SUPPLEMENTAL APPENDIX FOR "ROBUSTNESS, INFINITESIMAL NEIGHBORHOODS, AND MOMENT RESTRICTIONS"

YUICHI KITAMURA, TAISUKE OTSU, AND KIRILL EVDOKIMOV

A.1. Proofs of Main Results

This Appendix presents the proofs of some of the results presented in the previous sections.

A.1.1. Proof of Lemma 2.1.

We first show the claim for $\alpha < \frac{1}{2}$, that is,

$$(1-\alpha) I_\alpha (P,Q) - \frac{1}{2} I_{\frac{1}{2}} (P,Q) \geq 0. \tag{A.1}$$

Let $H_\alpha (x) = \frac{1}{\alpha}(1-x^\alpha) - 2\left(1 - x^{\frac{1}{2}}\right), \quad 0 \leq x \leq \infty$, then the above inequality becomes

$$\int H_\alpha \left(\frac{p}{q}\right) q d\nu \geq 0. \tag{A.2}$$

Note

$$\frac{d}{dx} H_\alpha (x) = -x^{\alpha-1} + x^{-\frac{1}{2}} \begin{cases} > 0 & \text{if } x > 1 \\ = 0 & \text{if } x = 1 \\ < 0 & \text{if } x < 1. \end{cases}$$

The above holds for the case with $\alpha = 0$ as well, since $H_0(x) = -\log x - 2\left(1 - x^{\frac{1}{2}}\right)$. Moreover, $H_\alpha(1) = 0$. Therefore $H_\alpha(x) \geq 0$ for all $x \geq 0$, and the desired inequality (A.2) follows immediately. Next, we prove the case with $\alpha > \frac{1}{2}$, that is,

$$\alpha I_\alpha (P,Q) \geq \frac{1}{2} I_{\frac{1}{2}} (P,Q).$$

Let $\beta = 1 - \alpha < \frac{1}{2}$, then the above inequality becomes

$$(1-\beta) I_{1-\beta} (P,Q) \geq \frac{1}{2} I_{\frac{1}{2}} (P,Q). \tag{A.3}$$

By (A.1) and the symmetry of the Hellinger distance,

$$(1-\beta) I_\beta (Q,P) \geq \frac{1}{2} I_{\frac{1}{2}} (Q,P) = \frac{1}{2} I_{\frac{1}{2}} (P,Q).$$

But the equality $I_{1-\beta}(P,Q) = I_\beta(Q,P)$ holds for every $\beta \in \mathbb{R}$, and (A.3) follows.

Notation. Let C be a generic positive constant, $\|\cdot\|$ be the L_2-metric,

$$\begin{aligned}\theta_n &= \theta_0 + t/\sqrt{n}, \quad \bar{T}_{Q_n} = \bar{T}(Q_n), \quad \bar{T}_{P_n} = \bar{T}(P_n), \\ \bar{P}_{\theta,Q} &= \arg\min_{P \in \bar{\mathcal{P}}_\theta} H(P,Q), \quad R_n(Q,\theta,\gamma) = -\int \frac{1}{(1+\gamma' g_n(x,\theta))} dQ, \\ g_n(x,\theta) &= g(x,\theta)\mathbb{I}\{x \in \mathcal{X}_n\}, \quad \Lambda_n = G'\Omega^{-1} g_n(x,\theta_0), \quad \Lambda = G'\Omega^{-1} g(x,\theta_0), \\ \psi_{n,Q_n} &= -2\left(\int \Lambda_n \Lambda_n' dQ_n\right)^{-1} \int \Lambda_n \left\{dQ_n^{1/2} - d\bar{P}_{\theta_0,Q_n}^{1/2}\right\} dQ_n^{1/2}.\end{aligned}$$

A.1.2. **Proof of Theorem 3.1.**

A.1.2.1. *Proof of (i).* Pick arbitrary $r > 0$ and $t \in \mathbb{R}^p$. Consider the following parametric submodel having the likelihood ratio

$$\frac{dP_{\theta_n,\zeta_n}}{dP_0} = \frac{1+\zeta_n' g_n(x,\theta_n)}{\int (1+\zeta_n' g_n(x,\theta_n))\, dP_0} = f(x,\theta_n,\zeta_n), \tag{A.4}$$

where

$$\zeta_n = -E_{P_0}\left[g(x,\theta_n) g_n(x,\theta_n)'\right]^{-1} E_{P_0}[g(x,\theta_n)].$$

Note that $P_{\theta_0,0} = P_0$, $P_{\theta_n,\zeta_n} \in \mathcal{P}_{\theta_n}$ (by the definition of ζ_n), and $\zeta_n = O(n^{-1/2})$ (by the proof of Lemma A.4 (i)). Also, since $\sup_{x\in\mathcal{X}} |\zeta_n' g_n(x,\theta_n)| = O(n^{-1/2} m_n) = o(1)$, the likelihood ratio $\frac{dP_{\theta_n,\zeta_n}}{dP_0}$ is well-defined for all n large enough. So, for this submodel the mapping T_a must satisfy (3.1).
We now evaluate the Hellinger distance between P_{θ_n,ζ_n} and P_0. An expansion around $\zeta_n = 0$ yields

$$H(P_{\theta_n,\zeta_n}, P_0) = \left\| \zeta_n' \left.\frac{\partial f(x,\theta_n,\zeta_n)^{1/2}}{\partial \zeta_n}\right|_{\zeta_n=0} dP_0^{1/2} + \frac{1}{2}\zeta_n' \left.\frac{\partial^2 f(x,\theta_n,\zeta_n)^{1/2}}{\partial\zeta_n \partial\zeta_n'}\right|_{\zeta_n=\dot{\zeta}_n} \zeta_n dP_0^{1/2} \right\|,$$

where $\dot{\zeta}_n$ is a point on the line joining ζ_n and 0, and

$$\left.\frac{\partial f(x,\theta_n,\zeta_n)^{1/2}}{\partial \zeta_n}\right|_{\zeta_n=0} = \frac{1}{2}\{g_n(x,\theta_n) - E_{P_0}[g_n(x,\theta_n)]\},$$

$$\begin{aligned}\frac{\partial^2 f(x,\theta_n,\zeta_n)^{1/2}}{\partial\zeta_n\partial\zeta_n'} &= -\frac{1}{4}\left(1+\zeta_n' g_n(x,\theta_n)\right)^{-3/2}\left(1+\zeta_n' E_{P_0}[g_n(x,\theta_n)]\right)^{-1/2} g_n(x,\theta_n) g_n(x,\theta_n)' \\ &\quad -\frac{1}{2}\left(1+\zeta_n' g_n(x,\theta_n)\right)^{-1/2}\left(1+\zeta_n' E_{P_0}[g_n(x,\theta_n)]\right)^{-3/2} g_n(x,\theta_n) E_{P_0}[g_n(x,\theta_n)]' \\ &\quad +\frac{3}{4}\left(1+\zeta_n' g_n(x,\theta_n)\right)^{1/2}\left(1+\zeta_n' E_{P_0}[g_n(x,\theta_n)]\right)^{-5/2} E_{P_0}[g_n(x,\theta_n)] E_{P_0}[g_n(x,\theta_n)]'.\end{aligned}$$

Thus, a lengthy but straightforward calculation combined with Lemma A.4, $\zeta_n = O\left(n^{-1/2}\right)$, and $\sup_{x\in\mathcal{X}} |\zeta_n' g_n(x,\theta_n)| = o(1)$ implies

$$\text{(A.5)} \qquad nH\left(P_{\theta_n,\zeta_n}, P_0\right)^2 = n\left\|\frac{1}{2}\zeta_n'\left(g_n(x,\theta_n) - E_{P_0}\left[g_n(x,\theta_n)\right]\right)dP_0^{1/2}\right\|^2 + o(1) \to \frac{1}{4}t'\Sigma^{-1}t.$$

Based on this limit, a lower bound of the maximum bias of T_a is obtained as (see, Rieder (1994, eq. (56) on p. 180))

$$\begin{aligned}
& \liminf_{n\to\infty} \sup_{Q\in B_H\left(P_0,r/\sqrt{n}\right)} n\left(\tau\circ T_a(Q) - \tau(\theta_0)\right)^2 \\
\geq\; & \liminf_{n\to\infty} \sup_{\left\{t\in\mathbb{R}^p : P_{\theta_n,\zeta_n}\in B_H\left(P_0,r/\sqrt{n}\right)\right\}} n\left(\tau\circ T_a\left(P_{\theta_n,\zeta_n}\right) - \tau(\theta_0)\right)^2 \\
\geq\; & \max_{\left\{t\in\mathbb{R}^p : \frac{1}{4}t'\Sigma t\leq r^2-\epsilon\right\}} \left(\left(\frac{\partial\tau(\theta_0)}{\partial\theta}\right)' t\right)^2 = 4\left(r^2-\epsilon\right)\left(\frac{\partial\tau(\theta_0)}{\partial\theta}\right)'\Sigma^{-1}\left(\frac{\partial\tau(\theta_0)}{\partial\theta}\right),
\end{aligned}$$

for each $\epsilon \in \left(0, r^2\right)$, where the first inequality follows from the set inclusion relationship, the second inequality follows from (3.1) and (A.5), and the equality follows from the Kuhn-Tucker theorem. Since ϵ can be arbitrarily small, we obtain the conclusion.

A.1.2.2. *Proof of (ii).* Pick arbitrary $r > 0$ and sequence $Q_n \in B_H\left(P_0, r/\sqrt{n}\right)$. We first show the Fisher consistency of $\bar{T}$. From Lemma A.2 (note: $P_{\theta_n,\zeta_n} \in B_H\left(P_0, r/\sqrt{n}\right)$ for all n large enough),

$$\begin{aligned}
\sqrt{n}\left(\bar{T}\left(P_{\theta_n,\zeta_n}\right) - \theta_0\right) &= -\sqrt{n}\Sigma^{-1}\int \Lambda_n dP_{\theta_n,\zeta_n} + o(1) \\
&= \Sigma^{-1}G'\Omega^{-1}\int \partial g\left(x,\dot{\theta}\right)/\partial\theta dP_{\theta_n,\zeta_n} t + o(1) \\
&\to t
\end{aligned}$$

for all n large enough, where $\dot{\theta}$ is a point on the line joining θ_n and θ_0, the second equality follows from $\int g(x,\theta_0)\mathbb{I}\{x\notin\mathcal{X}_n\}dP_{\theta_n,\zeta_n} = o\left(n^{-1/2}\right)$ (by a similar argument to (A.4)), $\int g(x,\theta_n)dP_{\theta_n,\zeta_n} = 0$ (by $P_{\theta_n,\zeta_n} \in \mathcal{P}_{\theta_n}$), and an expansion around $\theta_n = \theta_0$, and the convergence follows from the last statement of Lemma A.4 (i). Therefore, $\bar{T}$ is Fisher consistent.

