Econometrica Supplementary Material

SUPPLEMENT TO "ON THE ASYMPTOTIC OPTIMALITY OF EMPIRICAL LIKELIHOOD FOR TESTING MOMENT RESTRICTIONS"

(*Econometrica*, Vol. 80, No. 1, January 2012, 413–423)

BY YUICHI KITAMURA, ANDRES SANTOS AND AZEEM M. SHAIKH

APPENDIX

THE FOLLOWING TABLE of notation and definitions will be used throughout this appendix:

co(A)**:** the convex hull of a set A,
supp(Q)**:** the support of a measure $Q \in \mathbf{M}$,
$\text{supp}_P(g(X, \theta))$**:** the support of $g(X, \theta)$ when X is distributed according to $P \in \mathbf{M}$,
$s(Q, \theta)$**:** the dimension of the $\text{co}(\text{supp}_P(g(X, \theta)))$.

The principal challenge in deriving our optimality result is establishing part (a) of Theorem 3.1. For ease of exposition, we provide an outline of the proof of this claim before its formal derivation:

Step 1. First we note that $\Lambda_1(\eta) \subseteq \ddot{\Lambda}_1(\eta)$, where for

$$\begin{aligned}\ddot{\mathbf{M}}_0(Q) = \{&P \in \mathbf{M} : P \ll Q, Q \ll P, s(Q, \theta) = m, E_P[g(X, \theta)] \\ &= 0 \text{ for some } \theta \in \Theta\},\end{aligned}$$

the set $\ddot{\Lambda}_1(\eta)$ is given by

$$\ddot{\Lambda}_1(\eta) = \left\{Q \in \mathbf{M} : \inf_{P \in \ddot{\mathbf{M}}_0(Q)} I(Q|P) \geq \eta\right\}.$$

Step 2. Lemma A.2 exploits Lemma A.1 to show that $\ddot{\Lambda}_1(\eta)$ is closed in the weak topology. Sanov's theorem (see Theorem 6.2.10 in Dembo and Zeitouni (1998)) then implies that

$$\begin{aligned}\text{(A1)} \qquad \limsup_{n\to\infty} \frac{1}{n} \log P^n\{\hat{P}_n \in \Lambda_1(\eta)\} &\leq \limsup_{n\to\infty} \frac{1}{n} \log P^n\{\hat{P}_n \in \ddot{\Lambda}_1(\eta)\} \\ &\leq -\inf_{Q \in \ddot{\Lambda}_1(\eta)} I(Q|P).\end{aligned}$$

Step 3. In Lemma A.3, we derive conditions (see (A37)) under which, for any $P \in \mathbf{P}_0$, there is $\bar{\eta}(P) > 0$ such that

$$\text{(A2)} \qquad \inf_{Q \in \ddot{\Lambda}_1(\eta)} I(Q|P) \geq \eta$$

 DOI: 10.3982/ECTA8773

for all $\eta \leq \bar{\eta}(P)$, which, in light of equation (A1), establishes the desired result pointwise in $P \in \mathbf{P}_0$.

Step 4. Part (a) of Theorem 3.1 is then established by showing that (A2) in fact holds uniformly in $P \in \mathbf{P}_0$. In particular, we show that for all $\eta > 0$ sufficiently small,

$$\inf_{P \in \mathbf{P}_0} \inf_{Q \in \ddot{A}_1(\eta)} I(Q|P) \geq \eta. \tag{A3}$$

These derivations exploit Lemmas A.4–A.9.

LEMMA A.1: *Let* $\{B_1, \dots, B_{d+1}\}$ *be a collection of closed and bounded balls in* $\mathbf{R}^d$ *such that for any collection of points* $\{g_1, \dots, g_{d+1}\}$ *with* $g_i \in B_i$ *for each* $1 \leq i \leq d+1$,

$$0 \in \operatorname{int}\bigl(\operatorname{co}(\{g_1, \dots, g_{d+1}\})\bigr) \tag{A4}$$

(*relative to the topology on* $\mathbf{R}^d$). *Then there exists* $\varepsilon > 0$ *such that for all* $0 \neq \gamma \in \mathbf{R}^d$, *there exists* $j = j(\gamma) \in \{1, \dots, d+1\}$ *such that* $\gamma' g < 0$ *and* $|\gamma' g| \geq |g||\gamma|\varepsilon$ *for all* $g \in B_j$.

PROOF: Let $\mathbf{B}(\gamma)$ be the maximal subset of $\{B_1, \dots, B_{d+1}\}$ such that for all $B \in \mathbf{B}(\gamma)$, we have that $\gamma' g < 0$ for all $g \in B$. The desired claim follows if we can show (i) $\mathbf{B}(\gamma)$ is nonempty for any $0 \neq \gamma \in \mathbf{R}^d$ and (ii)

$$\inf_{0 \neq \gamma \in \mathbf{R}^d} \varepsilon(\gamma) > 0,$$

where

$$\varepsilon(\gamma) = \max_{B \in \mathbf{B}(\gamma)} \inf_{g \in B} \frac{|\gamma' g|}{|g||\gamma|}.$$

To establish (i), consider the hyperplane $H_\gamma = \{g \in \mathbf{R}^d : \gamma' g = 0\}$ and note that if $\gamma \neq 0$, then H_γ must strongly separate at least two balls $B_i, B_k \in \{B_1, \dots, B_{d+1}\}$ with $i \neq k$. Otherwise, for either $\bar{\gamma} = \gamma$ or $\bar{\gamma} = -\gamma$, there exists a collection of points $\{g_1, \dots, g_{d+1}\}$ with $g_i \in B_i$ and $\bar{\gamma}' g_i \geq 0$ for each $1 \leq i \leq d+1$, which contradicts (A4). Therefore, since H_γ strongly separates at least two balls $B_i, B_k \in \{B_1, \dots, B_{d+1}\}$ with $i \neq k$, it follows that there exists a $j = j(\gamma)$ such that $\gamma' g_j < 0$ for all $g \in B_j$.

To establish (ii), note that we may assume without loss of generality that $|\gamma| = 1$ and suppose by way of contradiction that there exists a sequence $\{\gamma_n\}_{n=1}^\infty$ such that $\varepsilon(\gamma_n) \to 0$. Since $|\gamma_n| = 1$, we have that there exists a subsequence $\{\gamma_{n_k}\}_{k=1}^\infty$ such that $\gamma_{n_k} \to \gamma^*$ and $|\gamma^*| = 1$. Moreover, since $\mathbf{B}(\gamma^*) \subseteq \mathbf{B}(\gamma_{n_k})$ for all k sufficiently large, it follows that for such k,

$$\varepsilon(\gamma_{n_k}) \geq \max_{B \in \mathbf{B}(\gamma^*)} \inf_{g \in B} \frac{|\gamma'_{n_k} g|}{|g|}. \tag{A5}$$

Next note that

$$\text{(A6)} \qquad \inf_{g \in B_i} |g| > 0$$

for $1 \leq i \leq d+1$. To see this, note that if (A6) fails, there exists $1 \leq i^* \leq d+1$ such that $0 \in B_{i^*}$ since each B_i is closed. In this case, any collection of points $\{g_1, \ldots, g_{d+1}\}$ with $g_i \in B_i$ for $1 \leq i \leq d+1$ and $g_{i^*} = 0$ will not satisfy (A4). It thus follows that $|\gamma'_{n_k} g|/|g| \to |\gamma^{*\prime} g|/|g|$ uniformly over $g \in B$ for each $B \in \mathbf{B}(\gamma^*)$. The right hand side of (A5) therefore tends to $\varepsilon(\gamma^*)$. But since each $B \in \mathbf{B}(\gamma^*)$ is compact and there are only finitely many such B, $\varepsilon(\gamma^*) > 0$. Hence, $\varepsilon(\gamma_n) \not\to 0$, from which the desired claim follows. *Q.E.D.*

LEMMA A.2: *Let Assumptions* 3.1 *and* 3.2 *hold*, *and define the set of probability measures*

$$\text{(A7)} \qquad \ddot{\mathbf{M}}_0(Q) = \big\{P \in \mathbf{M} : P \ll Q, Q \ll P, s(Q, \theta) = m, \\ E_P[g(X, \theta)] = 0 \textit{ for some } \theta \in \Theta\big\}.$$

Accordingly, *also denote the rejection region resulting from employing* $\ddot{\mathbf{M}}_0(Q)$ *instead of* $\mathbf{M}_0(Q)$ *by*

$$\text{(A8)} \qquad \ddot{\Lambda}_1(\eta) = \Big\{Q \in \mathbf{M} : \inf_{P \in \ddot{\mathbf{M}}_0(Q)} I(Q|P) \geq \eta\Big\}.$$

It then follows that $\ddot{\Lambda}_1(\eta)$ *is closed under the weak topology for any* $\eta > 0$.

PROOF: Let $\{Q_n\}_{n=1}^{\infty}$ be a sequence such that $Q_n \to Q$ and $Q_n \in \ddot{\Lambda}_1(\eta)$ for all n. We wish to show that $Q \in \ddot{\Lambda}_1(\eta)$. Note that if $\ddot{\mathbf{M}}_0(Q) = \emptyset$, then

$$\inf_{P \in \ddot{\mathbf{M}}_0(Q)} I(Q|P) = +\infty,$$

so $Q \in \ddot{\Lambda}_1(\eta)$. We may therefore assume further that $\ddot{\mathbf{M}}_0(Q) \neq \emptyset$.

