_2<0$ is

$$\begin{aligned}&\mathbb{E}\left[\begin{array}{c}n^{1/4}\left(\left(y-\mathbf{x}'\overline{\beta}_0\right)\mathbf{x}1(\gamma_0+\frac{v_1}{n^{1/2}}<q\le\gamma_0)-\mathbb{E}\left[\left(y-\mathbf{x}'\overline{\beta}_0\right)\mathbf{x}1(\gamma_0+\frac{v_1}{n^{1/2}}<q\le\gamma_0)\right]\right)\\ n^{1/4}\left(\left(y-\mathbf{x}'\overline{\beta}_0\right)\mathbf{x}'1(\gamma_0+\frac{v_2}{n^{1/2}}<q\le\gamma_0)-\mathbb{E}\left[\left(y-\mathbf{x}'\overline{\beta}_0\right)\mathbf{x}'1(\gamma_0+\frac{v_2}{n^{1/2}}<q\le\gamma_0)\right]\right)\end{array}\right]\\ &=n^{1/2}\mathbb{E}\left[\left(y-\mathbf{x}'\overline{\beta}_0\right)^2\mathbf{x}\mathbf{x}'1\left(\gamma_0+\tfrac{v_1\vee v_2}{n^{1/2}}<q\le\gamma_0\right)\right]\\ &\quad-n^{1/4}\mathbb{E}\left[\left(y-\mathbf{x}'\overline{\beta}_0\right)\mathbf{x}1(\gamma_0+\tfrac{v_1}{n^{1/2}}<q\le\gamma_0)\right]n^{1/4}\mathbb{E}\left[\left(y-\mathbf{x}'\overline{\beta}_0\right)\mathbf{x}'1(\gamma_0+\tfrac{v_2}{n^{1/2}}<q\le\gamma_0)\right]\\ &\to|v_1\vee v_2|\,f_0\mathbb{E}\left[\left(y-\mathbf{x}'\overline{\beta}_0\right)^2\mathbf{x}\mathbf{x}'|q=\gamma_0-\right],\end{aligned}$$

which is the covariance kernel of

$$\Xi_1(v)=\begin{cases}f_0^{1/2}\mathbb{E}\left[\left(y-\mathbf{x}'\overline{\beta}_0\right)^2\mathbf{x}\mathbf{x}'|q=\gamma_0-\right]^{1/2}\mathbf{B}_1(-v)=:\sqrt{\Omega_-}\mathbf{B}_1(-v) & \text{if } v\le 0,\\ f_0^{1/2}\mathbb{E}\left[\left(y-\mathbf{x}'\overline{\beta}_0\right)^2\mathbf{x}\mathbf{x}'|q=\gamma_0+\right]^{1/2}\mathbf{B}_2(v)=:\sqrt{\Omega_+}\mathbf{B}_2(v) & \text{if } v>0.\end{cases}$$

as $v\le 0$, where $\mathbf{B}_1(v)$ and $\mathbf{B}_2(v)$ are two independent standard $(d+1)$-dimensional Brownian motions on $[0,\infty)$. For $v_1,v_2<0$, the covariance between $\Xi_{1n}(v_1)$ and $\Xi_{2n}(v_2)$ is

$$\begin{aligned}&\mathbb{E}\left[\begin{array}{c}n^{1/4}\left(\left(y-\mathbf{x}'\overline{\beta}_0\right)\mathbf{x}1(\gamma_0+\frac{v_1}{n^{1/2}}<q\le\gamma_0)-\mathbb{E}\left[\left(y-\mathbf{x}'\overline{\beta}_0\right)\mathbf{x}1(\gamma_0+\frac{v_1}{n^{1/2}}<q\le\gamma_0)\right]\right)\\ n^{3/4}\left(\left(y-\mathbf{x}'\overline{\beta}_0\right)q\delta_{q0}1(\gamma_0+\frac{v_2}{n^{1/2}}<q\le\gamma_0)-\mathbb{E}\left[\left(y-\mathbf{x}'\overline{\beta}_0\right)q\delta_{q0}1(\gamma_0+\frac{v_2}{n^{1/2}}<q\le\gamma_0)\right]\right)\end{array}\right]\\ &=n\mathbb{E}\left[\left(y-\mathbf{x}'\overline{\beta}_0\right)^2\mathbf{x}q1\left(\gamma_0+\tfrac{v_1\vee v_2}{n^{1/2}}<q\le\gamma_0\right)\right]\delta_{q0}\\ &\quad-n^{1/4}\mathbb{E}\left[\left(y-\mathbf{x}'\overline{\beta}_0\right)\mathbf{x}1(\gamma_0+\tfrac{v_1}{n^{1/2}}<q\le\gamma_0)\right]n^{3/4}\mathbb{E}\left[\left(y-\mathbf{x}'\overline{\beta}_0\right)q\delta_{q0}1(\gamma_0+\tfrac{v_2}{n^{1/2}}<q\le\gamma_0)\right]\\ &\to-|v_1\vee v_2|^2\,\tfrac{f_0}{2}\mathbb{E}\left[\left(y-\mathbf{x}'\overline{\beta}_0\right)^2\mathbf{x}|q=\gamma_0-\right]\delta_{q0}=:|v_1\vee v_2|^2\,\Upsilon_-.\end{aligned}$$

In summary, $\begin{pmatrix}\Xi_{1n}(v)\\ \Xi_{2n}(v)\end{pmatrix}\rightsquigarrow\begin{pmatrix}\Xi_1(v)\\ \Xi_2(v)\end{pmatrix}$, where $\begin{pmatrix}\Xi_1(v)\\ \Xi_2(v)\end{pmatrix}$ is as stated in the lemma. ■

Additional References

Bhattacharya, P.K. and P.J. Brockwell, 1976, The Minimum of an Additive Process with Applications to Signal Estimation and Storage Theory, *Z. Wahrschein. Verw. Gebiete*, 37, 51-75.

Kim, J. and D. Pollard, 1990, Cube Root Asymptotics, *Annals of Statistics*, 18, 191-219.

Newey, W.K. and D.L. McFadden, 1994, Large Sample Estimation and Hypothesis Testing, *Handbook of Econometrics*, Vol. 4, R.F. Eagle and D.L. McFadden, eds., Elsevier Science B.V., Ch. 36, 2111-2245.

Van der Vaart, A.W. and J.A. Wellner, 1996, *Weak Convergence and Empirical Processes: With Applications to Statistics*, New York: Springer.