We next show (3.1). An expansion of $\tau\circ\bar{T}_{Q_n}$ around $\bar{T}_{Q_n} = \theta_0$, Lemmas A.1 (ii) and A.2, and Assumption 3.1 (viii) imply

$$\begin{aligned}
& \sqrt{n}\left(\tau\circ\bar{T}_{Q_n} - \tau(\theta_0)\right) = -\sqrt{n}\left(\frac{\partial\tau(\theta_0)}{\partial\theta}\right)'\Sigma^{-1}\int \Lambda_n dQ_n + o(1) \\
=\; & -\sqrt{n}\nu_0'\int \Lambda_n\left\{dQ_n^{1/2} - dP_0^{1/2}\right\}dQ_n^{1/2} - \sqrt{n}\nu_0'\int \Lambda_n dP_0^{1/2}\left\{dQ_n^{1/2} - dP_0^{1/2}\right\} + o(1),
\end{aligned}$$

where we denote $\nu_0' = \left(\frac{\partial\tau(\theta_0)}{\partial\theta}\right)'\Sigma^{-1}$. From the triangle inequality,

$$\begin{aligned} & n\left(\tau\circ\bar{T}_{Q_n} - \tau(\theta_0)\right)^2 \\ \leq\ & n\left\{\begin{array}{c} \left|\nu_0'\int\Lambda_n\left\{dQ_n^{1/2} - dP_0^{1/2}\right\}dQ_n^{1/2}\right|^2 + \left|\nu_0'\int\Lambda_n\left\{dQ_n^{1/2} - dP_0^{1/2}\right\}dP_0^{1/2}\right|^2 \\ +2\left|\nu_0'\int\Lambda_n\left\{dQ_n^{1/2} - dP_0^{1/2}\right\}dQ_n^{1/2}\right|\left|\nu_0'\int\Lambda_n\left\{dQ_n^{1/2} - dP_0^{1/2}\right\}dP_0^{1/2}\right| \end{array}\right\} + o(1) \\ =\ & n\left\{A_1 + A_2 + 2A_3\right\} + o(1). \end{aligned}$$

For A_1, observe that

$$A_1 \leq \left|\nu_0'\int\Lambda_n\Lambda_n' dQ_n\nu_0\right|\left|\int\left\{dQ_n^{1/2} - dP_0^{1/2}\right\}^2\right| \leq B^*\frac{r^2}{n} + o\left(n^{-1}\right),$$

where the first inequality follows from the Cauchy-Schwarz inequality, and the second inequality follows from Lemma A.5 (i) and $Q_n \in B_H(P_0, r/\sqrt{n})$. Similarly, we have $A_2 \leq B^*\frac{r^2}{n} + o\left(n^{-1}\right)$ and $A_3 \leq B^*\frac{r^2}{n} + o\left(n^{-1}\right)$. Combining these terms,

$$\limsup_{n\to\infty} n\left(\tau\circ\bar{T}_{Q_n} - \tau(\theta_0)\right)^2 \leq 4r^2B^*, \tag{A.6}$$

for any sequence $Q_n \in B_H(P_0, r/\sqrt{n})$ and $r > 0$. Pick any $r > 0$. Since the supremum $\sup_{Q\in B_H\left(P_0,\frac{r}{\sqrt{n}}\right)} n\left(\tau\circ\bar{T}(Q) - \tau(\theta_0)\right)^2$ is finite for all n large enough (from Lemma A.1 (i)), there exists a sequence $Q_n^* \in B_H(P_0, r/\sqrt{n})$ such that

$$\limsup_{n\to\infty} n\left(\tau\circ\bar{T}_{Q_n^*} - \tau(\theta_0)\right)^2 = \limsup_{n\to\infty}\sup_{Q\in B_H\left(P_0,\frac{r}{\sqrt{n}}\right)} n\left(\tau\circ\bar{T}(Q) - \tau(\theta_0)\right)^2.$$

Therefore, the conclusion follows by (A.6).

A.1.3. **Proof of Theorem 3.2.**

A.1.3.1. *Proof of (i).* Pick arbitrary $\epsilon \in \left(0, r^2\right)$ and $r > 0$. Consider the parametric submodel P_{θ_n,ζ_n} defined in (A.4). The convolution theorem (Theorem 25.20 of van der Vaart (1998)) implies that for each $t \in \mathbb{R}^p$, there exists a probability measure M_0 which does not depend on t and satisfies

$$\sqrt{n}\left(\tau\circ T_a(P_n) - \tau\circ T_a\left(P_{\theta_n,\zeta_n}\right)\right) \xrightarrow{d} M_0 * N(0, B^*) \quad \text{under } P_{\theta_n,\zeta_n}. \tag{A.7}$$

Let

$$t^* = \arg\max_{\left\{t\in\mathbb{R}^p:\frac{1}{4}t'\Sigma t\leq r^2-\epsilon\right\}}\left(\left(\frac{\partial\tau(\theta_0)}{\partial\theta}\right)' t\right)^2 \quad \text{s.t.} \quad \left(\frac{\partial\tau(\theta_0)}{\partial\theta}\right)' t^*\int\xi dM_0 * N(0, B^*) \geq 0.$$

Since the integral $\int\xi dM_0 * N(0, B^*)$ does not depend on t, such t^* always exists. From $\frac{1}{4}t^{*\prime}\Sigma t^* \leq r^2 - \epsilon$ and (A.5), it holds that $P_{\theta_0+t^*/\sqrt{n},\zeta_n} \in B_H(P_0, r/\sqrt{n})$ for all n large enough. Also, note that

$E_{P_{\theta_n,\zeta_n}}\left[\sup_{\theta\in\Theta}|g(x,\theta)|^{\eta}\right]<\infty$ for all n large enough (by $\sup_{x\in\mathcal{X}}|\zeta_n' g_n(x,\theta_n)|=o(1)$ and Assumption 3.1 (v)). Thus, $P_{\theta_0+t^*/\sqrt{n},\zeta_n}\in\bar{B}_H(P_0,r/\sqrt{n})$ for all n large enough, and we have

$$
\begin{aligned}
& \lim_{b\to\infty}\liminf_{n\to\infty}\sup_{Q\in\bar{B}_H\left(P_0,r/\sqrt{n}\right)}\int b\wedge n\left(\tau\circ T_a(P_n)-\tau(\theta_0)\right)^2 dQ^{\otimes n} \\
\geq\; & \lim_{b\to\infty}\liminf_{n\to\infty}\int b\wedge n\left(\tau\circ T_a(P_n)-\tau(\theta_0)\right)^2 dP^{\otimes n}_{\theta_0+t^*/\sqrt{n},\zeta_n} \\
=\; & \lim_{b\to\infty}\liminf_{n\to\infty}\int b\wedge n\left(\xi+\left(\frac{\partial\tau(\theta_0)}{\partial\theta}\right)' t^*\right)^2 dM_0*N(0,B^*) \\
=\; & \int\xi^2 dM_0*N(0,B^*)+\left(\left(\frac{\partial\tau(\theta_0)}{\partial\theta}\right)' t^*\right)^2+2\left(\frac{\partial\tau(\theta_0)}{\partial\theta}\right)' t^*\int\xi dM_0*N(0,B^*) \\
\geq\; & \left\{1+4\left(r^2-\epsilon\right)\right\}B^*,
\end{aligned}
$$

where the first equality follows from the Fisher consistency of T_a, (A.9), and the continuous mapping theorem, the second equality follows from the monotone convergence theorem, and the second inequality follows from the definition of t^*. Since ϵ can be arbitrarily small, we obtain the conclusion.

A.1.3.2. *Proof of (ii).* Pick arbitrary $r>0$ and $b>0$. Applying the inequality $b\wedge(c_1+c_2)\leq b\wedge c_1+b\wedge c_2$ for any $c_1,c_2\geq 0$,

$$
\begin{aligned}
& \limsup_{n\to\infty}\sup_{Q\in\bar{B}_H\left(P_0,r/\sqrt{n}\right)}\int b\wedge n\left(\tau\circ T(P_n)-\tau(\theta_0)\right)^2 dQ^{\otimes n} \\
\leq\; & \limsup_{n\to\infty}\sup_{Q\in\bar{B}_H\left(P_0,r/\sqrt{n}\right)}\int b\wedge n\left(\tau\circ T(P_n)-\tau\circ\bar{T}(P_n)\right)^2 dQ^{\otimes n} \\
& +2\limsup_{n\to\infty}\sup_{Q\in\bar{B}_H\left(P_0,r/\sqrt{n}\right)}\int b\wedge\left\{n\left|\tau\circ T(P_n)-\tau\circ\bar{T}(P_n)\right|\left|\tau\circ\bar{T}(P_n)-\tau(\theta_0)\right|\right\}dQ^{\otimes n} \\
& +\limsup_{n\to\infty}\sup_{Q\in\bar{B}_H\left(P_0,r/\sqrt{n}\right)}\int b\wedge n\left(\tau\circ\bar{T}(P_n)-\tau(\theta_0)\right)^2 dQ^{\otimes n} \\
=\; & A_1+2A_2+A_3, \qquad\text{(A.8)}
\end{aligned}
$$

For A_1,

$$
\begin{aligned}
A_1 &\leq b\times\limsup_{n\to\infty}\sup_{Q\in\bar{B}_H\left(P_0,r/\sqrt{n}\right)}\int_{(x_1,\ldots,x_n)\notin\mathcal{X}_n^n} dQ^{\otimes n} \\
&\leq b\times\limsup_{n\to\infty}\sup_{Q\in\bar{B}_H\left(P_0,r/\sqrt{n}\right)}\sum_{i=1}^{n}\int_{x_i\notin\mathcal{X}_n} dQ \\
&\leq b\times\limsup_{n\to\infty}\sup_{Q\in\bar{B}_H\left(P_0,r/\sqrt{n}\right)} n m_n^{-\eta}E_Q\left[\sup_{\theta\in\Theta}|g(x,\theta)|^{\eta}\right]=0, \qquad\text{(A.9)}
\end{aligned}
$$

where the first inequality follows from $T(P_n) = \bar{T}(P_n)$ for all $(x_1, \ldots, x_n) \in \mathcal{X}_n^n$, the second inequality follows from a set inclusion relation, the third inequality follows from the Markov inequality, and the equality follows from Assumption 3.1 (vii) and $E_Q[\sup_{\theta\in\Theta} |g(x,\theta)|^\eta] < \infty$ for all $Q \in \bar{B}_H(P_0, r/\sqrt{n})$. Similarly, we have $A_2 = 0$.

We now consider A_3. Note that the mapping $f_{b,n}(Q) = \int b \wedge n\left(\tau \circ \bar{T}(P_n) - \tau(\theta_0)\right)^2 dQ^{\otimes n}$ is continuous in $Q \in B_H(P_0, r/\sqrt{n})$ under the Hellinger distance for each n, and the set $B_H(P_0, r/\sqrt{n})$ (not $\bar{B}_H(P_0, r/\sqrt{n})$) is compact under the Hellinger distance for each n. Thus, there exists $\tilde{Q}_{b,n} \in B_H(P_0, r/\sqrt{n})$ such that $\sup_{Q \in B_H(P_0, r/\sqrt{n})} f_n(Q) = f_n\left(\tilde{Q}_{b,n}\right)$ for each n. Then we have

$$
\begin{aligned}
A_3 &\leq \limsup_{n\to\infty} \sup_{Q\in B_H(P_0,r/\sqrt{n})} \int b \wedge n\left(\tau \circ \bar{T}(P_n) - \tau(\theta_0)\right)^2 dQ^{\otimes n} \\
&= \limsup_{n\to\infty} \int b \wedge n\left(\tau \circ \bar{T}(P_n) - \tau(\theta_0)\right)^2 d\tilde{Q}_{b,n}^{\otimes n} \\
&= \int b \wedge \left(\xi + \tilde{t}_b\right)^2 dN(0, B^*) \\
&\leq B^* + \tilde{t}_b^2 \\
&\leq \left(1 + 4r^2\right) B^*,
\end{aligned}
$$

where $\tilde{t}_b = \limsup_{n\to\infty} \sqrt{n}\left(\tau \circ \bar{T}\left(\tilde{Q}_{b,n}\right) - \tau(\theta_0)\right)$, the first inequality follows from $\bar{B}_H(P_0, r/\sqrt{n}) \subseteq B_H(P_0, r/\sqrt{n})$, the second equality follows from Lemma A.8 (with $Q_n = \tilde{Q}_{b,n}$) and the continuous mapping theorem, the second inequality follows from $b \wedge c \leq c$ and a direct calculation, and the last inequality follows from Theorem 3.1 (ii). Combining these results, the conclusion is obtained.