Now suppose by way of contradiction that $Q \notin \ddot{\Lambda}_1(\eta)$. Define the set

$$\text{(A9)} \qquad \ddot{\mathbf{M}}_0(Q, \theta) = \big\{P \in \mathbf{M} : P \ll Q, Q \ll P, s(Q, \theta) = m, E_P[g(X, \theta)] = 0\big\}$$

and note that $\ddot{\mathbf{M}}_0(Q) = \bigcup_{\theta \in \Theta} \ddot{\mathbf{M}}_0(Q, \theta)$. Further define the set

$$\text{(A10)} \qquad \Theta(Q) = \{\theta \in \Theta : \ddot{\mathbf{M}}_0(Q, \theta) \neq \emptyset\}.$$

Since $\ddot{\mathbf{M}}_0(Q) \neq \emptyset$, it follows that $\Theta(Q) \neq \emptyset$ and, therefore, the primal constraint qualification of Theorem 3.4 of Borwein and Lewis (1993) is satisfied for all $\theta \in \Theta(Q)$. Hence,

$$\text{(A11)} \qquad \inf_{P \in \ddot{\mathbf{M}}_0(Q)} I(Q|P) = \inf_{\theta \in \Theta(Q)} \max_{\gamma \in \mathbf{R}^m} \int \log(1 + \gamma' g(x, \theta))\, dQ.$$

Therefore, since $Q \notin \ddot{\Lambda}_1(\eta)$ by hypothesis, it follows from (A11) that there exists $\theta^* \in \Theta(Q)$ such that

$$\max_{\gamma \in \mathbf{R}^m} \int \log(1 + \gamma' g(x, \theta^*))\, dQ < \eta. \tag{A12}$$

Further notice that since $\ddot{\mathbf{M}}_0(Q, \theta^*) \neq \emptyset$, by virtue of $\theta^* \in \Theta(Q)$, it follows that

$$s(Q, \theta^*) = m. \tag{A13}$$

Next, we argue that

$$0 \in \operatorname{int}\big(\operatorname{co}\big(\operatorname{supp}_Q(g(X, \theta^*))\big)\big) \tag{A14}$$

(relative to the topology on $\mathbf{R}^m$). If this is not the case, then there exists a $0 \neq \gamma \in \mathbf{R}^m$ such that $\gamma' g(x, \theta^*) \geq 0$ for all $x \in \operatorname{supp}(Q)$. Moreover, it must be the case that $\gamma' g(X, \theta^*) > 0$ with positive probability under Q, for otherwise $\operatorname{supp}_Q(g(X, \theta^*))$ will be contained in an $m - 1$ dimensional subspace of $\mathbf{R}^m$, which contradicts (A13). For such γ, we have for scalar α,

$$\lim_{\alpha \to \infty} \int \log(1 + \alpha \gamma' g(x, \theta^*))\, dQ = \infty,$$

which contradicts (A12), so (A14) is thus established.

We now show $\ddot{\mathbf{M}}_0(Q_n, \theta^*) \neq \emptyset$ for n sufficiently large. It follows from (A14) that there exists a collection of points $\{g_1, \ldots, g_{s(Q,\theta^*)+1}\}$ in $\operatorname{supp}_Q(g(X, \theta^*))$ such that

$$0 \in \operatorname{int}\big(\operatorname{co}\big(\{g_1, \ldots, g_{s(Q,\theta^*)+1}\}\big)\big) \tag{A15}$$

(relative to the topology on $\mathbf{R}^m$). For $1 \leq i \leq s(Q, \theta^*) + 1$, let B_i be an open neighborhood of g_i so small that any collection of points $\{\tilde{g}_1, \ldots, \tilde{g}_{s(Q,\theta^*)+1}\}$ with $\tilde{g}_i \in \bar{B}_i$ for $1 \leq i \leq s(Q, \theta^*) + 1$ will also satisfy (A15) with $\tilde{g}_i$ in place of g_i. For $1 \leq i \leq s(Q, \theta^*) + 1$, let

$$B_i^{-1} = \{x \in \mathcal{X} : g(x, \theta^*) \in B_i\}. \tag{A16}$$

Since each B_i is open and $g(x, \theta^*)$ is continuous, each B_i^{-1} is also open. Moreover, since each B_i is an open neighborhood of a point in the support of $g(X, \theta^*)$ under Q,

$$Q\{X \in B_i^{-1}\} = Q\{g(X, \theta^*) \in B_i\} > 0. \tag{A17}$$

By the portmanteau lemma, we therefore have that for all n sufficiently large,

$$Q_n\{X \in B_i^{-1}\} > 0 \tag{A18}$$

for all $1 \leq i \leq s(Q, \theta^*) + 1$. Thus, for n sufficiently large, (A14) holds with Q_n in place of Q. It follows that $\ddot{\mathbf{M}}_0(Q_n, \theta^*) \neq \emptyset$ for n sufficiently large. Hence, the primal constraint qualification of Theorem 3.4 of Borwein and Lewis (1993) is satisfied for such values of n, from which it follows that

$$\inf_{P \in \ddot{\mathbf{M}}_0(Q_n, \theta^*)} I(Q|P) = \max_{\gamma \in \mathbf{R}^m} \int \log(1 + \gamma' g(x, \theta^*))\, dQ_n. \tag{A19}$$

Next let

$$\gamma_n^* \in \arg\max_{\gamma \in \mathbf{R}^m} \int \log(1 + \gamma' g(x, \theta^*))\, dQ_n. \tag{A20}$$

We now argue that the $\{\gamma_n^*\}_{n=1}^{\infty}$ are uniformly bounded. If this were not the case, then for each $M > 0$, there would exist a subsequence $\{\gamma_{n_k}^*\}_{k=1}^{\infty}$ for which $|\gamma_{n_k}^*| > M$ for all k. By Lemma A.1, there is then an $\varepsilon > 0$ and $j(\gamma_{n_k}^*) \in \{1, \dots, s(Q, \theta^*) + 1\}$ such that

$$\gamma_{n_k}^{*\prime} g < 0 \quad \text{and} \quad \left|\gamma_{n_k}^{*\prime} g\right| \geq |g| \left|\gamma_{n_k}^*\right| \varepsilon \tag{A21}$$

for all $g \in B_{j(\gamma_{n_k}^*)}$. There exists a further subsequence $\{\gamma_{n_{k_\ell}}^*\}_{\ell=1}^{\infty}$ along which $j(\gamma_{n_{k_\ell}}^*)$ is constant. Let $j^* = j(\gamma_{n_{k_\ell}}^*)$. For x such that $g(x, \theta^*) \in B_{j^*}$, we have from (A21) that

$$\gamma_{n_{k_\ell}}^{*\prime} g(x, \theta^*) < 0 \quad \text{and} \quad \left|\gamma_{n_{k_\ell}}^{*\prime} g(x, \theta^*)\right| \geq |g(x, \theta^*)| \left|\gamma_{n_{k_\ell}}^*\right| \varepsilon. \tag{A22}$$

Moreover, we also have from (A20) that

$$Q_{n_{k_\ell}} \left\{1 + \gamma_{n_{k_\ell}}^{*\prime} g(X, \theta^*) > 0\right\} = 1,$$

which, together with (A22), implies that

$$Q_{n_{k_\ell}} \left\{g(X, \theta^*) \in B_{j^*}, |g(X, \theta^*)| \left|\gamma_{n_{k_\ell}}^*\right| \varepsilon > 1\right\} = 0. \tag{A23}$$

Hence, we deduce from $|\gamma_{n_{k_\ell}}^*| > M$ for all ℓ and result (A23) that

$$\begin{aligned} &Q_{n_{k_\ell}} \left\{g(X, \theta^*) \in B_{j^*}, |g(X, \theta^*)| > \frac{1}{\varepsilon M}\right\} \\ &\quad \leq Q_{n_{k_\ell}} \left\{g(X, \theta^*) \in B_{j^*}, |g(X, \theta^*)| \left|\gamma_{n_{k_\ell}}^*\right| \varepsilon > 1\right\} = 0. \end{aligned}$$

Thus, by the portmanteau lemma, we further conclude

$$Q\left\{g(X, \theta^*) \in B_{j^*}, |g(X, \theta^*)| > \frac{1}{\varepsilon M}\right\} = 0,$$

which implies that

$$Q\{g(X,\theta^*)\in B_{j^*}\} = Q\left\{g(X,\theta^*)\in B_{j^*}, |g(X,\theta^*)| \le \frac{1}{\varepsilon M}\right\}.$$

Letting $M\to\infty$, we conclude from (A17) that $0\in B_{j^*}$, which contradicts the requirement that any collection of points $\{\tilde g_1,\dots,\tilde g_{s(Q,\theta^*)+1}\}$ with $\tilde g_i\in \bar B_i$ for $1\le i\le s(Q,\theta^*)+1$ must satisfy (A15) with $\tilde g_i$ in place of g_i. Hence, it must be the case that the sequence $\{\gamma_n^*\}_{n=1}^\infty$ is uniformly bounded.

Therefore, there exists a subsequence $\{\gamma_{n_k}^*\}_{k=1}^\infty$ such that $\gamma_{n_k}^*\to\gamma^*$ and $\gamma^*\in\mathbf{R}^m$. We now show that

$$\text{(A24)}\qquad Q\{1+\gamma^{*\prime}g(X,\theta^*)>0\}=1.$$

To this end, for $\delta>0$, let

$$\text{(A25)}\qquad \begin{aligned} R_\delta^- &= \{x\in\mathcal{X}: 1+\gamma^{*\prime}g(x,\theta^*)<\delta\},\\ R_\delta^+ &= \{x\in\mathcal{X}: 1+\gamma^{*\prime}g(x,\theta^*)\ge\delta\},\end{aligned}$$

and note that

$$\text{(A26)}\qquad \begin{aligned} &\int \log\bigl(1+\gamma_{n_k}^{*\prime}g(x,\theta^*)\bigr)\,dQ_{n_k}\\ &\quad= \int_{R_\delta^-}\log\bigl(1+\gamma_{n_k}^{*\prime}g(x,\theta^*)\bigr)\,dQ_{n_k} + \int_{R_\delta^+}\log\bigl(1+\gamma_{n_k}^{*\prime}g(x,\theta^*)\bigr)\,dQ_{n_k}\\ &\quad\le \int_{R_\delta^-}\log\bigl(1+\gamma_{n_k}^{*\prime}g(x,\theta^*)\bigr)\,dQ_{n_k} + \int_{R_\delta^+}\log\bigl(1+\bigl|\gamma_{n_k}^*\bigr||g(x,\theta^*)|\bigr)\,dQ_{n_k}\\ &\quad\le \int_{R_\delta^-}\log\bigl(1+\gamma_{n_k}^{*\prime}g(x,\theta^*)\bigr)\,dQ_{n_k} + \log\Bigl(1+M\max_{x\in\mathcal{X}}|g(x,\theta^*)|\Bigr),\end{aligned}$$

where the equality holds by inspection, the first inequality holds by the Cauchy–Schwarz inequality, and the second inequality holds because $|\gamma_{n_k}^*|\le M$ for all k and some M. Since $1+\gamma_{n_k}^{*\prime}g(x,\theta^*)\to 1+\gamma^{*\prime}g(x,\theta^*)$ uniformly for $x\in\mathcal{X}$, we have that for k sufficiently large, the integrand in the first term in (A26) is bounded above by $\log(2\delta)$. Thus, for k sufficiently large, (A26) is bounded above by

$$\text{(A27)}\qquad Q_{n_k}\{X\in R_\delta^-\}\log(2\delta)+\log\Bigl(1+M\max_{x\in\mathcal{X}}|g(x,\theta^*)|\Bigr).$$

But

$$\text{(A28)}\qquad \liminf_{n_k\to\infty} Q_{n_k}\{X\in R_\delta^-\}\ge Q\{X\in R_\delta^-\}\ge Q\{X\in\overline{R_0^-}\},$$

where the first inequality follows from the portmanteau lemma and the second inequality follows from the fact that $\overline{R_0^-} \subseteq R_\delta^-$ for all $\delta > 0$. If (A24) fails, then from (A28) we have that

$$\inf_{\delta>0} \liminf_{n_k\to\infty} Q_{n_k}\{X \in R_\delta^-\} > 0.$$

It now follows from (A26), (A27), and (A28) that for k sufficiently large,

$$\int \log\bigl(1+\gamma_{n_k}^{*\prime} g(x,\theta^*)\bigr)\, dQ_{n_k} < 0$$

for $\delta > 0$ sufficiently small, which contradicts (A20). Hence, (A24) is established.

We complete the proof by establishing

$$\int \log(1+\gamma^{*\prime} g(x,\theta^*))\, dQ \geq \eta, \tag{A29}$$

which is a contradiction to (A12). To this end, note that

$$\int \max\bigl\{\log(1+\gamma^{*\prime} g(x,\theta^*)), \log(\delta)\bigr\}\, dQ \tag{A30}$$

$$= \int \max\bigl\{\log(1+\gamma^{*\prime} g(x,\theta^*)), \log(\delta)\bigr\}\bigl(dQ - dQ_{n_k}\bigr) \tag{A31}$$

$$+ \int \max\bigl\{0, \log(\delta) - \log\bigl(1+\gamma_{n_k}^{*\prime} g(x,\theta^*)\bigr)\bigr\}\, dQ_{n_k} \tag{A32}$$

$$\begin{aligned} &+ \int \bigl(\max\bigl\{\log(1+\gamma^{*\prime} g(x,\theta^*)), \log(\delta)\bigr\} \\ &\quad - \max\bigl\{\log\bigl(1+\gamma_{n_k}^{*\prime} g(x,\theta^*)\bigr), \log(\delta)\bigr\}\bigr)\, dQ_{n_k} \end{aligned} \tag{A33}$$

$$+ \int \log\bigl(1+\gamma_{n_k}^{*\prime} g(x,\theta^*)\bigr)\, dQ_{n_k}. \tag{A34}$$

By virtue of $Q_n \to Q$, (A31) tends to zero, while (A32) is nonnegative. Since

$$\begin{aligned} &\max\bigl\{\log\bigl(1+\gamma_{n_k}^{*\prime} g(x,\theta^*)\bigr), \log(\delta)\bigr\} \\ &\quad \to \max\bigl\{\log(1+\gamma^{*\prime} g(x,\theta^*)), \log(\delta)\bigr\} \end{aligned}$$

uniformly on $x \in \mathcal{X}$, (A33) tends to zero. Finally, because of (A19), (A20), and the fact that $Q_{n_k} \in \ddot{\Lambda}_1(\eta)$ for all k, (A34) is weakly greater than η. Thus, (A30) is weakly greater than η. By letting $\delta \searrow 0$, we see from (A24) and the monotone convergence theorem that (A29) holds, which contradicts (A12) as desired. *Q.E.D.*

LEMMA A.3: *Suppose X_i, $i = 1, \dots, n$, is an i.i.d. sequence of random variables with distribution P on $\mathcal{X}$. Suppose further that Assumptions* 3.1, 3.2, *and* 3.4 *hold, that $P \in \mathbf{P}_0$, and that there exists $\omega > 0$ such that all $Q \ll P$ with $P\{X \in \operatorname{supp}(Q)\} > \exp(-\omega)$ satisfy $\ddot{\mathbf{M}}_0(Q, \theta_0(P)) \neq \emptyset$, where*

$$\ddot{\mathbf{M}}_0(Q, \theta) = \bigl\{P \in \mathbf{M} : P \ll Q, Q \ll P, s(Q, \theta) = m, E_P[g(X, \theta)] = 0\bigr\}. \tag{A35}$$

Also let

$$\Gamma(\eta, P) = \bigl\{\gamma \in \mathbf{R}^m : e^{-\eta} \le P\bigl\{1 + \gamma' g(X, \theta_0(P)) \ge 0\bigr\} \le 1\bigr\}. \tag{A36}$$

If $\eta < \omega$ and η satisfies

$$\begin{aligned} &\inf_{\gamma \in \Gamma(\eta, P)} \sup_{\lambda_0, \lambda_1 \ge 0} \lambda_0 + \eta \lambda_1 \\ &\qquad - \exp(\lambda_0 - 1) \int \exp\bigl(\lambda_1 \log\bigl(1 + \gamma' g(x, \theta_0(P))\bigr)\bigr) \\ &\qquad \times I\bigl\{1 + \gamma' g(x, \theta_0(P)) > 0\bigr\}\, dP \\ &\quad \ge \eta, \end{aligned} \tag{A37}$$

then it follows that

$$\limsup_{n \to \infty} \frac{1}{n} \log P^n\{\hat{P}_n \in \Lambda_1(\eta)\} \le -\eta.$$

PROOF: Let $\ddot{\Lambda}_1(\eta, P) = \{Q \in \mathbf{M} : \inf_{R \in \ddot{\mathbf{M}}_0(Q, \theta_0(P))} I(Q|R) \ge \eta\}$ and note that

$$\begin{aligned} \limsup_{n \to \infty} \frac{1}{n} \log P^n\{\hat{P}_n \in \ddot{\Lambda}_1(\eta)\} &\le -\inf_{Q \in \ddot{\Lambda}_1(\eta)} I(Q|P) \\ &\le -\inf_{Q \in \ddot{\Lambda}_1(\eta, P)} I(Q|P), \end{aligned}$$

where the first inequality follows from Lemma A.2 and Sanov's theorem (see Theorem 6.2.10 in Dembo and Zeitouni (1998)), while the final inequality follows from $\ddot{\Lambda}_1(\eta) \subseteq \ddot{\Lambda}_1(\eta, P)$. To complete the proof, it therefore suffices to show that if $\eta < \omega$ satisfies (A37), then

$$\inf_{Q \in \ddot{\Lambda}_1(\eta, P)} I(Q|P) \ge \eta. \tag{A38}$$

For $\mathcal{S} \subseteq \operatorname{supp}(P)$, let $\mathbf{M}(\mathcal{S}) = \{Q \in \mathbf{M} : \operatorname{supp}(Q) \subseteq \mathcal{S}\}$. Note that

$$\inf_{Q \in \mathbf{M}(\mathcal{S})} I(Q|P) = -\log(P\{X \in \mathcal{S}\}). \tag{A39}$$

To see this, observe that the left hand side of (A39) is greater than or equal to the right hand side of (A39) by Jensen's inequality and that $I(Q|P) = -\log(P\{X \in \mathcal{S}\})$ for Q given by the distribution P conditional on $\mathcal{S}$. Next note that for any Q such that $P\{X \in \operatorname{supp}(Q)\} \leq \exp(-\eta)$, we have that

$$I(Q|P) \geq \inf_{\mu \in \mathbf{M}(\operatorname{supp}(Q))} I(\mu|P) = -\log\bigl(P\{X \in \operatorname{supp}(Q)\}\bigr) \geq \eta. \tag{A40}$$

Note further that if Q is not dominated by P, then $I(Q|P) = +\infty$. Hence, for

$$\tilde{\Lambda}_1(\eta, P) = \bigl\{Q \in \ddot{\Lambda}_1(\eta, P) : Q \ll P, P\{X \in \operatorname{supp}(Q)\} \geq \exp(-\eta)\bigr\},$$

we have that

$$\inf_{Q \in \ddot{\Lambda}_1(\eta, P)} I(Q|P) \geq \min\Bigl\{\inf_{Q \in \tilde{\Lambda}_1(\eta, P)} I(Q|P), \eta\Bigr\}. \tag{A41}$$