A.1.4. **Proof of Theorem 3.3.**

A.1.4.1. *Proof of (i).* Consider the parametric submodel P_{θ_n,ζ_n} defined in (A.4). Since ℓ is uniformly continuous on $\bar{\mathbb{R}}^p$ (by Assumption 3.2) and T_a is Fisher consistent,

$$
b \wedge \ell\left(\sqrt{n}\left\{S_n - \tau \circ T_a\left(P_{\theta_n,\zeta_n}\right)\right\}\right) - b \wedge \ell\left(\sqrt{n}\left\{S_n - \tau(\theta_0)\right\} - \left(\frac{\partial \tau(\theta_0)}{\partial \theta}\right)' t\right) \to 0,
$$

uniformly in t, $|t| < c$ and $\{S_n\}_{n\in\mathbb{N}}$ for each $c > 0$ and $b > 0$. Thus,

(A.10)
$$
\inf_{S_n\in\mathcal{S}} \sup_{|t|\leq c} \int b\wedge\ell\left(\sqrt{n}\left\{S_n - \tau \circ T_a\left(P_{\theta_n,\zeta_n}\right)\right\}\right) dP_{\theta_n,\zeta_n}^{\otimes n} = \inf_{R_n\in\mathcal{R}} \sup_{|t|\leq c} \int b\wedge\ell\left(R_n - \left(\frac{\partial \tau(\theta_0)}{\partial \theta}\right)' t\right) dP_{\theta_n,\zeta_n}^{\otimes n} + o(1),
$$

for each $c > 0$, where $R_n = \sqrt{n}\{S_n - \tau(\theta_0)\}$ is a standardized estimator and $\mathcal{R} = \{\sqrt{n}\{S_n - \tau(\theta_0)\} : S_n \in \mathcal{S}\}$. By expanding the log likelihood ratio $\log \frac{dP^{\otimes n}_{\theta_n,\zeta_n}}{dP^{\otimes n}_0}$ around $\zeta_n = 0$,

$$\begin{aligned}\log \frac{dP^{\otimes n}_{\theta_n,\zeta_n}}{dP^{\otimes n}_0} &= \zeta_n' \sum_{i=1}^n \{g_n(x_i,\theta_n) - E_{P_0}[g_n(x,\theta_n)]\} \\ &\quad -\frac{\zeta_n' \sum_{i=1}^n g_n(x_i,\theta_n) g_n(x_i,\theta_n) \zeta_n}{2\left(1+\dot{\zeta}_n' g_n(x_i,\theta_n)\right)^2} + \frac{n\zeta_n' E_{P_0}[g_n(x,\theta_n)] E_{P_0}[g_n(x,\theta_n)]' \zeta_n}{2\left(1+\ddot{\zeta}_n' \int g_n(x,\theta_n)\right)^2} \\ &= L_1 - L_2 + L_3.\end{aligned}$$

where $\dot{\zeta}_n$ and $\ddot{\zeta}_n$ are points on the line joining ζ_n and 0. For L_1, an expansion of $g_n(x,\theta_n)$ (in ζ_n) around $\theta_n = \theta_0$ combined with Lemma A.4 (i) implies that under P_0,

$$L_1 = -t'G'\Omega^{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^n \{g_n(x_i,\theta_n) - E_{P_0}[g_n(x,\theta_n)]\} + o_p(1).$$

Also, Lemma A.4 (i) and $\sup_{x\in\mathcal{X}} |\zeta_n' g_n(x,\theta_n)| = o(1)$ imply that under P_0,

$$L_2 \xrightarrow{p} \frac{1}{2}t'\Sigma t, \quad L_3 \to 0.$$

Therefore, in the terminology of Rieder (1994, Definition 2.2.9), the parametric model P_{θ_n,ζ_n} is asymptotically normal with the asymptotic sufficient statistic $-G'\Omega^{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^n \{g_n(x_i,\theta_n) - E_{P_0}[g_n(x,\theta_n)]\}$ and the asymptotic covariance matrix Σ. Note that this is essentially the LAN (local asymptotic normality) condition introduced by LeCam. If P_{θ_n,ζ_n} is asymptotically normal in this sense, we can directly apply the result of the minimax risk bound by Rieder (1994, Theorem 3.3.8 (a)), that is

$$\text{(A.11)} \qquad \lim_{b\to\infty}\lim_{c\to\infty}\liminf_{n\to\infty}\inf_{S_n\in\mathcal{S}}\sup_{|t|\le c}\int b\wedge\ell\left(R_n - \left(\frac{\partial\tau(\theta_0)}{\partial\theta}\right)' t\right) dP^{\otimes n}_{\theta_n,\zeta_n} \ge \int \ell dN(0,B^*)$$

(see also Theorem 1 in LeCam and Yang (1990)). From (A.10) and (A.11),

$$\lim_{b\to\infty}\lim_{c\to\infty}\liminf_{n\to\infty}\inf_{S_n\in\mathcal{S}}\sup_{|t|\le c}\int b\wedge\ell\left(\sqrt{n}\{S_n - \tau\circ T_a(P_{\theta_n,\zeta_n})\}\right) dP^{\otimes n}_{\theta_n,\zeta_n} \ge \int \ell dN(0,B^*).$$

Finally, since $E_{P_{\theta_n,\zeta_n}}[\sup_{\theta\in\Theta}|g(x,\theta)|^\eta] < \infty$ for all n large enough (by $\sup_{x\in\mathcal{X}}|\zeta_n' g_n(x,\theta_n)| = o(1)$ and Assumption 3.1 (v)), we have $P_{\theta_n,\zeta_n} \in \bar{B}_H(P_0, r/\sqrt{n})$ for all t satisfying $\frac{1}{4}t'\Sigma t \le r^2 - \epsilon$ with any $\epsilon \in (0, r^2)$ and all n large enough. Therefore, the set inclusion relation yields

$$\begin{aligned}&\lim_{b\to\infty}\lim_{r\to\infty}\liminf_{n\to\infty}\inf_{S_n\in\mathcal{S}}\sup_{Q\in\bar{B}_H(P_0,r/\sqrt{n})}\int b\wedge\ell\left(\sqrt{n}\{S_n - \tau\circ T_a(Q)\}\right) dQ^{\otimes n} \\ \ge\ &\lim_{b\to\infty}\lim_{c\to\infty}\liminf_{n\to\infty}\inf_{S_n\in\mathcal{S}}\sup_{|t|\le c}\int b\wedge\ell\left(\sqrt{n}\{S_n - \tau\circ T_a(P_{\theta_n,\zeta_n})\}\right) dP^{\otimes n}_{\theta_n,\zeta_n},\end{aligned}$$

which implies the conclusion.

A.1.4.2. *Proof of (ii).* Pick arbitrary $r > 0$ and $b > 0$. Since $T(P_n) = \bar{T}(P_n)$ for all $(x_1, \ldots, x_n) \in \mathcal{X}_n^n$,

$$\begin{aligned}
&\lim_{n\to\infty} \sup_{Q\in \bar{B}_H(P_0, r/\sqrt{n})} \int b \wedge \ell\left(\sqrt{n}\left\{\tau \circ T(P_n) - \tau \circ \bar{T}(Q)\right\}\right) dQ^{\otimes n} \\
\leq\ & \lim_{n\to\infty} \sup_{Q\in \bar{B}_H(P_0, r/\sqrt{n})} \int_{(x_1,\ldots,x_n)\notin \mathcal{X}_n^n} b \wedge \ell\left(\sqrt{n}\left\{\tau \circ T(P_n) - \tau \circ \bar{T}(Q)\right\}\right) dQ^{\otimes n} \\
&+ \lim_{n\to\infty} \sup_{Q\in \bar{B}_H(P_0, r/\sqrt{n})} \int_{(x_1,\ldots,x_n)\in \mathcal{X}_n^n} b \wedge \ell\left(\sqrt{n}\left\{\tau \circ \bar{T}(P_n) - \tau \circ \bar{T}(Q)\right\}\right) dQ^{\otimes n}.
\end{aligned} \tag{A.12}$$

An argument similar to (A.9) implies that the first term of (A.12) is zero. From $\mathcal{X}_n^n \subseteq \mathcal{X}^n$ and $\bar{B}_H(P_0, r/\sqrt{n}) \subseteq B_H(P_0, r/\sqrt{n})$, the second term of (A.12) is bounded from above by

$$\lim_{n\to\infty} \sup_{Q\in B_H(P_0, r/\sqrt{n})} \int b \wedge \ell\left(\sqrt{n}\left\{\tau \circ \bar{T}(P_n) - \tau \circ \bar{T}(Q)\right\}\right) dQ^{\otimes n} = \int b \wedge \ell dN(0, B^*),$$

where the equality follows from Lemma A.8, the uniform continuity of ℓ over $\bar{\mathbb{R}}^p$, and compactness of $B_H(P_0, r/\sqrt{n})$ under the Hellinger distance. Let $b \to \infty$ and the conclusion follows.

A.2. Auxiliary Lemmas

Lemma A.1. *Suppose that Assumption 3.1 holds. Then*

- **(i):** *for each $r > 0$, $\bar{T}(Q)$ exists for all $Q \in B_H(P_0, r/\sqrt{n})$ and all n large enough,*
- **(ii):** *$\bar{T}_{Q_n} \to \theta_0$ as $n \to \infty$ for each $r > 0$ and sequence $Q_n \in B_H(P_0, r/\sqrt{n})$.*

Proof of (i). The proof is split into several steps. Let $\mathcal{G}(\theta, Q)$ be the convex hull of the support of $g(x, \theta)$ under $x \sim Q$.

In the first step, we show $0 \in \text{int}\mathcal{G}(\theta_0, P_0)$. If $0 \notin \mathcal{G}(\theta_0, P_0)$, then we have $E_{P_0}[g(x, \theta_0)] \neq 0$, which is a contradiction. Thus, it is enough to show that 0 is not on the boundary of $\mathcal{G}(\theta_0, P_0)$. Suppose 0 is indeed on the boundary of $\mathcal{G}(\theta_0, P_0)$. In this case, we have two cases: (a) there exists a constant m-vector $a \neq 0$ such that $a'g \geq 0$ for all $g \in \mathcal{G}(\theta_0, P_0)$ and $P_0\{g \in \mathcal{G}(\theta_0, P_0) : a'g > 0\} > 0$, or (b) there exists $a \neq 0$ such that $a'g = 0$ for all $g \in \mathcal{G}(\theta_0, P_0)$. For the case (a), we have $a'E_{P_0}[g(x, \theta_0)] > 0$, which contradicts with $E_{P_0}[g(x, \theta_0)] = 0$. For the case (b), we have $a'E_{P_0}\left[g(x, \theta_0) g(x, \theta_0)'\right] a = 0$, which contradicts with Assumption 3.1 (vi).