We may assume that $\tilde{\Lambda}_1(\eta, P) \neq \emptyset$, for otherwise the right hand side of (A41) equals η, thus establishing (A38). Furthermore, since $\eta < \omega$, we also have by assumption that for any $Q \in \tilde{\Lambda}_1(\eta, P)$, $\ddot{\mathbf{M}}_0(Q, \theta_0(P)) \neq \emptyset$. Hence, the primal constraint qualification of Theorem 3.4 of Borwein and Lewis (1993) is satisfied, so that for all $Q \in \tilde{\Lambda}_1(\eta, P)$, we have

$$\inf_{R \in \ddot{\mathbf{M}}_0(Q, \theta_0(P))} I(R|Q) = \max_{\gamma \in \mathbf{R}^m} \int \log\bigl(1 + \gamma' g(x, \theta_0(P))\bigr)\, dQ \geq \eta,$$

where the inequality is implied by $Q \in \ddot{\Lambda}_1(\eta, P)$. Next we define

$$\begin{aligned}
\Gamma = \Bigl\{&\gamma \in \mathbf{R}^m : \exists Q \in \tilde{\Lambda}_1(\eta, P) \\
&\text{s.t. } \gamma \in \arg\max_{\lambda \in \mathbf{R}^m} \int \log\bigl(1 + \lambda' g(x, \theta_0(P))\bigr)\, dQ\Bigr\}, \\
\mathbf{S}(\gamma) = \Bigl\{&\mathcal{S} \subseteq \operatorname{supp}(P) : \exists Q \in \tilde{\Lambda}_1(\eta, P) \text{ s.t. } \mathcal{S} = \operatorname{supp}(Q), \\
&\gamma \in \arg\max_{\lambda \in \mathbf{R}^m} \int \log\bigl(1 + \lambda' g(x, \theta_0(P))\bigr)\, dQ\Bigr\}, \\
\mathbf{R}(\gamma, \mathcal{S}) = \Bigl\{&Q \in \tilde{\Lambda}_1(\eta, P) : \gamma \in \arg\max_{\lambda \in \mathbf{R}^m} \int \log\bigl(1 + \lambda' g(x, \theta_0(P))\bigr)\, dQ, \\
&\mathcal{S} = \operatorname{supp}(Q)\Bigr\}.
\end{aligned}$$

With these definitions, we write

$$\tilde{\Lambda}_1(\eta, P) = \bigcup_{\gamma \in \Gamma} \bigcup_{\mathcal{S} \in \mathbf{S}(\gamma)} \mathbf{R}(\gamma, \mathcal{S}).$$

Hence,

$$\inf_{Q \in \tilde{\Lambda}_1(\eta, P)} I(Q|P) = \inf_{\gamma \in \Gamma} \inf_{\mathcal{S} \in \mathbf{S}(\gamma)} \inf_{Q \in \mathbf{R}(\gamma, \mathcal{S})} I(Q|P). \tag{A42}$$

Note that if $Q \in \mathbf{R}(\gamma, \mathcal{S})$, then (i) $Q \ll P$, (ii) $\mathcal{S} = \operatorname{supp}(Q)$, and (iii) $\int \log(1 + \gamma' g(x, \theta_0(P)))\, dQ \geq \eta$. We therefore have for $\delta > 0$ sufficiently small, that

$$\begin{aligned} &\inf_{Q \in \mathbf{R}(\gamma, \mathcal{S})} I(Q|P) \\ &\quad \geq \inf\biggl\{ \int_{\mathcal{S}} \log(\phi(x))\phi(x)\, dP : \phi \in L^1(\mathcal{S}), \phi > 0, \\ &\qquad \int_{\mathcal{S}} \log\bigl(1 + \gamma' g(x, \theta_0(P))\bigr)\phi(x)\, dP \geq \eta, \int_{\mathcal{S}} \phi(x)\, dP = 1 \biggr\} \\ &\quad \geq \inf\biggl\{ \int_{\mathcal{S}} \log(\phi(x))\phi(x)\, dP : \phi \in L^1(\mathcal{S}), \phi > 0, \\ &\qquad \int_{\mathcal{S}} \log\bigl(1 + \gamma' g(x, \theta_0(P))\bigr) I\{x \in R_\delta^+\}\phi(x)\, dP \geq \eta, \\ &\qquad \int_{\mathcal{S}} \phi(x)\, dP = 1 \biggr\}, \end{aligned} \tag{A43}$$

where the first inequality follows from the preceding statements (i), (ii), and (iii), and the second inequality follows from the definition of R_δ^+ in (A25) but with $(\theta_0(P), \gamma)$ in place of (θ^*, γ^*).

We now use Corollary 4.8 of Borwein and Lewis (1992a) and part (vi) of Example 6.5 of Borwein and Lewis (1992b) to find the dual problem of (A43). To this end, first note that since $\tilde{\Lambda}_1(\eta, P) \neq \emptyset$, we have that $\mathbf{R}(\gamma, \mathcal{S}) \neq \emptyset$ for at least one $\gamma \in \Gamma$ and $\mathcal{S} \in \mathbf{S}(\gamma)$. For any such γ and $\mathcal{S}$, we have as a result that there exists a ϕ satisfying the constraints of (A43). Next note that the map $A : L^1(\mathcal{S}) \to \mathbf{R}$ defined by

$$A(\phi) = \int_{\mathcal{S}} \log\bigl(1 + \gamma' g(x, \theta_0(P))\bigr) I\{x \in R_\delta^+\}\phi(x)\, dP$$

is continuous because $\log(1 + \gamma' g(x, \theta_0(P)))I\{x \in R_\delta^+\}$ lies in $L^\infty(\mathcal{S})$ as a result of $\mathcal{S}$ being a subset of the compact set $\mathcal{X}$ and $g(x, \theta_0(P))$ being continuous on $\mathcal{X}$. Using Corollary 4.8 of Borwein and Lewis (1992a) and part (vi) of Example 6.5 of Borwein and Lewis (1992b) to find the dual problem of (A43) implies

$$\inf_{Q\in\mathbf{R}(\gamma,\mathcal{S})} I(Q|P) \geq \liminf_{\delta\searrow 0} \sup_{\lambda_0,\lambda_1\geq 0} \lambda_0+\eta\lambda_1 - \exp(\lambda_0-1)\int_{\mathcal{S}} \exp\bigl(\lambda_1\log\bigl(1+\gamma' g(x,\theta_0(P))\bigr) I\{x\in R_\delta^+\}\bigr)\,dP. \tag{A44}$$

By definition, for every $\mathcal{S}\in\mathbf{S}(\gamma)$, there exists Q such that $\mathcal{S}=\operatorname{supp}(Q)$ and

$$\gamma\in\arg\max_{\lambda\in\mathbf{R}^m}\int\log\bigl(1+\lambda' g(x,\theta_0(P))\bigr)\,dQ. \tag{A45}$$

For any such Q, we must have that

$$Q\bigl\{1+\gamma' g(X,\theta_0(P))\leq 0\bigr\}=Q\bigl\{1+\gamma' g(X,\theta_0(P))\leq 0, X\in\mathcal{S}\bigr\}=0, \tag{A46}$$

from which it follows that

$$P\bigl\{1+\gamma' g(X,\theta_0(P))\leq 0, X\in\mathcal{S}\bigr\}=0 \tag{A47}$$

as well. Hence, by letting $\delta\searrow 0$, we see by the monotone convergence theorem that

$$\int_{\mathcal{S}} \exp\bigl(\lambda_1\log\bigl(1+\gamma' g(x,\theta_0(P))\bigr) I\{x\in R_\delta^+\}\bigr)\,dP$$

is right-continuous at $\delta=0$. Following arguments as in Lemma 17.29 in Aliprantis and Border (2006), it is possible to show that the supremum in (A44) is lower semicontinuous at $\delta=0$ as well. Hence, the right hand side of (A44) is greater than or equal to

$$\sup_{\lambda_0,\lambda_1\geq 0} \lambda_0+\eta\lambda_1-\exp(\lambda_0-1) \times\int_{\mathcal{S}} \exp\bigl(\lambda_1\log\bigl(1+\gamma' g(x,\theta_0(P))\bigr) I\bigl\{1+\gamma' g(x,\theta_0(P))\geq 0\bigr\}\bigr)\,dP. \tag{A48}$$

Since the integrand in (A48) is nonnegative, we have from (A44) and (A48) that

$$\inf_{\mathcal{S}\in\mathbf{S}(\gamma)}\inf_{Q\in\mathbf{R}(\gamma,\mathcal{S})} I(Q|P) \geq \inf_{\mathcal{S}\in\mathbf{S}(\gamma)}\sup_{\lambda_0,\lambda_1\geq 0} \lambda_0+\eta\lambda_1 \tag{A49}$$

$$
\begin{aligned}
&\quad - \exp(\lambda_0 - 1) \int_{\mathcal{S}} \exp\bigl(\lambda_1 \log\bigl(1 + \gamma' g(x, \theta_0(P))\bigr)\bigr) \\
&\quad \times I\bigl\{1 + \gamma' g(x, \theta_0(P)) > 0\bigr\}\, dP \\
&\ge \sup_{\lambda_0, \lambda_1 \ge 0} \lambda_0 + \eta \lambda_1 \\
&\quad - \exp(\lambda_0 - 1) \int \exp\bigl(\lambda_1 \log\bigl(1 + \gamma' g(x, \theta_0(P))\bigr)\bigr) \\
&\quad \times I\bigl\{1 + \gamma' g(x, \theta_0(P)) > 0\bigr\}\, dP.
\end{aligned}
$$

By definition, for every $\gamma \in \Gamma$, there exists a $Q \in \tilde{\Lambda}_1(\eta, P)$ such that γ satisfies (A45). Thus, as before, (A46) holds, from which it follows that

$$
\operatorname{supp}(Q) \subseteq \bigl\{x \in \mathbf{R}^d : 1 + \gamma' g(x, \theta_0(P)) \ge 0\bigr\}.
$$

Therefore,

$$
P\bigl\{1 + \gamma' g(X, \theta_0(P)) \ge 0\bigr\} \ge P\{X \in \operatorname{supp}(Q)\} \ge \exp(-\eta)
$$

by $Q \in \tilde{\Lambda}_1(\eta, P)$. Hence, $\gamma \in \Gamma(\eta, P)$, which implies $\Gamma \subseteq \Gamma(\eta, P)$. It therefore follows from (A49) that

$$
\begin{aligned}
&\inf_{\gamma \in \Gamma} \inf_{\mathcal{S} \in \mathbf{S}(\gamma)} \inf_{Q \in \mathbf{R}(\gamma, \mathcal{S})} I(Q|P) \\
&\quad \ge \inf_{\gamma \in \Gamma(\eta, P)} \sup_{\lambda_0, \lambda_1 \ge 0} \lambda_0 + \eta \lambda_1 \\
&\quad - \exp(\lambda_0 - 1) \int \exp\bigl(\lambda_1 \log\bigl(1 + \gamma' g(x, \theta_0(P))\bigr)\bigr) \\
&\quad \times I\bigl\{1 + \gamma' g(x, \theta_0(P)) > 0\bigr\}\, dP.
\end{aligned}
$$

The desired claim (A38) thus follows for $\eta < \omega$, satisfying (A37). *Q.E.D.*

LEMMA A.4: *If Assumptions* 3.1, 3.2, 3.3, *and* 3.4 *hold*, *then* $\mathbf{P}_0$ *is closed under the weak topology and*, *in addition*, $\theta_0(P)$ *is continuous on* $\mathbf{P}_0$ *under the weak topology*.

PROOF: Let $P_n \to P$ with $P_n \in \mathbf{P}_0$ for all n and denote

$$
\theta_n = \theta_0(P_n),
$$

where $\theta_0(P_n)$ is a singleton by $P_n \in \mathbf{P}_0$ and Assumption 3.4. Let θ^* be a limit point of $\{\theta_n\}_{n=1}^{\infty}$ and let $\{\theta_{n_k}\}_{k=1}^{\infty}$ be a subsequence such that $\theta_{n_k} \to \theta^*$. It then follows that

$$\text{(A50)} \quad \begin{aligned} \left|\int g(x,\theta^*)\,dP\right| &= \lim_{n_k\to\infty}\left|\int g(x,\theta^*)\,dP_{n_k}\right| \\ &= \lim_{n_k\to\infty}\left|\int \big(g(x,\theta^*) - g\big(x,\theta_{n_k}\big)\big)\,dP_{n_k}\right| \\ &\le \lim_{n_k\to\infty}\sup_{x\in\mathcal{X}}\big|g(x,\theta^*) - g\big(x,\theta_{n_k}\big)\big| \\ &= 0, \end{aligned}$$

where the first equality follows by $P_n \to P$ and $g(x, \theta^*)$ continuous and bounded. The second equality is implied by $\theta_{n_k} = \theta_0(P_{n_k})$, the inequality follows by inspection, and the final result is due to the uniform continuity of $g(x, \theta)$. However, since $\mathbf{P}$ is closed, it follows that $P \in \mathbf{P}$ and, therefore, (A50) implies $P \in \mathbf{P}_0$ which establishes that $\mathbf{P}_0$ is closed as claimed. Moreover, we also conclude from (A50) and $P \in \mathbf{P}_0$ that

$$\text{(A51)} \quad \theta^* = \theta_0(P).$$

Therefore, $\theta_0(P)$ is the unique limit point of $\{\theta_n\}_{n=1}^{\infty}$, which establishes the continuity of $\theta_0(P)$. *Q.E.D.*

LEMMA A.5: *If Assumptions* 3.1–3.5 *hold, then there exists a* $\varsigma > 0$ *such that*

$$\text{(A52)} \quad \sup_{P\in\mathbf{P}_0}\sup_{\gamma\neq 0} P\big\{\gamma' g(X, \theta_0(P)) \ge 0\big\} \le 1 - \varsigma.$$

PROOF: We establish the claim by contradiction. Suppose that contrary to (A52), we have

$$\text{(A53)} \quad \sup_{P\in\mathbf{P}_0}\sup_{|\gamma|=1} P\big\{\gamma' g(X, \theta_0(P)) \ge 0\big\} = 1.$$

By Lemma A.4, $\mathbf{P}_0 \subset \mathbf{M}$ is closed in the weak topology. Therefore, since $\mathbf{M}$ is compact in the weak topology by Theorem 15.11 in Aliprantis and Border (2006), it follows that $\mathbf{P}_0$ is compact as well. Letting $\mathbf{S}^m$ denote the unit sphere on $\mathbf{R}^m$, we then obtain that $\mathbf{P}_0 \times \mathbf{S}^m$ is compact and hence by (A53), there exists a sequence $\{(P_n, \gamma_n)\}_{n=1}^{\infty}$ satisfying $(P_n, \gamma_n) \in \mathbf{P}_0 \times \mathbf{S}^m$ for all n, $(P_n, \gamma_n) \to (P^*, \gamma^*) \in \mathbf{P}_0 \times \mathbf{S}^m$, and

$$\text{(A54)} \quad \lim_{n\to\infty} P_n\big\{\gamma_n' g(X, \theta_0(P_n)) \ge 0\big\} = 1.$$

Defining the sets $A_n^+ = \{x \in \mathcal{X} : \gamma_n' g(x, \theta(P_n)) > 0\}$ and $A_n^- = \{x \in \mathcal{X} : \gamma_n' g(x, \theta(P_n)) < 0\}$, we then obtain from (A54), $\int g(x, \theta_0(P_n))\,dP_n = 0$, and $g(x, \theta)$

bounded on $\mathcal{X} \times \Theta$ that

$$\begin{aligned}
\text{(A55)} \quad & \limsup_{n\to\infty} \int \left|\gamma_n' g(x, \theta_0(P_n))\right| dP_n \\
&= \limsup_{n\to\infty} \left\{ \int_{A_n^+} \gamma_n' g(x, \theta_0(P_n))\, dP_n - \int_{A_n^-} \gamma_n' g(x, \theta_0(P_n))\, dP_n \right\} \\
&= \limsup_{n\to\infty} \int_{A_n^-} 2\left|\gamma_n' g(x, \theta_0(P_n))\right| dP_n \\
&\le \limsup_{n\to\infty} \sup_{x\in\mathcal{X},\theta\in\Theta} 2|g(x,\theta)| \times P_n\{A_n^-\} \\
&= 0.
\end{aligned}$$

Since $(P_n, \gamma_n) \to (P^*, \gamma^*)$, Lemma A.4 and compactness imply $\sup_{x\in\mathcal{X}} |\gamma_n' g(x, \theta_0(P_n)) - \gamma^{*\prime} g(x, \theta_0(P^*))| \to 0$. Hence, (A55), $P_n \to P^*$, and $g(x, \theta_0(P^*))$ continuous and bounded yield

$$\begin{aligned}
\text{(A56)} \quad & \int \left|\gamma^{*\prime} g(x, \theta_0(P^*))\right| dP^* \\
&\le \limsup_{n\to\infty} \int \left|\gamma^{*\prime} g(x, \theta_0(P^*))\right| (dP^* - dP_n) \\
&\quad + \limsup_{n\to\infty} \int \left|\gamma^{*\prime} g(x, \theta_0(P^*)) - \gamma_n' g(x, \theta_0(P_n))\right| dP_n \\
&\quad + \limsup_{n\to\infty} \int \left|\gamma_n' g(x, \theta_0(P_n))\right| dP_n \\
&= 0.
\end{aligned}$$

It follows from (A56) that $P^* \in \mathbf{D}_0$, which contradicts $P^* \in \mathbf{P}_0$ and Assumption 3.5. *Q.E.D.*

LEMMA A.6: *If Assumptions* 3.1–3.5 *hold and* ς *is as in Lemma* A.5, *then for any* δ *such that* $0 < \delta < \varsigma$,

$$\text{(A57)} \quad \Gamma(\eta, P) = \left\{\gamma \in \mathbf{R}^m : e^{-\eta} \le P\left\{1 + \gamma' g(X, \theta_0(P)) \ge 0\right\} \le 1\right\}$$

is nonempty, compact valued, and upper hemicontinuous on $(\eta, P) \in [0, -\log(1 - \varsigma + \delta)] \times \mathbf{P}_0$ *under the product of the topology on* $\mathbf{R}$ *and the weak topology.*

PROOF: The correspondence $\Gamma(\eta, P)$ is clearly not empty since $0 \in \Gamma(\eta, P)$ for all $(\eta, P) \in [0, -\log(1 - \varsigma + \delta)] \times \mathbf{P}_0$. To establish upper hemicontinuity, we wish to show that if $P_n \to P$ and $\eta_n \to \eta$ with $(\eta_n, P_n) \in [0, -\log(1 - \varsigma +$

$\delta)] \times \mathbf{P}_0$ for all n, then any sequence $\{\gamma_n\}_{n=1}^{\infty}$ with $\gamma_n \in \Gamma(\eta_n, P_n)$ for all n has a limit point in $\Gamma(\eta, P)$. For this purpose, we first show that the sequence $\{\gamma_n\}_{n=1}^{\infty}$ is uniformly bounded. Suppose by way of contradiction that

$$\limsup_{n\to\infty} |\gamma_n| = +\infty. \tag{A58}$$

It follows that there exists a subsequence $\{\gamma_{n_k}\}_{k=1}^{\infty}$ satisfying, for all k,

$$|\gamma_{n_k}| \geq n_k. \tag{A59}$$

In addition, by compactness there exists an additional subsequence $\{\gamma_{n_{k_l}}\}_{l=1}^{\infty}$ such that

$$\frac{\gamma_{n_{k_l}}}{|\gamma_{n_{k_l}}|} \to \gamma_1. \tag{A60}$$

Along such a subsequence, however, we have

$$\begin{aligned}
e^{-\eta} &= \lim_{n_{k_l}\to\infty} e^{-\eta_{n_{k_l}}} \\
&\leq \limsup_{n_{k_l}\to\infty} P_{n_{k_l}}\{1 + \gamma'_{n_{k_l}} g(X, \theta_0(P_{n_{k_l}})) \geq 0\} \\
&= \limsup_{n_{k_l}\to\infty} P_{n_{k_l}}\left\{\frac{\gamma'_{n_{k_l}}}{|\gamma_{n_{k_l}}|} g(X, \theta_0(P_{n_{k_l}})) \geq -\frac{1}{|\gamma_{n_{k_l}}|}\right\} \\
&\leq \liminf_{\varepsilon\searrow 0} \limsup_{n_{k_l}\to\infty} P_{n_{k_l}}\left\{\frac{\gamma'_{n_{k_l}}}{|\gamma'_{n_{k_l}}|} g(X, \theta_0(P_{n_{k_l}})) \geq -\varepsilon\right\} \\
&\leq \liminf_{\varepsilon\searrow 0} \limsup_{n_{k_l}\to\infty} P_{n_{k_l}}\{\gamma'_1 g(X, \theta_0(P)) \geq -2\varepsilon\} \\
&\leq \liminf_{\varepsilon\searrow 0} P\{\gamma'_1 g(X, \theta_0(P)) \geq -2\varepsilon\} \\
&= P\{\gamma'_1 g(X, \theta_0(P)) \geq 0\},
\end{aligned} \tag{A61}$$

where the first equality follows by assumption and the first inequality follows by $\gamma_{n_{k_l}} \in \Gamma(\eta_{n_{k_l}}, P_{n_{k_l}})$ for all l. The second equality follows by inspection. The second inequality is implied by (A59) and the third inequality is implied by $\theta_0(P_{n_{k_l}}) \to \theta_0(P)$ by Lemma A.4, (A60), and the uniform continuity of $g(x, \theta)$. The final inequality and equality follow by the portmanteau and the bounded convergence theorems, respectively. Hence,

$$1 - \varsigma < e^{-\eta} \leq P\{\gamma'_1 g(X, \theta_0(P)) \geq 0\} \tag{A62}$$

by (A61) and $\eta_{n_{k_l}} \in [0, -\log(1-\varsigma+\delta)]$ for all l. Result (A62), however, contradicts $P \in \mathbf{P}_0$ by Lemma A.5.