In the second step, we show that for each $r > 0$, there exists $\delta > 0$ such that $0 \in \text{int}\mathcal{G}(\theta, Q)$ for all $|\theta - \theta_0| \leq \delta$ and all $Q \in B_H(P_0, \delta)$. Pick any $r > 0$. From the first step, we can find $m+1$ points $\{\tilde{g}_1, \ldots, \tilde{g}_{m+1}\} = \{g(\tilde{x}_1, \theta_0), \ldots, g(\tilde{x}_{m+1}, \theta_0)\}$ in the support of $g(x, \theta_0)$ under $x \sim P_0$ such that 0 is interior of the convex hull of $\{\tilde{g}_1, \ldots, \tilde{g}_{m+1}\}$. From the property of the convex hull (Rockafeller, 1970, Corollary 2.3.1), we can take $c_r > 0$ such that for any points $\{g_1, \ldots, g_{m+1}\}$ satisfying $|g_j - \tilde{g}_j| \leq c_r$

for $j=1,\ldots,m+1$, the interior of the convex hull of $\{g_1,\ldots,g_{m+1}\}$ contains 0. Let us take any $j=1,\ldots,m+1$. For the second step, it is sufficient to show that there exists $\delta_j>0$ such that $Q\{|g(x,\theta)-\tilde{g}_j|\le c_r\}>0$ for all $|\theta-\theta_0|\le\delta_j$ and all $Q\in B_H(P_0,\delta_j)$. Suppose this is false, i.e., for any $\delta_j>0$, we can take a pair (Q_j,θ_j) such that $H(Q_j,P_0)\le\delta_j$, $|\theta_j-\theta_0|\le\delta_j$, and $Q_j\{|g(x,\theta_j)-\tilde{g}_j|\le c_r\}=0$. Then we have

$$\delta_j\ge H(Q_j,P_0)\ge\sqrt{\int_{\{|g(x,\theta_j)-\tilde{g}_j|\le c_r\}}\left(\sqrt{dQ_j}-\sqrt{dP_0}\right)^2}=\sqrt{P_0\{|g(x,\theta_j)-\tilde{g}_j|\le c_r\}}.$$

On the other hand, by Assumption 3.1 (iv), the dominated convergence theorem guarantees

$$P_0\{|g(x,\theta_j)-\tilde{g}_j|\le c_r\}\to P_0\{|g(x,\theta_0)-\tilde{g}_j|\le c_r\}>0\qquad\text{as }\theta_j\to\theta_0.$$

Since δ_j can be arbitrarily small, we have a contradiction. This completes the second step.

In the third step, we show that for each $r>0$, there exists $\delta>0$ such that $R_n(\theta,Q)=\inf_{P\in\bar{\mathcal{P}}_\theta,P\ll Q}H(P,Q)$ has a minimum on $\{\theta\in\Theta:|\theta-\theta_0|\le\delta\}$ for all $Q\in B_H\left(P_0,\frac{r}{\sqrt{n}}\right)$ and all n large enough. Let us take $\delta>0$ to satisfy the conclusion of the second step. By Assumption 3.1 (iv), we can take N_δ to satisfy $\max_{1\le j\le m+1}\sup_{\theta\in\Theta,|\theta-\theta_0|\le\delta}|g(\tilde{x}_j,\theta)|\le m_{N_\delta}$. Thus, letting $\mathcal{G}_n(\theta,Q)$ be the convex hull of the support of $g_n(x,\theta)$ under $x\sim Q$, the second step also guarantees that for each $r>0$, there exists $\delta>0$ such that $0\in\operatorname{int}\mathcal{G}_n(\theta,Q)$ for all $|\theta-\theta_0|\le\delta$, all $Q\in B_H(P_0,\delta)$, and all $n\ge N_\delta$. Based on this, the convex duality result in Borwein and Lewis (1993, Theorem 3.4) implies $R_n(\theta,Q)=\sup_{\gamma\in\mathbb{R}^m}-\int\frac{1}{(1+\gamma' g_n(x,\theta))}dQ$ for all $|\theta-\theta_0|\le\delta$, all $Q\in B_H(P_0,\delta)$, and all $n\ge N_\delta$. Since $\sup_{\gamma\in\mathbb{R}^m}-\int\frac{1}{(1+\gamma' g_n(x,\theta))}dQ$ is continuous at all θ with $|\theta-\theta_0|\le\delta$ (by the maximum theorem), the Weierstrass theorem completes the third step.

Finally, based on the third step, it is sufficient for the conclusion to show that for every $r>0$, there exists $N\in\mathbb{N}$ such that $R_n(\theta_0,Q)<\inf_{\theta\in\Theta:|\theta-\theta_0|>\delta}R_n(\theta,Q)$ for all $n\ge N$ and all $Q\in B_H\left(P_0,\frac{r}{\sqrt{n}}\right)$. Pick any $r>0$. We first derive an upper bound of $R_n(\theta_0,Q)=\sup_{\gamma\in\mathbb{R}^m}-\int\frac{1}{(1+\gamma' g_n(x,\theta_0))}dQ$. From Lemma A.5 (ii), $\gamma_n(\theta_0,Q)=\arg\max_{\gamma\in\mathbb{R}^m}-\int\frac{1}{(1+\gamma' g_n(x,\theta_0))}dQ$ exists and $\sup_{x\in\mathcal{X}}\left|\gamma_n(\theta_0,Q)' g_n(x,\theta_0)\right|\le\frac{1}{2}$ for all n large enough and all $Q\in B_H\left(P_0,\frac{r}{\sqrt{n}}\right)$. Thus, by a second-order expansion around $\gamma_n(\theta_0,Q)=0$, we have

$$R_n(\theta_0,Q)\quad\le\quad -1+\gamma_n(\theta_0,Q)'\int g_n(x,\theta_0)\,dQ.$$

Define $C^*=\inf_{\theta\in\Theta:|\theta-\theta_0|>\delta}|E_{P_0}[g(x,\theta)]|^2/(1+|E_{P_0}[g(x,\theta)]|)>0$. From Lemma A.5 and $m_n n^{-1/2}\to 0$, it holds

$$\text{(A.1)}\qquad m_n(R(\theta_0,Q)+1)\le m_n\left|\gamma_n(\theta_0,Q)'\int g_n(x,\theta_0)\,dQ\right|<\frac{C^*}{4},$$

for all n large enough and all $Q \in B_H\left(P_0, \frac{r}{\sqrt{n}}\right)$. We now derive a lower bound of $R_n(\theta, Q)$ with $|\theta - \theta_0| > \delta$. Pick any $\theta \in \Theta$ such that $|\theta - \theta_0| > \delta$, and take any n large enough and $Q \in B_H\left(P_0, \frac{r}{\sqrt{n}}\right)$ to satisfy (A.1). If $0 \notin \mathcal{G}_n(\theta, Q)$, then $R_n(\theta, Q) = +\infty$. Thus, we concentrate on the case of $0 \in \mathcal{G}_n(\theta, Q)$, which guarantees $R_n(\theta, Q) = \sup_{\gamma \in \mathbb{R}^m} -\int \frac{1}{(1+\gamma' g_n(x,\theta))} dQ$ (Borwein and Lewis, 1993, Theorem 3.4). Let $\gamma_0(\theta) = E_{P_0}[g(x,\theta)] / (1 + |E_{P_0}[g(x,\theta)]|)$. Observe that

$$
\begin{aligned}
R_n(\theta, Q) &\geq -\int \frac{1}{\left(1 + m_n^{-1} \gamma_0(\theta)' g_n(x,\theta)\right)} dQ \\
&= -1 + m_n^{-1} \gamma_0(\theta)' \int g_n(x,\theta)\, dQ - m_n^{-2} \int \frac{\left(\gamma_0(\theta)' g_n(x,\theta)\right)^2}{\left(1 + t(x)\, m_n^{-1} \gamma_0(\theta)' g_n(x,\theta)\right)^3} dQ,
\end{aligned}
$$

where the second equality follows from an expansion ($t(x) \in (0,1)$ for almost every x under Q). From a similar argument to Lemma A.5 with $\sup_{\theta \in \Theta} |\gamma_0(\theta)| \leq 1$ and $m_n \to \infty$,

$$
\sup_{\theta \in \Theta} \left| \int g_n(x,\theta)\, dQ - \int g(x,\theta)\, dP_0 \right| \leq \frac{C^*}{4}, \qquad m_n^{-1} \sup_{\theta \in \Theta} \left| \int \frac{\left(\gamma_0(\theta)' g_n(x,\theta)\right)^2}{\left(1 + t_1(x)\, m_n^{-1} \gamma_0(\theta)' g_n(x,\theta)\right)^3} dQ \right| \leq \frac{C^*}{4},
$$

for all n large enough and all $Q \in B_H\left(P_0, \frac{r}{\sqrt{n}}\right)$. Combining these results and using the definition of C^*, we obtain

$$
\inf_{\theta \in \Theta : |\theta - \theta_0| > \delta} m_n \left(R_n(\theta, Q) + 1\right) \geq \frac{C^*}{2}, \tag{A.2}
$$

for all n large enough and all $Q \in B_H\left(P_0, \frac{r}{\sqrt{n}}\right)$. Therefore, (A.1) and (A.2) complete the proof of the final step.

Proof of (ii). Pick arbitrary $r > 0$ and sequence $Q_n \in B_H(P_0, r/\sqrt{n})$. From the triangle inequality,

(A.3)

$$
\sup_{\theta \in \Theta} \left| E_{Q_n}[g_n(x,\theta)] - E_{P_0}[g(x,\theta)] \right| \leq \sup_{\theta \in \Theta} \left| E_{Q_n}[g_n(x,\theta)] - E_{P_0}[g_n(x,\theta)] \right| + \sup_{\theta \in \Theta} \left| E_{P_0}[g(x,\theta)\, \mathbb{I}\{x \notin \mathcal{X}_n\}] \right|.
$$

The first term of (A.3) satisfies

$$
\begin{aligned}
& \sup_{\theta \in \Theta} \left| E_{Q_n}[g_n(x,\theta)] - E_{P_0}[g_n(x,\theta)] \right| \\
\leq\ & \sup_{\theta \in \Theta} \left| \int g_n(x,\theta) \left\{ dQ_n^{1/2} - dP_0^{1/2} \right\}^2 \right| + 2 \sup_{\theta \in \Theta} \left| \int g_n(x,\theta)\, dP_0^{1/2} \left\{ dQ_n^{1/2} - dP_0^{1/2} \right\} \right| \\
\leq\ & m_n \frac{r^2}{n} + 2 \sqrt{E_{P_0}\left[\sup_{\theta \in \Theta} |g(x,\theta)|^2\right]} \frac{r}{\sqrt{n}} = O\left(n^{-1/2}\right),
\end{aligned}
$$

where the first inequality follows from the triangle inequality, the second inequality follows from $Q_n \in B_H(P_0, r/\sqrt{n})$ and the Cauchy-Schwarz inequality, and the equality follows from Assumption

3.1 (v) and (vii). The second term of (A.3) satisfies

$$
\begin{aligned}
&\sup_{\theta\in\Theta}\left|E_{P_0}\left[g(x,\theta)\mathbb{I}\{x\notin\mathcal{X}_n\}\right]\right| \\
\leq\;& \left(\int \sup_{\theta\in\Theta}|g(x,\theta)|^{\eta}dP_0\right)^{1/\eta}\left(\int \mathbb{I}\{x\notin\mathcal{X}_n\}dP_0\right)^{(\eta-1)/\eta} \\
\leq\;& \left(E_{P_0}\left[\sup_{\theta\in\Theta}|g(x,\theta)|^{\eta}\right]\right)^{1/\eta}\left(m_n^{-\eta}E_{P_0}\left[\sup_{\theta\in\Theta}|g(x,\theta)|^{\eta}\right]\right)^{(\eta-1)/\eta}=o\left(n^{-1/2}\right), \qquad \text{(A.4)}
\end{aligned}
$$

where the first inequality follows from the Hölder inequality, and the second inequality follows from the Markov inequality, and the equality follows from Assumption 3.1 (v) and (vii). Combining these results, we obtain the uniform convergence $\sup_{\theta\in\Theta}|E_{Q_n}[g_n(x,\theta)]-E_{P_0}[g(x,\theta)]|\to 0$. Therefore, from the triangle inequality and $\left|E_{Q_n}\left[g_n\left(x,\bar{T}_{Q_n}\right)\right]\right|=O\left(n^{-1/2}\right)$ (Lemma A.6 (i)),

$$
\left|E_{P_0}\left[g\left(x,\bar{T}_{Q_n}\right)\right]\right|\leq\left|E_{P_0}\left[g\left(x,\bar{T}_{Q_n}\right)\right]-E_{Q_n}\left[g_n\left(x,\bar{T}_{Q_n}\right)\right]\right|+\left|E_{Q_n}\left[g_n\left(x,\bar{T}_{Q_n}\right)\right]\right|\to 0.
$$

The conclusion follows from Assumption 3.1 (iii).