Because the sequence $\{\gamma_n\}_{n=1}^{\infty}$ is uniformly bounded, it follows that there exists a subsequence such that

$$\lim_{n_j \to \infty} \gamma_{n_j} = \gamma_2. \tag{A63}$$

To conclude establishing upper hemicontinuity, we show that $\gamma_2 \in \Gamma(\eta, P)$, which is implied by

$$\begin{aligned} e^{-\eta} &= \lim_{n_j \to \infty} e^{-\eta_{n_j}} \\ &\le \limsup_{n_j \to \infty} P_{n_j}\bigl\{1 + \gamma_{n_j}' g\bigl(X, \theta_0\bigl(P_{n_j}\bigr)\bigr) \ge 0\bigr\} \\ &\le \liminf_{\varepsilon \searrow 0} \limsup_{n_j \to \infty} P_{n_j}\bigl\{1 + \gamma_2' g(X, \theta_0(P)) \ge -\varepsilon\bigr\} \\ &\le \liminf_{\varepsilon \searrow 0} P\bigl\{1 + \gamma_2' g(X, \theta_0(P)) \ge -\varepsilon\bigr\} \\ &= P\bigl\{1 + \gamma_2' g(X, \theta_0(P)) \ge 0\bigr\}, \end{aligned} \tag{A64}$$

where the first equality follows by assumption and the first inequality follows by $\gamma_{n_j} \in \Gamma(\eta_{n_j}, P_{n_j})$ for all j. By Lemma A.4, $\theta_0(P_{n_j}) \to \theta_0(P)$ and, therefore, the second inequality follows by the uniform continuity of $g(x, \theta)$. The final inequality and equality follow by the portmanteau and the bounded convergence theorems, respectively.

The arguments in (A58)–(A61), but for $\{\gamma_n\}_{n=1}^{\infty}$ an unbounded sequence in $\Gamma(\eta, P)$ and $\eta_n = \eta$, $P_n = P$ for all n, show that $\Gamma(\eta, P)$ is bounded. Similarly, the arguments in (A64), but with $\eta_n = \eta$ and $P_n = P$ for all n, show that $\Gamma(\eta, P)$ is closed. Hence, $\Gamma(\eta, P)$ is compact. *Q.E.D.*

LEMMA A.7: *If Assumptions* 3.1, 3.2, 3.3, *and* 3.4 *hold*, *then the function*

$$f(\lambda_1, \gamma, P) = \int \bigl(1 + \gamma' g(x, \theta_0(P))\bigr)^{\lambda_1} I\bigl\{1 + \gamma' g(x, \theta_0(P)) > 0\bigr\}\, dP$$

is lower semicontinuous on $(\lambda_1, \gamma, P) \in \mathbf{R}_+ \times \mathbf{R}^m \times \mathbf{P}_0$, *where* $\mathbf{P}_0$ *is endowed with the weak topology*.

PROOF: Let $(\lambda_{1,n}, \gamma_n, P_n) \to (\lambda_1, \gamma, P)$. To establish the lemma, we aim to show that

$$\liminf_{n \to \infty} f(\lambda_{1,n}, \gamma_n, P_n) \ge f(\lambda_1, \gamma, P). \tag{A65}$$

For this purpose, we define the auxiliary variable

$$\varepsilon_m \equiv \sup_{x\in\mathcal{X}, m_0\geq m} \left|\gamma'_{m_0} g\big(x, \theta_0\big(P_{m_0}\big)\big) - \gamma' g(x, \theta_0(P))\right|. \tag{A66}$$

Notice that due to Lemma A.4 and Assumption 3.1, we have $\lim_{m\to\infty} \varepsilon_m = 0$. Also define

$$\begin{aligned} \bar{\lambda}_{1,m} &\equiv \sup_{m_0\geq m} \lambda_{m_0}, \\ \underline{\lambda}_{1,m} &\equiv \inf_{m_0\geq m} \lambda_{m_0} \end{aligned} \tag{A67}$$

as well as the function

$$L_m(u) = u^{\underline{\lambda}_{1,m}} I\{u > 1\} + u^{\bar{\lambda}_{1,m}} I\{0 < u \leq 1\}. \tag{A68}$$

Then notice that pointwise in $x \in \mathcal{X}$, we have that

$$\begin{aligned} &\inf_{m_0\geq m} \big(1 + \gamma'_{m_0} g\big(x, \theta_0\big(P_{m_0}\big)\big)\big)^{\lambda_{m_0}} I\big\{1 + \gamma'_{m_0} g\big(x, \theta_0\big(P_{m_0}\big)\big) > 0\big\} \\ &\quad\geq \inf_{m_0\geq m} \big(1 + \gamma'_{m_0} g\big(x, \theta_0\big(P_{m_0}\big)\big)\big)^{\lambda_{m_0}} I\big\{1 + \gamma' g(x, \theta_0(P)) > \varepsilon_m\big\} \\ &\quad\geq \inf_{m_0\geq m} \big(1 + \gamma' g(x, \theta_0(P)) - \varepsilon_m\big)^{\lambda_{m_0}} I\big\{1 + \gamma' g(x, \theta_0(P)) > \varepsilon_m\big\} \\ &\quad\geq L_m\big(1 + \gamma' g(x, \theta_0(P)) - \varepsilon_m\big), \end{aligned} \tag{A69}$$

where the first two inequalities are implied by (A66), and the final inequality follows by (A68) and direct calculation. Next, exploiting standard manipulations and (A69), we are able to conclude

$$\begin{aligned} &\liminf_{n\to\infty} f(\lambda_{1,n}, \gamma_n, P_n) \\ &\quad= \lim_{n\to\infty} \inf_{n_0\geq n} \int \big(1 + \gamma'_{n_0} g\big(x, \theta_0\big(P_{n_0}\big)\big)\big)^{\lambda_{n_0}} \\ &\qquad \times I\big\{1 + \gamma'_{n_0} g\big(x, \theta_0\big(P_{n_0}\big)\big) > 0\big\}\, dP_{n_0} \\ &\quad\geq \lim_{n\to\infty} \inf_{n_0\geq n} \inf_{m_0\geq m} \int \big(1 + \gamma'_{m_0} g\big(x, \theta_0\big(P_{m_0}\big)\big)\big)^{\lambda_{m_0}} \\ &\qquad \times I\big\{1 + \gamma'_{m_0} g\big(x, \theta_0\big(P_{m_0}\big)\big) > 0\big\}\, dP_{n_0} \\ &\quad\geq \liminf_{m\to\infty} \liminf_{n\to\infty} \inf_{m_0\geq m} \int \big(1 + \gamma'_{m_0} g\big(x, \theta_0\big(P_{m_0}\big)\big)\big)^{\lambda_{m_0}} \\ &\qquad \times I\big\{1 + \gamma'_{m_0} g\big(x, \theta_0\big(P_{m_0}\big)\big) > 0\big\}\, dP_n \end{aligned} \tag{A70}$$

$$
\begin{aligned}
&\geq \liminf_{m\to\infty}\liminf_{n\to\infty}\int \inf_{m_0\geq m}\bigl(1+\gamma'_{m_0}g\bigl(x,\theta_0\bigl(P_{m_0}\bigr)\bigr)\bigr)^{\lambda_{m_0}} \\
&\quad \times I\bigl\{1+\gamma'_{m_0}g\bigl(x,\theta_0\bigl(P_{m_0}\bigr)\bigr)>0\bigr\}\,dP_n \\
&\geq \liminf_{m\to\infty}\liminf_{n\to\infty}\int L_m\bigl(1+\gamma' g(x,\theta_0(P))-\varepsilon_m\bigr)\,dP_n.
\end{aligned}
$$

Further, observe from (A68) that if $\bar{\lambda}_{1,m}>0$, then $L_m(u)$ is continuous, while if $\bar{\lambda}_{1,m}=0$, then we have $L_m(u)=I\{u>0\}$. In both cases, since $g(x,\theta_0(P))$ is continuous and $\mathcal{X}$ is compact, we obtain by the portmanteau lemma and $P_n\to P$ in the weak topology,

$$
\begin{aligned}
\text{(A71)}\quad &\liminf_{m\to\infty}\liminf_{n\to\infty}\int L_m\bigl(1+\gamma' g(x,\theta_0(P))-\varepsilon_m\bigr)\,dP_n \\
&\quad\geq \liminf_{m\to\infty}\int L_m\bigl(1+\gamma' g(x,\theta_0(P))-\varepsilon_m\bigr)\,dP \\
&\quad\geq \int \liminf_{m\to\infty} L_m\bigl(1+\gamma' g(x,\theta_0(P))-\varepsilon_m\bigr)\,dP,
\end{aligned}
$$

where the second inequality follows by Fatou's lemma. Finally, by $\bar{\lambda}_{1,m}\to\lambda_1$, $\underline{\lambda}_{1,m}\to\lambda_1$, and $\varepsilon_m\to 0$, direct calculation reveals that pointwise in $x\in\mathcal{X}$, we have

$$
\begin{aligned}
\text{(A72)}\quad &\liminf_{m\to\infty} L_m\bigl(1+\gamma' g(x,\theta_0(P))-\varepsilon_m\bigr) \\
&\quad\geq \bigl(1+\gamma' g(x,\theta_0(P))\bigr)^{\lambda_1} I\bigl\{1+\gamma' g(x,\theta_0(P))>0\bigr\}.
\end{aligned}
$$

Combining (A70), (A71), and (A72) establishes the claim of the lemma. *Q.E.D.*

LEMMA A.8: *Suppose Assumptions* 3.1, 3.2, 3.3, *and* 3.4 *hold and for* $(\lambda_0,\lambda_1,\eta,\gamma,P)\in[0,2]^2\times\mathbf{R}_+\times\mathbf{R}^m\times\mathbf{P}_0$ *with* $\mathbf{P}_0$ *endowed with the weak topology*, *define the function*

$$
\begin{aligned}
&F(\lambda_0,\lambda_1,\eta,\gamma,P) \\
&\quad=\lambda_0+\eta(\lambda_1-1) \\
&\qquad - e^{\lambda_0-1}\int \bigl(1+\gamma' g(x,\theta_0(P))\bigr)^{\lambda_1} I\bigl\{1+\gamma' g(x,\theta_0(P))>0\bigr\}\,dP.
\end{aligned}
$$

In addition, *consider the correspondences*

$$
\begin{aligned}
&E(\eta,\gamma,P) \\
&\quad=\{(\lambda_0,\lambda_1,y)\in[0,2]^2\times\mathbf{R}: y\leq F(\lambda_0,\lambda_1,\eta,\gamma,P)\},
\end{aligned}
$$

$$\Pi(\eta, \gamma, P) \\ = \{y \in \mathbf{R} \colon (\lambda_0, \lambda_1, y) \in E(\eta, \gamma, P) \textit{ for some } (\lambda_0, \lambda_1) \in [0, 2]^2\}.$$

It then follows that $\Pi(\eta, \gamma, P)$ is lower hemicontinuous on $\mathbf{R}_+ \times \mathbf{R} \times \mathbf{P}_0$.

PROOF: As in Lemma A.7, we define the function

$$f(\lambda_1, \gamma, P) = \int \big(1 + \gamma' g(x, \theta_0(P))\big)^{\lambda_1} I\big\{1 + \gamma' g(x, \theta_0(P)) > 0\big\}\, dP. \tag{A73}$$

We first show that $f(\lambda_1, \gamma, P)$ is continuous at all points on $[0, 2] \times \mathbf{R}^m \times \mathbf{P}_0$ with $\lambda_1 \neq 0$. For this purpose, let $(\lambda_{1,n}, \gamma_n, P_n) \to (\lambda_1, \gamma, P)$, and note that by Lemma A.4 and $\mathcal{X}$ being compact, we have

$$\lim_{n\to\infty} \sup_{x\in\mathcal{X}} \big|\gamma_n' g(x, \theta_0(P_n)) - \gamma' g(x, \theta_0(P))\big| = 0. \tag{A74}$$

Further, notice that since $\lambda_1 > 0$, then by $\lambda_{1,n} \to \lambda_1$, we have $\lambda_{1,n} > 0$ for n large enough, which implies

$$\lim_{n\to\infty} \sup_{x\in\mathcal{X}} \big|\big(1 + \gamma_n' g(x, \theta_0(P_n))\big)^{\lambda_{1,n}} \\ \times I\big\{1 + \gamma_n' g(x, \theta_0(P_n)) > 0, 1 + \gamma' g(x, \theta_0(P)) \leq 0\big\}\big| = 0 \tag{A75}$$

as a result of (A74). By direct calculations, we then obtain from (A74) and (A75) that

$$\lim_{n\to\infty} \sup_{x\in\mathcal{X}} \big|\big(1 + \gamma_n' g(x, \theta_0(P_n))\big)^{\lambda_{1,n}} I\big\{1 + \gamma_n' g(x, \theta_0(P_n)) > 0\big\} \\ - \big(1 + \gamma' g(x, \theta_0(P))\big)^{\lambda_1} I\big\{1 + \gamma' g(x, \theta_0(P)) > 0\big\}\big| = 0. \tag{A76}$$

By (A76) and noting that the integrand is a continuous bounded function for $\lambda_1 > 0$, $P_n \to P$ establishes

$$f(\lambda_{1,n}, \gamma_n, P_n) \\ = \int \big(1 + \gamma' g(x, \theta_0(P))\big)^{\lambda_1} I\big\{1 + \gamma' g(x, \theta_0(P)) > 0\big\}\, dP_n + o(1) \\ \to f(\lambda_1, \gamma, P), \tag{A77}$$

hence proving the desired continuity of $f(\lambda_1, \gamma, P)$ at all points $(\lambda_1, \gamma, P) \in [0, 2] \times \mathbf{R}_+ \times \mathbf{P}_0$ with $\lambda_1 > 0$.

We now establish lower hemicontinuity of $\Pi(\eta, \gamma, P)$. This requires showing that for any $y \in \Pi(\eta, \gamma, P)$ and $(\eta_n, \gamma_n, P_n) \to (\eta, \gamma, P)$, there exists a subsequence $\{\eta_{n_k}, \gamma_{n_k}, P_{n_k}\}_{k=1}^{\infty}$ and $y_{n_k} \in \Pi(\eta_{n_k}, \gamma_{n_k}, P_{n_k})$ with $y_{n_k} \to y$. Since $y \in \Pi(\eta, \gamma, P)$, there exists a $(\lambda_0(y), \lambda_1(y)) \in [0, 2]^2$ with

$$y \leq F(\lambda_0(y), \lambda_1(y), \eta, \gamma, P). \tag{A78}$$

If $\lambda_1(y) > 0$, then we immediately have from (A77) that

$$F(\lambda_0(y), \lambda_1(y), \eta_n, \gamma_n, P_n) \to F(\lambda_0(y), \lambda_1(y), \eta, \gamma, P), \tag{A79}$$

from which it follows that there exists $y_n \in \Pi(\eta_n, \gamma_n, P_n)$ with $y_n \to y$. To address the case $\lambda_1(y) = 0$, notice that

$$\begin{aligned} &\limsup_{n\to\infty} F(\lambda_0(y), 0, \eta_n, \gamma_n, P_n) \\ &\quad = \lambda_0(y) + \eta - e^{\lambda_0(y)-1} \times \liminf_{n\to\infty} P_n\big\{1 + \gamma_n' g(X, \theta_0(P_n)) > 0\big\} \\ &\quad \geq \lambda_0(y) + \eta - e^{\lambda_0(y)-1} \times P\big\{1 + \gamma' g(X, \theta_0(P)) > 0\big\} \\ &\quad = F(\lambda_0(y), 0, \eta, \gamma, P), \end{aligned} \tag{A80}$$

where the inequality is implied by $P_n \to P$, (A74), Theorem 1.11.1 in van der Vaart and Wellner (1996), and the portmanteau lemma. The final equality in (A80) is definitional. The existence of a subsequence $\{\gamma_{n_k}, \eta_{n_k}, P_{n_k}\}_{k=1}^{\infty}$ with $y_{n_k} \in \Pi(\gamma_{n_k}, \eta_{n_k}, P_{n_k})$ and $y_{n_k} \to y$ then follows. *Q.E.D.*

LEMMA A.9: *If Assumptions* 3.1–3.5 *hold, then for every* $Q \in \mathbf{P}_0$, *there exists an open neighborhood* $N(Q)$ *in* $\mathbf{P}_0$ *with respect to the weak topology and a* $\bar{\eta}(Q) > 0$ *such that for all* $\eta \in [0, \bar{\eta}(Q)]$,

$$\begin{aligned} &\inf_{P\in N(Q)} \inf_{\gamma\in\Gamma(\eta,P)} \sup_{\lambda_0,\lambda_1\geq 0} \lambda_0 + \eta(\lambda_1 - 1) \\ &\qquad - e^{\lambda_0-1} \int \big(1 + \gamma' g(x, \theta_0(P))\big)^{\lambda_1} I\big\{1 + \gamma' g(x, \theta_0(P)) > 0\big\}\, dP \geq 0. \end{aligned} \tag{A81}$$

PROOF: First notice that since by Lemma A.6, the correspondence $\Gamma(0, Q)$ is compact valued, there exists a compact set A such that,

$$\Gamma(0, Q) \subset A.$$

Furthermore, since by Lemma A.6, $\Gamma(\eta, P)$ is also upper hemicontinuous at $(\eta, P) = (0, Q)$, there exists a $\alpha(Q) > 0$ and an open neighborhood $B(Q)$ in $\mathbf{P}_0$ such that for all $0 \leq \eta \leq \alpha(Q)$ and $P \in B(Q)$, we have

$$\Gamma(\eta, P) \subset A. \tag{A82}$$

Thus, since $[0,2]^2 \subset \mathbf{R} \times \mathbf{R}_+$, it immediately follows that for all $0 \le \eta \le \alpha(Q)$ and $P \in B(Q)$,

$$\text{(A83)}\quad \inf_{\gamma \in \Gamma(\eta,P)} \sup_{\lambda_0,\lambda_1 \ge 0} \lambda_0 + \eta(\lambda_1 - 1) - e^{\lambda_0 - 1} \int \big(1+\gamma' g(x,\theta_0(P))\big)^{\lambda_1} I\big\{1+\gamma' g(x,\theta_0(P)) > 0\big\}\, dP$$
$$\ge \inf_{\gamma \in A} \sup_{(\lambda_0,\lambda_1) \in [0,2]^2} \lambda_0 + \eta(\lambda_1 - 1) - e^{\lambda_0 - 1} \int \big(1+\gamma' g(x,\theta_0(P))\big)^{\lambda_1} I\big\{1+\gamma' g(x,\theta_0(P)) > 0\big\}\, dP.$$