Lemma A.2. *Suppose that Assumption 3.1 holds. Then for each* $r>0$ *and sequence* $Q_n\in B_H(P_0,r/\sqrt{n})$,

$$
\sqrt{n}\left(\bar{T}_{Q_n}-\theta_0\right)=-\sqrt{n}\Sigma^{-1}\int\Lambda_n dQ_n+o(1). \qquad \text{(A.5)}
$$

Proof. The proof is based on Rieder (1994, proofs of Theorems 6.3.4 and Theorem 6.4.5). Pick arbitrary $r>0$ and $Q_n\in B_H(P_0,r/\sqrt{n})$. Observe that

$$
\begin{aligned}
&\left\|dQ_n^{1/2}-d\bar{P}_{\theta_0,Q_n}^{1/2}+\frac{1}{2}\left(\bar{T}_{Q_n}-\theta_0\right)'\Lambda_n dQ_n^{1/2}\right\|^2 \\
=\;& \left\|dQ_n^{1/2}-d\bar{P}_{\theta_0,Q_n}^{1/2}+\frac{1}{2}\psi_{n,Q_n}'\Lambda_n dQ_n^{1/2}\right\|^2+\left\|\frac{1}{2}\left(\bar{T}_{Q_n}-\theta_0-\psi_{n,Q_n}\right)'\Lambda_n dQ_n^{1/2}\right\|^2 \\
&+\left\{\int\left(dQ_n^{1/2}-d\bar{P}_{\theta_0,Q_n}^{1/2}+\frac{1}{2}\psi_{n,Q_n}'\Lambda_n dQ_n^{1/2}\right)\Lambda_n' dQ_n^{1/2}\right\}\left(\bar{T}_{Q_n}-\theta_0-\psi_{n,Q_n}\right) \\
=\;& \left\|dQ_n^{1/2}-d\bar{P}_{\theta_0,Q_n}^{1/2}+\frac{1}{2}\psi_{n,Q_n}'\Lambda_n dQ_n^{1/2}\right\|^2+\left\|\frac{1}{2}\left(\bar{T}_{Q_n}-\theta_0-\psi_{n,Q_n}\right)'\Lambda_n dQ_n^{1/2}\right\|^2, \qquad \text{(A.6)}
\end{aligned}
$$

where the second equality follows from

$$
\begin{aligned}
&\int\left\{dQ_n^{1/2}-d\bar{P}_{\theta_0,Q_n}^{1/2}+\frac{1}{2}\psi_{n,Q_n}'\Lambda_n dQ_n^{1/2}\right\}\Lambda_n' dQ_n^{1/2} \\
=\;& \int\Lambda_n'\left\{dQ_n^{1/2}-d\bar{P}_{\theta_0,Q_n}^{1/2}\right\}dQ_n^{1/2}+\frac{1}{2}\psi_{n,Q_n}'\int\Lambda_n\Lambda_n' dQ_n=0.
\end{aligned}
$$

The left hand side of (A.6) satisfies

$$\begin{aligned}
& \left\| dQ_n^{1/2} - d\bar{P}_{\theta_0,Q_n}^{1/2} + \frac{1}{2}\left(\bar{T}_{Q_n} - \theta_0\right)' \Lambda_n dQ_n^{1/2} \right\| \\
\leq & \left\| dQ_n^{1/2} - d\bar{P}_{\bar{T}_{Q_n},Q_n}^{1/2} \right\| + o\left(\left|\bar{T}_{Q_n} - \theta_0\right|\right) + o\left(n^{-1/2}\right) \\
\leq & \left\| dQ_n^{1/2} - d\bar{P}_{\theta_0+\psi_{n,Q_n},Q_n}^{1/2} \right\| + o\left(\left|\bar{T}_{Q_n} - \theta_0\right|\right) + o\left(n^{-1/2}\right) \\
\leq & \left\| dQ_n^{1/2} - d\bar{P}_{\theta_0,Q_n}^{1/2} + \frac{1}{2}\psi_{n,Q_n}' \Lambda_n dQ_n^{1/2} \right\| + o\left(\left|\bar{T}_{Q_n} - \theta_0\right|\right) + o\left(\left|\psi_{n,Q_n}\right|\right) + o\left(n^{-1/2}\right), \qquad \text{(A.7)}
\end{aligned}$$

where the first inequality follows from the triangle inequality and Lemma A.3 (i), the second inequality follows from $\bar{T}_{Q_n} = \arg\min_{\theta\in\Theta} \left\| dQ_n^{1/2} - d\bar{P}_{\theta,Q_n}^{1/2} \right\|$, and the third inequality follows from the triangle inequality and Lemma A.3 (ii). From (A.6) and (A.7),

$$\begin{aligned}
& \left| \left\| dQ_n^{1/2} - d\bar{P}_{\theta_0,Q_n}^{1/2} + \frac{1}{2}\psi_{n,Q_n}' \Lambda_n dQ_n^{1/2} \right\|^2 + \left\| \frac{1}{2}\left(\bar{T}_{Q_n} - \theta_0 - \psi_{n,Q_n}\right)' \Lambda_n dQ_n^{1/2} \right\|^2 \right|^{1/2} \\
\leq & \left\| dQ_n^{1/2} - d\bar{P}_{\theta_0,Q_n}^{1/2} + \frac{1}{2}\psi_{n,Q_n}' \Lambda_n dQ_n^{1/2} \right\| + o\left(\left|\bar{T}_{Q_n} - \theta_0\right|\right) + o\left(\left|\psi_{n,Q_n}\right|\right) + o\left(n^{-1/2}\right).
\end{aligned}$$

This implies

$$\begin{aligned}
& o\left(\left|\bar{T}_{Q_n} - \theta_0\right|\right) + o\left(\left|\psi_{n,Q_n}\right|\right) + o\left(n^{-1/2}\right) \\
\geq & \sqrt{\frac{1}{4}\left(\bar{T}_{Q_n} - \theta_0 - \psi_{n,Q_n}\right)' \int \Lambda_n \Lambda_n' dQ_n \left(\bar{T}_{Q_n} - \theta_0 - \psi_{n,Q_n}\right)} \geq C\left|\bar{T}_{Q_n} - \theta_0 - \psi_{n,Q_n}\right|, \qquad \text{(A.8)}
\end{aligned}$$

for all n large enough, where the second inequality follows from Lemma A.5 (i) and Assumption 3.1 (vi).

We now analyze ψ_{n,Q_n}. From the definition of ψ_{n,Q_n},

$$\begin{aligned}
\psi_{n,Q_n} = & -2\left\{\left(\int \Lambda_n \Lambda_n' dQ_n\right)^{-1} - \Sigma^{-1}\right\} \int \Lambda_n \left\{dQ_n^{1/2} - d\bar{P}_{\theta_0,Q_n}^{1/2}\right\} dQ_n^{1/2} \\
& -2\Sigma^{-1} \int \Lambda_n \left\{dQ_n^{1/2} - d\bar{P}_{\theta_0,Q_n}^{1/2}\right\} dQ_n^{1/2}. \qquad \text{(A.9)}
\end{aligned}$$

From this and Lemma A.5 (i), the first term of (A.9) is $o\left(n^{-1/2}\right)$. The second term of (A.9) satisfies

$$\begin{aligned}
& -2\Sigma^{-1}\int \Lambda_n \left\{dQ_n^{1/2} - d\bar{P}_{\theta_0,Q_n}^{1/2}\right\} dQ_n^{1/2} \\
= \;& -2\Sigma^{-1}G'\Omega^{-1}\left(\int g_n\left(x,\theta_0\right) g_n\left(x,\theta_0\right)' dQ_n\right)\gamma_n\left(\theta_0,Q_n\right) \\
& +2\Sigma^{-1}G'\Omega^{-1}\left(\int \frac{\gamma_n\left(\theta_0,Q_n\right)' g_n\left(x,\theta_0\right)}{1+\gamma_n\left(\theta_0,Q_n\right)' g_n\left(x,\theta_0\right)} g_n\left(x,\theta_0\right) g_n\left(x,\theta_0\right)' dQ_n\right)\gamma_n\left(\theta_0,Q_n\right) \\
= \;& -\Sigma^{-1}G'\Omega^{-1}\left\{\int g_n\left(x,\theta_0\right) dQ_n + \frac{1}{2}\int \varrho_n\left(x,\theta_0,Q_n\right) g_n\left(x,\theta_0\right) dQ_n\right\} + o\left(n^{-1/2}\right) \\
= \;& -\Sigma^{-1}\int \Lambda_n dQ_n + o\left(n^{-1/2}\right),
\end{aligned}$$

where the first equality follows from (A.10), the second equality follows from (A.11) and Lemma A.5, and the third equality follows from Lemma A.5. Therefore,

$$\sqrt{n}\psi_{n,Q_n} = -\sqrt{n}\Sigma^{-1}\int \Lambda_n dQ_n + o\left(1\right),$$

which also implies $\left|\psi_{n,Q_n}\right| = O\left(n^{-1/2}\right)$ (by Lemma A.5 (i)). Combining this with (A.8),

$$\sqrt{n}\left(\bar{T}_{Q_n} - \theta_0\right) = \sqrt{n}\psi_{n,Q_n} + o\left(\sqrt{n}\left|\bar{T}_{Q_n} - \theta_0\right|\right) + o\left(1\right).$$

By solving this equation for $\sqrt{n}\left(\bar{T}_{Q_n} - \theta_0\right)$, the conclusion is obtained.

Lemma A.3. *Suppose that Assumption 3.1 holds. Then for each $r > 0$ and sequence $Q_n \in B_H\left(P_0, r/\sqrt{n}\right)$,*

(i): $\left\|d\bar{P}_{\bar{T}_{Q_n},Q_n}^{1/2} - d\bar{P}_{\theta_0,Q_n}^{1/2} + \frac{1}{2}\left(\bar{T}_{Q_n} - \theta_0\right)' \Lambda_n dQ_n^{1/2}\right\| = o\left(\left|\bar{T}_{Q_n} - \theta_0\right|\right) + o\left(n^{-1/2}\right)$,

(ii): $\left\|d\bar{P}_{\theta_0+\psi_{n,Q_n},Q_n}^{1/2} - d\bar{P}_{\theta_0,Q_n}^{1/2} + \frac{1}{2}\psi_{n,Q_n}' \Lambda_n dQ_n^{1/2}\right\| = o\left(\left|\psi_{n,Q_n}\right|\right) + o\left(n^{-1/2}\right)$.

Proof of (i). From the convex duality of partially finite programming (Borwein and Lewis (1993)), the Radon-Nikodym derivative $d\bar{P}_{\theta,Q}/dQ$ is written as

$$\text{(A.10)} \qquad \frac{d\bar{P}_{\theta,Q}}{dQ} = \frac{1}{\left(1+\gamma_n\left(\theta,Q\right)' g_n\left(x,\theta\right)\right)^2},$$

for each $n \in \mathbb{N}$, $\theta \in \Theta$, and $Q \in \mathcal{M}$, where $\gamma_n\left(\theta,Q\right)$ solves

$$\text{(A.11)} \quad 0 = \int \frac{g_n\left(x,\theta\right)}{\left(1+\gamma_n\left(\theta,Q\right)' g_n\left(x,\theta\right)\right)^2} dQ = E_Q\left[g_n\left(x,\theta\right)\left\{1 - 2\gamma_n\left(\theta,Q\right)' g_n\left(x,\theta\right) + \varrho_n\left(x,\theta,Q\right)\right\}\right],$$

with

$$\varrho_n\left(x,\theta,Q\right) = \frac{3\left(\gamma_n\left(\theta,Q\right)' g_n\left(x,\theta\right)\right)^2 + 2\left(\gamma_n(\theta,Q)' g_n\left(x,\theta\right)\right)^3}{\left(1+\gamma_n\left(\theta,Q\right)' g_n\left(x,\theta\right)\right)^2}.$$

Denote $t_n = \bar{T}_{Q_n} - \theta_0$. Pick arbitrary $r > 0$ and sequence $Q_n \in B_H(P_0, r/\sqrt{n})$. From the triangle inequality and (A.10),

$$
\begin{aligned}
& \left\| d\bar{P}^{1/2}_{\bar{T}_{Q_n},Q_n} - d\bar{P}^{1/2}_{\theta_0,Q_n} + \frac{1}{2} t_n' \Lambda_n dQ_n^{1/2} \right\| \\
\leq & \left\| \left\{ \gamma_n(\theta_0, Q_n)' g_n(x,\theta_0) - \gamma_n\left(\bar{T}_{Q_n}, Q_n\right)' g_n\left(x, \bar{T}_{Q_n}\right) \right\} dQ_n^{1/2} + \frac{1}{2} t_n' \Lambda_n dQ_n^{1/2} \right\| \\
& + \left\| \begin{array}{c} \left\{ \gamma_n(\theta_0, Q_n)' g_n(x,\theta_0) - \gamma_n\left(\bar{T}_{Q_n}, Q_n\right)' g_n\left(x, \bar{T}_{Q_n}\right) \right\} \\ \times \left\{ \frac{1}{\left(1+\gamma_n\left(\bar{T}_{Q_n},Q_n\right)' g_n\left(x,\bar{T}_{Q_n}\right)\right)\left(1+\gamma_n(\theta_0,Q_n)' g_n(x,\theta_0)\right)} - 1 \right\} dQ_n^{1/2} \end{array} \right\| = T_1 + T_2.
\end{aligned}
$$

For T_2, Lemmas A.5 and A.6 imply $T_2 = o\left(n^{-1/2}\right)$. For T_1, the triangle inequality and (A.11) yield

$$
\begin{aligned}
T_1 \leq & \left\| \left\{ \begin{array}{c} -\frac{1}{2} E_{Q_n}\left[g_n\left(x,\bar{T}_{Q_n}\right)\right]' E_{Q_n}\left[g_n\left(x,\bar{T}_{Q_n}\right) g_n\left(x,\bar{T}_{Q_n}\right)'\right]^{-1} g_n\left(x,\bar{T}_{Q_n}\right) \\ +\frac{1}{2} E_{Q_n}\left[g_n(x,\theta_0)\right]' E_{Q_n}\left[g_n(x,\theta_0) g_n(x,\theta_0)'\right]^{-1} g_n(x,\theta_0) + \frac{1}{2} t_n' \Lambda_n \end{array} \right\} dQ_n^{1/2} \right\| \\
& + \left\| E_{Q_n}\left[\varrho_n(x,\theta_0,Q_n) g_n(x,\theta_0)\right]' E_{Q_n}\left[g_n(x,\theta_0) g_n(x,\theta_0)'\right]^{-1} g_n(x,\theta_0)\, dQ_n^{1/2} \right\| \\
& + \left\| E_{Q_n}\left[\varrho_n\left(x,\bar{T}_{Q_n},Q_n\right) g_n\left(x,\bar{T}_{Q_n}\right)\right]' E_{Q_n}\left[g_n\left(x,\bar{T}_{Q_n}\right) g_n\left(x,\bar{T}_{Q_n}\right)'\right]^{-1} g_n(x,\theta_0)\, dQ_n^{1/2} \right\| \\
= & \; T_{11} + T_{12} + T_{13}.
\end{aligned}
$$

Lemmas A.5 and A.6 imply that $T_{12} = o\left(n^{-1/2}\right)$ and $T_{13} = o\left(n^{-1/2}\right)$. For T_{11}, expansions of $g_n\left(x, \bar{T}_{Q_n}\right)$ around $\bar{T}_{Q_n} = \theta_0$ yield

$$
\begin{aligned}
T_{11} \leq & \left\| -\frac{1}{2} E_{Q_n}\left[g_n\left(x,\bar{T}_{Q_n}\right)\right]' \left(\begin{array}{c} E_{Q_n}\left[g_n\left(x,\bar{T}_{Q_n}\right) g_n\left(x,\bar{T}_{Q_n}\right)'\right]^{-1} \\ -E_{Q_n}\left[g_n(x,\theta_0) g_n(x,\theta_0)'\right]^{-1} \end{array} \right) g_n\left(x,\bar{T}_{Q_n}\right) dQ_n^{1/2} \right\| \\
& + \left\| -\frac{1}{2} E_{Q_n}\left[g_n\left(x,\bar{T}_{Q_n}\right)\right]' E_{Q_n}\left[g_n(x,\theta_0) g_n(x,\theta_0)'\right]^{-1} \left\{ g_n\left(x,\bar{T}_{Q_n}\right) - g_n(x,\theta_0) \right\} dQ_n^{1/2} \right\| \\
& + \left\| -\frac{1}{2} t_n' \left(\int \frac{\partial g_n\left(x,\dot{\theta}\right)}{\partial \theta'} dQ_n - G \right)' E_{Q_n}\left[g_n(x,\theta_0) g_n(x,\theta_0)'\right]^{-1} g_n(x,\theta_0)\, dQ_n^{1/2} \right\| \\
& + \left\| \frac{1}{2} t_n' G' \left(\Omega^{-1} - E_{Q_n}\left[g_n(x,\theta_0) g_n(x,\theta_0)'\right]^{-1} \right) g_n(x,\theta_0)\, dQ_n^{1/2} \right\| \\
= & \; o\left(n^{-1/2}\right) + o(t_n),
\end{aligned}
$$

where $\dot{\theta}$ is a point on the line joining θ_0 and $\bar{T}_{Q_n}$, and the equality follows from Lemmas A.5 (i) and A.6 (i).

Proof of (ii). Similar to the proof of Part (i) of this lemma.

Lemma A.4. *Suppose that Assumption 3.1 hold. Then for each $t \in \mathbb{R}^p$,*

(i): $|E_{P_0}[g_n(x,\theta_0)]| = o\left(n^{-1/2}\right)$, $|E_{P_0}[g_n(x,\theta_n)]| = O\left(n^{-1/2}\right)$, $\left|E_{P_0}\left[g_n(x,\theta_n)\, g_n(x,\theta_n)'\right] - \Omega\right| = o(1)$, *and* $|E_{P_0}[\partial g_n(x,\theta_n)/\partial\theta'] - G| = o(1)$,

(ii): $\gamma_n(\theta_n, P_0) = \arg\max_{\gamma\in\mathbb{R}^m} - \int \frac{1}{(1+\gamma' g_n(x,\theta_n))} dP_0$ *exists for all n large enough,* $|\gamma_n(\theta_n, P_0)| = O\left(n^{-1/2}\right)$, *and* $\sup_{x\in\mathcal{X}} \left|\gamma_n(\theta_n, P_0)' g_n(x,\theta_n)\right| = o(1)$.

Proof of (i). Proof of the first statement. The same argument as (A.4) with Assumption 3.1 (iii) yields the conclusion.

Proof of the second statement. Pick an arbitrary $t \in \mathbb{R}^p$. From the triangle inequality,

$$|E_{P_0}[g_n(x,\theta_n)]| \le |E_{P_0}[g(x,\theta_n)\, \mathbb{I}\{x \notin \mathcal{X}_n\}]| + |E_{P_0}[g(x,\theta_n)]|. \tag{A.12}$$

By the same argument as (A.4) and $E_{P_0}[|g(x,\theta_n)|^\eta] < \infty$ (from Assumption 3.1 (v)), the first term of (A.12) is $o\left(n^{-1/2}\right)$. The second term of (A.12) satisfies

$$|E_{P_0}[g(x,\theta_n)]| \le E_{P_0}\left[\sup_{\theta\in\mathcal{N}}\left|\frac{\partial g(x,\theta)}{\partial\theta'}\right|\right]\left|\frac{t}{\sqrt{n}}\right| = O\left(n^{-1/2}\right),$$

for all n large enough, where the inequality follows from a Taylor expansion around $t = 0$ and Assumption 3.1 (iii), and the equality follows from Assumption 3.1 (v). Combining these results, the conclusion is obtained.

Proof of the third statement. Pick an arbitrary $t \in \mathbb{R}^p$. From the triangle inequality,

$$\begin{aligned} & \left|E_{P_0}\left[g_n(x,\theta_n)\, g_n(x,\theta_n)'\right] - \Omega\right| \\ \le\ & \left|E_{P_0}\left[g_n(x,\theta_n)\, g_n(x,\theta_n)'\right] - E_{P_0}\left[g(x,\theta_n)\, g(x,\theta_n)'\right]\right| + \left|E_{P_0}\left[g(x,\theta_n)\, g(x,\theta_n)'\right] - \Omega\right|. \end{aligned}$$

The first term is $o\left(n^{-1/2}\right)$ by the same argument as (A.4) and the second term converges to zero by the continuity of $g(x,\theta)$ at θ_0.

Proof of the fourth statement. Similar to the proof of the third statement.

Proof of (ii). Pick an arbitrary $t \in \mathbb{R}^p$. Let $\Gamma_n = \{\gamma \in \mathbb{R}^m : |\gamma| \le a_n\}$ with a positive sequence $\{a_n\}_{n\in\mathbb{N}}$ satisfying $a_n m_n \to 0$ and $a_n n^{1/2} \to \infty$. Observe that

$$\sup_{\gamma\in\Gamma_n, x\in\mathcal{X}, \theta\in\Theta} \left|\gamma' g_n(x,\theta)\right| \le a_n m_n \to 0. \tag{A.13}$$

Since $R_n(P_0, \theta_n, \gamma)$ is twice continuously differentiable with respect to γ and Γ_n is compact, $\tilde{\gamma} = \arg\max_{\gamma \in \Gamma_n} R_n(P_0, \theta_n, \gamma)$ exists for each $n \in \mathbb{N}$. A Taylor expansion around $\tilde{\gamma} = 0$ yields

$$
\begin{aligned}
-1 &= R_n(P_0, \theta_n, 0) \le R_n(P_0, \theta_n, \tilde{\gamma}) = -1 + \tilde{\gamma}' E_{P_0}[g_n(x, \theta_n)] - \tilde{\gamma}' E_{P_0}\left[\frac{g_n(x, \theta_n) g_n(x, \theta_n)'}{(1 + \dot{\gamma}' g_n(x, \theta_n))^3}\right] \tilde{\gamma} \\
&\le -1 + \tilde{\gamma}' E_{P_0}[g_n(x, \theta_n)] - C \tilde{\gamma}' E_{P_0}\left[g_n(x, \theta_n) g_n(x, \theta_n)'\right] \tilde{\gamma} \\
&\le -1 + |\tilde{\gamma}| \, |E_{P_0}[g_n(x, \theta_n)]| - C |\tilde{\gamma}|^2, \qquad \text{(A.14)}
\end{aligned}
$$

for all n large enough, where $\dot{\gamma}$ is a point on the line joining 0 and $\tilde{\gamma}$, the second inequality follows from (A.13), and the last inequality follows from Lemma A.4 (i) and Assumption 3.1 (vi). Thus, Lemma A.4 (i) implies

$$
C |\tilde{\gamma}| \le |E_{P_0}[g_n(x, \theta_n)]| = O\left(n^{-1/2}\right). \qquad \text{(A.15)}
$$

From $a_n n^{1/2} \to \infty$, $\tilde{\gamma}$ is an interior point of Γ_n and satisfies the first-order condition $\partial R_n(Q_n, \theta_0, \tilde{\gamma}) / \partial \gamma = 0$ for all n large enough. Since $R_n(Q_n, \theta_0, \gamma)$ is concave in γ for all n large enough, $\tilde{\gamma} = \arg\max_{\gamma \in \mathbb{R}^m} R_n(P_0, \theta_n, \gamma)$ for all n large enough and the first statement is obtained. Thus, the second statement is obtained from (A.15). The third statement follows from (A.15) and Assumption 3.1 (vii).

Lemma A.5. *Suppose that Assumption 3.1 holds. Then for each $r > 0$ and sequence $Q_n \in B_H(P_0, r/\sqrt{n})$,*

- **(i):** $|E_{Q_n}[g_n(x, \theta_0)]| = O\left(n^{-1/2}\right)$, *and* $\left|E_{Q_n}\left[g_n(x, \theta_0) g_n(x, \theta_0)'\right] - \Omega\right| = o(1)$,
- **(ii):** $\gamma_n(\theta_0, Q_n) = \arg\max_{\gamma \in \mathbb{R}^m} - \int \frac{1}{(1 + \gamma' g_n(x, \theta_0))} dQ_n$ *exists for all n large enough, and* $|\gamma_n(\theta_0, Q_n)| = O\left(n^{-1/2}\right)$, *and* $\sup_{x \in \mathcal{X}} \left|\gamma_n(\theta_0, Q_n)' g_n(x, \theta_0)\right| = o(1)$.

Proof of (i). Proof of the first statement. Pick any $r > 0$ and sequence $Q_n \in B_H(P_0, r/\sqrt{n})$. We have

$$
\begin{aligned}
& |E_{Q_n}[g_n(x, \theta_0)]| \\
\le\ & \left| \int g_n(x, \theta_0) \{dQ_n - dP_0\} \right| + |E_{P_0}[g_n(x, \theta_0)]| \\
\le\ & \left| \int g_n(x, \theta_0) \left\{dQ_n^{1/2} - dP_0^{1/2}\right\}^2 \right| + 2 \left| \int g_n(x, \theta_0) \, dP_0^{1/2} \left\{dQ_n^{1/2} - dP_0^{1/2}\right\} \right| + o\left(n^{-1/2}\right) \\
\le\ & m_n \frac{r^2}{n} + 2 E_{P_0}\left[|g(x, \theta_0)|^2\right] \frac{r}{\sqrt{n}} + o\left(n^{-1/2}\right) = O\left(n^{-1/2}\right),
\end{aligned}
$$

where the first and second inequalities follow from the triangle inequality and Lemma A.4 (i), the third inequality follows from the Cauchy-Schwarz inequality and $Q_n \in B_H(P_0, r/\sqrt{n})$, and the equality follows from Assumption 3.1 (v) and (vii).

Proof of the second statement. Pick arbitrary $r > 0$ and sequence $Q_n \in B_H(P_0, r/\sqrt{n})$. From the triangle inequality,

$$\text{(A.16)}\quad \left|E_{Q_n}\left[g_n(x,\theta_0)\, g_n(x,\theta_0)'\right] - \Omega\right|$$
$$\leq \left|E_{Q_n}\left[g_n(x,\theta_0)\, g_n(x,\theta_0)'\right] - E_{P_0}\left[g_n(x,\theta_0)\, g_n(x,\theta_0)'\right]\right| + \left|E_{P_0}\left[g(x,\theta_0)\, g(x,\theta_0)'\, \mathbb{I}\{x \notin \mathcal{X}_n\}\right]\right|.$$

The first term of the RHS of (A.16) satisfies

$$\begin{aligned} & \left|E_{Q_n}\left[g_n(x,\theta_0)\, g_n(x,\theta_0)'\right] - E_{P_0}\left[g_n(x,\theta_0)\, g_n(x,\theta_0)'\right]\right| \\ \leq\ & \left|\int g_n(x,\theta_0)\, g_n(x,\theta_0)' \left\{dQ_n^{1/2} - dP_0^{1/2}\right\}^2\right| + 2\left|\int g_n(x,\theta_0)\, g_n(x,\theta_0)'\, dP_0^{1/2}\left\{dQ_n^{1/2} - dP_0^{1/2}\right\}\right| \\ \leq\ & m_n^2 \frac{r^2}{n} + 2E_{P_0}\left[|g(x,\theta_0)|^4\right] \frac{r}{\sqrt{n}} = o(1), \end{aligned}$$

where the first inequality follows from the triangle inequality, the second inequality follows from the Cauchy-Schwarz inequality and $Q_n \in B_H(P_0, r/\sqrt{n})$, and the equality follows from Assumption 3.1 (v) and (vii). The second term of (A.16) satisfies

$$\begin{aligned} & \left|E_{P_0}\left[g(x,\theta_0)\, g(x,\theta_0)'\, \mathbb{I}\{x \notin \mathcal{X}_n\}\right]\right| \\ \leq\ & \left(\int \left|g(x,\theta_0)\, g(x,\theta_0)'\right|^{1+\delta} dP_0\right)^{\frac{1}{1+\delta}} \left(\int \mathbb{I}\{x \notin \mathcal{X}_n\}\, dP_0\right)^{\frac{\delta}{1+\delta}} \\ \leq\ & \left(E_{P_0}\left[|g(x,\theta_0)|^{2+\delta}\right]\right)^{\frac{1}{1+\delta}} \left(m_n^{-\eta} E_{P_0}\left[|g(x,\theta_0)|^{\eta}\right]\right)^{\frac{\delta}{1+\delta}} = o(1), \end{aligned}$$

for sufficiently small $\delta > 0$, where the first inequality follows from the Hölder inequality, the second inequality follows from the Markov inequality, and the equality follows from Assumption 3.1 (vii).

Proof of (ii). Similar to the proof of Lemma A.4 (ii). Repeat the same argument with $R_n(Q_n, \theta_0, \gamma)$ instead of $R_n(P_0, \theta_n, \gamma)$.

Lemma A.6. *Suppose that Assumption 3.1 holds. Then for each $r > 0$ and sequence $Q_n \in B_H(P_0, r/\sqrt{n})$,*

- **(i):** $\left|E_{Q_n}\left[g_n\left(x, \bar{T}_{Q_n}\right)\right]\right| = O\left(n^{-1/2}\right)$, $\left|E_{Q_n}\left[g_n\left(x, \bar{T}_{Q_n}\right) g_n\left(x, \bar{T}_{Q_n}\right)'\right] - \Omega\right| = o(1)$, *and* $\left|E_{Q_n}\left[\partial g_n\left(x, \bar{T}_{Q_n}\right)/\partial\theta'\right] - G\right| = o(1)$,
- **(ii):** $\gamma_n\left(\bar{T}_{Q_n}, Q_n\right) = \arg\max_{\gamma\in\mathbb{R}^m} - \int \frac{1}{\left(1+\gamma' g_n\left(x,\bar{T}_{Q_n}\right)\right)} dQ_n$ *exists for all n large enough,* $\left|\gamma_n\left(\bar{T}_{Q_n}, Q_n\right)\right| = O\left(n^{-1/2}\right)$, *and* $\sup_{x\in\mathcal{X}} \left|\gamma_n\left(\bar{T}_{Q_n}, Q_n\right)' g_n\left(x, \bar{T}_{Q_n}\right)\right| = o(1)$.

Proof of (i). Proof of the first statement. Pick any $r > 0$ and sequence $Q_n \in B_H(P_0, r/\sqrt{n})$. Define $\tilde{\gamma} = \frac{E_{Q_n}\left[g_n\left(x,\bar{T}_{Q_n}\right)\right]}{\sqrt{n}\left|E_{Q_n}\left[g_n\left(x,\bar{T}_{Q_n}\right)\right]\right|}$. Since $|\tilde{\gamma}| = n^{-1/2}$,

$$\text{(A.17)}\qquad \sup_{x\in\mathcal{X},\theta\in\Theta} \left|\tilde{\gamma}' g_n(x,\theta)\right| \leq n^{-1/2} m_n \to 0.$$

Observe that

(A.18)

$$\begin{aligned}
&\left|E_{Q_n}\left[g_n\left(x,\bar{T}_{Q_n}\right)g_n\left(x,\bar{T}_{Q_n}\right)'\right]\right| \\
\leq & \int \sup_{\theta\in\Theta}\left|g_n\left(x,\theta\right)\right|^2\left\{dQ_n^{1/2}-dP_0^{1/2}\right\}^2+2\int \sup_{\theta\in\Theta}\left|g_n\left(x,\theta\right)\right|^2 dP_0^{1/2}\left\{dQ_n^{1/2}-dP_0^{1/2}\right\}+E_{P_0}\left[\sup_{\theta\in\Theta}\left|g_n\left(x,\theta\right)\right|^2\right] \\
\leq & m_n^2\frac{r^2}{n}+2m_n\sqrt{E_{P_0}\left[\sup_{\theta\in\Theta}\left|g\left(x,\theta\right)\right|^2\right]}\frac{r}{\sqrt{n}}+E_{P_0}\left[\sup_{\theta\in\Theta}\left|g\left(x,\theta\right)\right|^2\right]\leq CE_{P_0}\left[\sup_{\theta\in\Theta}\left|g\left(x,\theta\right)\right|^2\right],
\end{aligned}$$

for all n large enough, where the first inequality follows from the triangle inequality, the second inequality follows from the Cauchy-Schwarz inequality and $Q_n \in B_H\left(P_0, r/\sqrt{n}\right)$, and the last inequality follows from Assumption 3.1 (v) and (vii). Thus, an expansion around $\tilde{\gamma}=0$ yields

$$\begin{aligned}
R_n\left(Q_n,\bar{T}_{Q_n},\tilde{\gamma}\right) &= -1+\tilde{\gamma}'E_{Q_n}\left[g_n\left(x,\bar{T}_{Q_n}\right)\right]-\tilde{\gamma}'E_{Q_n}\left[\frac{g_n\left(x,\bar{T}_{Q_n}\right)g_n\left(x,\bar{T}_{Q_n}\right)'}{\left(1+\dot{\gamma}'g_n\left(x,\bar{T}_{Q_n}\right)\right)^3}\right]\tilde{\gamma} \\
&\geq -1+n^{-1/2}\left|E_{Q_n}\left[g_n\left(x,\bar{T}_{Q_n}\right)\right]\right|-C\tilde{\gamma}'E_{Q_n}\left[g_n\left(x,\bar{T}_{Q_n}\right)g_n\left(x,\bar{T}_{Q_n}\right)'\right]\tilde{\gamma} \\
&\geq -1+n^{-1/2}\left|E_{Q_n}\left[g_n\left(x,\bar{T}_{Q_n}\right)\right]\right|-Cn^{-1}, \qquad \text{(A.19)}
\end{aligned}$$

for all n large enough, where $\dot{\gamma}$ is a point on the line joining 0 and $\tilde{\gamma}$, the first inequality follows from (A.17), and the second inequality follows from $\tilde{\gamma}'\tilde{\gamma}=n^{-1}$ and (A.18). From the duality of partially finite programming (Borwein and Lewis (1993)), $\gamma_n\left(\bar{T}_{Q_n},Q_n\right)$ and $\bar{T}_{Q_n}$ are written as $\gamma_n\left(\bar{T}_{Q_n},Q_n\right)=\arg\max_{\gamma\in\mathbb{R}^m} R_n\left(Q_n,\bar{T}_{Q_n},\gamma\right)$ and $\bar{T}_{Q_n}=\arg\min_{\theta\in\Theta} R_n\left(Q_n,\theta,\gamma_n\left(\theta,Q_n\right)\right)$. Therefore, from (A.19),

$$\begin{aligned}
&-1+n^{-1/2}\left|E_{Q_n}\left[g_n\left(x,\bar{T}_{Q_n}\right)\right]\right|-Cn^{-1} \\
\leq\ & R_n\left(Q_n,\bar{T}_{Q_n},\tilde{\gamma}\right)\leq R_n\left(Q_n,\bar{T}_{Q_n},\gamma_n\left(\bar{T}_{Q_n},Q_n\right)\right)\leq R_n\left(Q_n,\theta_0,\gamma_n\left(\theta_0,Q_n\right)\right). \qquad \text{(A.20)}
\end{aligned}$$

By a similar argument to (A.14) combined with $\left|\gamma_n\left(\theta_0,Q_n\right)\right|=O\left(n^{-1/2}\right)$ and $\left|E_{Q_n}\left[g_n\left(x,\theta_0\right)\right]\right|=O\left(n^{-1/2}\right)$ (by Lemma A.5), we have

(A.21)

$$R_n\left(Q_n,\theta_0,\gamma_n\left(\theta_0,Q_n\right)\right)\leq -1+\left|\gamma_n\left(\theta_0,Q_n\right)\right|\left|E_{Q_n}\left[g_n\left(x,\theta_0\right)\right]\right|-C\left|\gamma_n\left(\theta_0,Q_n\right)\right|^2=-1+O\left(n^{-1}\right).$$

From (A.20) and (A.21), the conclusion follows.

Proof of the second statement. Similar to the proof of the second statement of Lemma A.5 (i).

Proof of the third statement. Pick arbitrary $r > 0$ and sequence $Q_n \in B_H(P_0, r/\sqrt{n})$. From the triangle inequality,

$$\begin{aligned} \left|E_{Q_n}\left[\partial g_n\left(x,\bar{T}_{Q_n}\right)/\partial\theta'\right]-G\right| \leq& \left|E_{Q_n}\left[\partial g_n\left(x,\bar{T}_{Q_n}\right)/\partial\theta'\right]-E_{P_0}\left[\partial g_n\left(x,\bar{T}_{Q_n}\right)/\partial\theta'\right]\right| \\ &+\left|E_{P_0}\left[\mathbb{I}\{x\notin\mathcal{X}_n\}\partial g\left(x,\bar{T}_{Q_n}\right)/\partial\theta'\right]\right|+\left|E_{P_0}\left[\partial g\left(x,\bar{T}_{Q_n}\right)/\partial\theta'\right]-G\right|. \end{aligned} \tag{A.22}$$

The first term of (A.22) satisfies

$$\begin{aligned} &\left|E_{Q_n}\left[\partial g_n\left(x,\bar{T}_{Q_n}\right)/\partial\theta'\right]-E_{P_0}\left[\partial g_n\left(x,\bar{T}_{Q_n}\right)/\partial\theta'\right]\right| \\ \leq&\left|\int\partial g_n\left(x,\bar{T}_{Q_n}\right)/\partial\theta'\left\{dQ_n^{1/2}-dP_0^{1/2}\right\}^2\right|+2\left|\int\partial g_n\left(x,\bar{T}_{Q_n}\right)/\partial\theta' dP_0^{1/2}\left\{dQ_n^{1/2}-dP_0^{1/2}\right\}\right| \\ \leq&\sup_{x\in\mathcal{X}_n,\theta\in\mathcal{N}}\left|\partial g_n(x,\theta)/\partial\theta'\right|\frac{r^2}{n}+2E_{P_0}\left[\sup_{\theta\in\mathcal{N}}\left|\partial g_n(x,\theta)/\partial\theta'\right|^2\right]\frac{r}{\sqrt{n}}=o(1), \end{aligned}$$

where the first inequality follows from the triangle inequality, the second inequality follows from the Cauchy-Schwarz inequality, and the equality follows from Assumption 3.1 (v) and (vii). The second term of (A.22) is $o(1)$ by the same argument as (A.4). The third term of (A.22) is $o(1)$ by the continuity of $\partial g(x,\theta)/\partial\theta'$ at θ_0 and Lemma A.1 (ii). Therefore, the conclusion is obtained.

Proof of (ii). Similar to the proof of Lemma A.4 (ii). Repeat the same argument with $R_n\left(Q_n, \bar{T}_{Q_n}, \gamma\right)$ instead of $R_n(P_0, \theta_n, \gamma)$.

Lemma A.7. *Suppose that Assumption 3.1 holds. Then for each sequence $Q_n \in B_H(P_0, r/\sqrt{n})$ and $r > 0$, $\bar{T}_{P_n} \xrightarrow{p} \theta_0$ under Q_n.*

Proof. Similar to the proof of Lemma A.1 (i).

Lemma A.8. *Suppose that Assumption 3.1 holds . Then for each $r > 0$ and sequence $Q_n \in B_H(P_0, r/\sqrt{n})$,*

$$\begin{aligned} \sqrt{n}\left(\bar{T}_{P_n}-\theta_0\right) &= -\sqrt{n}\Sigma^{-1}\int\Lambda_n dP_n+o_p(1) \quad \text{under } Q_n, \\ \sqrt{n}\left(\bar{T}_{P_n}-\bar{T}_{Q_n}\right) &\xrightarrow{d} N\left(0,\Sigma^{-1}\right) \quad \text{under } Q_n. \end{aligned}$$

Proof. The proof of the first statement is similar to that of Lemma A.2 (replace Q_n with P_n and use Lemmas A.9 and A.10 instead of Lemmas A.5 and A.6). For the second statement, Lemma A.2 and the first statement imply

$$\sqrt{n}\left(\bar{T}_{P_n}-\bar{T}_{Q_n}\right)=-\Sigma^{-1}G'\Omega^{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^n\left\{g_n(x_i,\theta_0)-E_{Q_n}\left[g_n(x,\theta_0)\right]\right\}+o_p(1),$$

under Q_n. Thus, it is sufficient to check that we can apply a central limit theorem to the triangular array $\{g_n(x_i,\theta_0)\}_{1\le i\le n,n}$. Observe that

$$\begin{aligned} & E_{Q_n}\left[|g_n(x,\theta_0)|^{2+\epsilon}\right] \\ = & \int |g_n(x,\theta_0)|^{2+\epsilon}\left\{dQ_n^{1/2}-dP_0^{1/2}\right\}^2 + 2\int |g_n(x,\theta_0)|^{2+\epsilon} dP_0^{1/2}\left\{dQ_n^{1/2}-dP_0^{1/2}\right\} + E_{P_0}\left[|g_n(x,\theta_0)|^{2+\epsilon}\right] \\ \le & m_n^{2+\epsilon}\frac{r^2}{n} + 2m_n^{1+\epsilon}E_{P_0}\left[|g(x,\theta_0)|^2\right]\frac{r}{\sqrt{n}} + E_{P_0}\left[|g(x,\theta_0)|^{2+\epsilon}\right] < \infty, \end{aligned}$$

for all n large enough, where the first inequality follows from the Cauchy-Schwarz inequality, and the second inequality follows from Assumption 3.1 (v) and (vii). Therefore, the conclusion is obtained.

Lemma A.9. *Suppose that Assumption 3.1 holds. Then for each $r>0$ and sequence $Q_n\in B_H(P_0,r/\sqrt{n})$, the followings hold under Q_n:*

- **(i):** $|E_{P_n}[g_n(x,\theta_0)]| = O_p(n^{-1/2})$, $|E_{P_n}[g_n(x,\theta_0)g_n(x,\theta_0)'] - \Omega| = o_p(1)$,
- **(ii):** $\gamma_n(\theta_0,P_n) = \arg\max_{\gamma\in\mathbb{R}^m} -\int \frac{1}{(1+\gamma' g_n(x,\theta_0))}dP_n$ *exists a.s. for all n large enough,* $|\gamma_n(\theta_0,P_n)| = O_p(n^{-1/2})$, *and* $\sup_{x\in\mathcal{X}}|\gamma_n(\theta_0,P_n)' g_n(x,\theta_0)| = o_p(1)$.

Proof of (i). Proof of the first statement. From the triangle inequality,

$$|E_{P_n}[g_n(x,\theta_0)]| \le |E_{P_n}[g_n(x,\theta_0)] - E_{Q_n}[g_n(x,\theta_0)]| + |E_{Q_n}[g_n(x,\theta_0)]|.$$

The first term is $O_p(n^{-1/2})$ by the central limit theorem for the triangular array $\{g_n(x_i,\theta_0)\}_{1\le i\le n,n}$. The second term is $O(n^{-1/2})$ by Lemma A.5 (i).

Proof of the second statement. From the triangle inequality,

$$\begin{aligned} & |E_{P_n}[g_n(x,\theta_0)g_n(x,\theta_0)' - \Omega]| \\ \le & |E_{P_n}[g_n(x,\theta_0)g_n(x,\theta_0)'] - E_{Q_n}[g_n(x,\theta_0)g_n(x,\theta_0)']| + |E_{Q_n}[g_n(x,\theta_0)g_n(x,\theta_0)'] - \Omega|. \end{aligned}$$

From a law of large numbers, the first term is $o_p(1)$. From Lemma A.5 (i), the second term is $o(1)$.

Proof of (ii). Similar to the proof of Lemma A.4 (ii) except using Lemma A.9 (i) instead of Lemma A.4 (i).

Lemma A.10. *Suppose that Assumption 3.1 holds. Then for each $r>0$ and sequence $Q_n\in B_H(P_0,r/\sqrt{n})$, the followings hold under Q_n:*

- **(i):** $|E_{P_n}[g_n(x,\bar{T}_{P_n})]| = O_p(n^{-1/2})$, $\left|E_{P_n}\left[g_n(x,\bar{T}_{P_n})g_n(x,\bar{T}_{P_n})'\right] - \Omega\right| = O_p(n^{-1/2})$, *and* $|E_{P_n}[\partial g_n(x,\bar{T}_{P_n})/\partial\theta'] - G| = o_p(1)$,

(ii): $\gamma_n\left(\bar{T}_{P_n}, P_n\right) = \arg\max_{\gamma\in\mathbb{R}^m} - \int \frac{1}{\left(1+\gamma' g_n\left(x,\bar{T}_{P_n}\right)\right)} dP_n$ *exists a.s. for all* n *large enough,* $\left|\gamma_n\left(\bar{T}_{P_n} P_n\right)\right| = O_p\left(n^{-1/2}\right)$*, and* $\sup_{x\in\mathcal{X}}\left|\gamma_n\left(\bar{T}_{P_n}, P_n\right)' g_n\left(x, \bar{T}_{P_n}\right)\right| = o_p(1)$.

Proof of (i). Similar to the proof of Lemma A.6 (i).

Proof of (ii). Similar to the proof of Lemma A.6 (ii).

References

Borwein, J. M., and A. S. Lewis (1993): "Partially-finite programming in L_1 and the existence of maximum entropy estimates," *SIAM Journal of Optimization*, 3, 248–267.

LeCam, L., and G. Yang (1990): *Asymptotics in Statistics: Some Basic Concepts*. Springer.

van der Vaart, A. (1998): *Asymptotic Statistics*. Cambridge University Press.

Cowles Foundation for Research in Economics, Yale University, New Haven, CT-06520.

E-mail address: yuichi.kitamura@yale.edu

Cowles Foundation for Research in Economics, Yale University, New Haven, CT-06520.

E-mail address: taisuke.otsu@yale.edu

Department of Economics, Princeton University, Princeton, NJ-08544.

E-mail address: kevdokim@princeton.edu