We establish the lemma by showing that for η sufficiently small, the right hand side of (A83) is nonnegative on an open neighborhood of Q. For this purpose, define the function

$$\text{(A84)}\quad F(\lambda_0,\lambda_1,\eta,\gamma,Q) = \lambda_0 + \eta(\lambda_1 - 1) - e^{\lambda_0 - 1} \int \big(1+\gamma' g(x,\theta_0(Q))\big)^{\lambda_1} I\big\{1+\gamma' g(x,\theta_0(Q)) > 0\big\}\, dQ.$$

By Lemmas A.7 and A.8 and Theorem 2 in Ausubel and Deneckere (1993), it follows that

$$\text{(A85)}\quad C(\gamma,\eta,Q) = \max_{(\lambda_0,\lambda_1) \in [0,2]^2} F(\lambda_0,\lambda_1,\eta,\gamma,Q)$$

is continuous on $(\gamma,\eta,Q) \in \mathbf{R}^m \times \mathbf{R}_+ \times \mathbf{P}_0$. Moreover, since A is compact, applying Berge's theorem of the maximum establishes that the correspondence

$$\text{(A86)}\quad \Xi(\eta,P) = \arg\min_{\gamma \in A} C(\gamma,\eta,P)$$

is well defined and upper hemicontinuous on $\mathbf{R}_+ \times \mathbf{P}_0$.

We now show $\Xi(0,Q) = \{0\}$. If $\gamma \in A \backslash \Gamma(0,Q)$, then $Q\{1+\gamma' g(X,\theta_0(Q)) \ge 0\} < 1$ and hence

$$\text{(A87)}\quad F(1,0,0,\gamma,Q) = 1 - Q\big\{1+\gamma' g(X,\theta_0(Q)) > 0\big\} > 0.$$

Alternatively, for any $0 \ne \gamma \in \Gamma(0,Q)$, we have $Q\{1+\gamma' g(X,\theta_0(Q)) \ge 0\} = 1$. Therefore,

$$\text{(A88)}\quad F(1,1,0,\gamma,Q) = 1 - \int \big(1+\gamma' g(X,\theta_0(Q))\big)\, dQ = 0$$

by virtue of $Q \in \mathbf{P}_0$. Further, since $Q \in \mathbf{P}_0$, Assumption 3.5 implies that for $\gamma \neq 0$,

$$0 < Q\{\gamma' g(X, \theta_0(Q)) \geq 0\} < 1. \tag{A89}$$

Next use the dominated convergence theorem to exchange the order of differentiation and integration in (A88) and conclude that for $0 \neq \gamma \in \Gamma(0, Q)$,

$$\begin{aligned} &\frac{\partial}{\partial \lambda_1} F(1, \lambda_1, 0, \gamma, Q)\bigg|_{\lambda_1 = 1} \\ &\quad = \int \big(1 + \gamma' g(x, \theta_0(Q))\big) \log\big(1 + \gamma' g(x, \theta_0(Q))\big) \\ &\qquad \times I\big\{1 + \gamma' g(x, \theta_0(Q)) > 0\big\}\, dQ \\ &\quad > 0, \end{aligned} \tag{A90}$$

where the inequality holds by (A89), which implies $\gamma' g(x, \theta_0(Q))$ is not constant on $\operatorname{supp}_Q(g(X, \theta_0(Q)))$ and, therefore, Jensen's inequality holds strictly. Hence, if $0 \neq \gamma \in \Gamma(0, Q)$, there exists $1 \leq \tilde{\lambda}_1 \leq 2$ such that

$$F(1, \tilde{\lambda}_1, 0, \gamma, Q) > 0. \tag{A91}$$

Thus, so far we have established through (A87) and (A91) that if $0 \neq \gamma \in A$, then

$$C(\gamma, 0, Q) > 0.$$

Alternatively, it follows from direct calculation that $C(0, 0, Q) = 0$, and hence we conclude

$$\Xi(0, Q) = \{0\}. \tag{A92}$$

Next notice that continuity of $g(x, \theta)$ in (x, θ) and compactness of $\mathcal{X}$ and Θ imply that

$$\sup_{\theta \in \Theta} \sup_{x \in \mathcal{X}} |g(x, \theta)| < \infty. \tag{A93}$$

Furthermore, since, as argued, $\Xi(\eta, P)$ is upper hemicontinuous at $(\eta, P) = (0, Q)$, it follows from (A92) and (A93) that there exist an $\alpha(Q) \geq \bar{\eta}(Q) > 0$ and an open neighborhood $N(Q) \subseteq B(Q)$ such that if $\eta \in [0, \bar{\eta}(Q)]$ and $P \in N(Q)$, then

$$\sup_{\gamma \in \Xi(\eta, P)} |\gamma| < \frac{1}{\sup_{x \in \mathcal{X}} |g(x, \theta_0(P))|}. \tag{A94}$$

We therefore conclude that if $0 \leq \eta \leq \bar{\eta}(Q)$, $P \in N(Q)$, and $\gamma \in \Xi(\eta, P)$, then

$$P\{1 + \gamma' g(X, \theta_0(P)) \geq 0\} = 1.$$

It follows that if $0 \leq \eta \leq \bar{\eta}(Q)$ and $P \in N(Q)$, then

$$\Xi(\eta, P) \subseteq \Gamma(0, P).$$

Consequently, we obtain that for all $0 \leq \eta \leq \bar{\eta}(Q)$ and $P \in N(Q)$,

$$\begin{aligned}
\text{(A95)} \quad & \min_{\gamma \in A} \max_{(\lambda_0, \lambda_1) \in [0,2]^2} \lambda_0 + \eta(\lambda_1 - 1) \\
& \qquad - e^{\lambda_0 - 1} \int \bigl(1 + \gamma' g(x, \theta_0(P))\bigr)^{\lambda_1} I\{1 + \gamma' g(x, \theta_0(P)) > 0\}\, dP \\
& = \min_{\gamma \in \Gamma(0,P)} \max_{(\lambda_0, \lambda_1) \in [0,2]^2} \lambda_0 + \eta(\lambda_1 - 1) \\
& \qquad - e^{\lambda_0 - 1} \int \bigl(1 + \gamma' g(x, \theta_0(P))\bigr)^{\lambda_1} I\{1 + \gamma' g(x, \theta_0(P)) > 0\}\, dP.
\end{aligned}$$

Arguing as in (A88), it then follows that $F(1, 1, \eta, \gamma, P) = 0$ for all $\gamma \in \Gamma(0, P)$ and any η. To conclude, note that since the minimum is attained, we establish using (A95) that

$$\begin{aligned}
\text{(A96)} \quad & \min_{\gamma \in \Gamma(0,P)} \max_{(\lambda_0, \lambda_1) \in [0,2]^2} \lambda_0 + \eta(\lambda_1 - 1) \\
& \quad - e^{\lambda_0 - 1} \int \bigl(1 + \gamma' g(x, \theta_0(P))\bigr)^{\lambda_1} I\{1 + \gamma' g(x, \theta_0(P)) > 0\}\, dP \geq 0.
\end{aligned}$$

Therefore (A83), (A95) and (A96) establish the claim of the lemma. *Q.E.D.*

PROOF OF THEOREM 3.1: (a) Let $\ddot{\mathbf{M}}_0(Q)$, $\ddot{\mathbf{M}}_0(Q, \theta)$, and $\ddot{\Lambda}_1(\eta)$ be as defined in (A7), (A9), and (A8), respectively. Observe that since $\ddot{\mathbf{M}}_0(Q) \subseteq \mathbf{M}_0(Q)$, it follows that $\Lambda_1(\eta) \subseteq \ddot{\Lambda}_1(\eta)$. Hence

$$\text{(A97)} \quad \limsup_{n \to \infty} \frac{1}{n} \log P^n\{\hat{P}_n \in \Lambda_1(\eta)\} \leq \limsup_{n \to \infty} \frac{1}{n} \log P^n\{\hat{P}_n \in \ddot{\Lambda}_1(\eta)\}.$$

The proof then proceeds by showing that the conditions of Lemma A.3 hold for all $P \in \mathbf{P}_0$ if $\eta > 0$ is sufficiently small. Toward this end, for ς as in Lemma A.5, define

$$\text{(A98)} \quad \omega_1 = -\log(1 - \varsigma).$$

We first show that for all $P \in \mathbf{P}_0$, if $Q \ll P$ and $P\{X \in \text{supp}(Q)\} > \exp(-\omega_1)$, then $\ddot{\mathbf{M}}_0(Q, \theta_0(P)) \neq \emptyset$. For this purpose, note that

$$\begin{aligned} \text{(A99)} \qquad \sup_{\gamma \neq 0} P\bigl\{X \in \text{supp}(Q), \gamma' g(X, \theta_0(P)) \geq 0\bigr\} &\leq \sup_{\gamma \neq 0} P\bigl\{\gamma' g(X, \theta_0(P)) \geq 0\bigr\} \\ &\leq 1 - \varsigma \\ &< P\{X \in \text{supp}(Q)\}, \end{aligned}$$

where the first inequality follows by inspection, the second inequality follows by $P \in \mathbf{P}_0$ and Lemma A.5, and the last inequality follows by hypothesis. It follows from (A99) that for all $\gamma \in \mathbf{R}^m$,

$$\text{(A100)} \qquad P\bigl\{X \in \text{supp}(Q), \gamma' g(X, \theta_0(P)) \geq 0\bigr\} > 0,$$

$$\text{(A101)} \qquad P\bigl\{X \in \text{supp}(Q), \gamma' g(X, \theta_0(P)) < 0\bigr\} > 0.$$

Hence, there exists no hyperplane separating $\text{supp}_Q(g(X, \theta_0(P)))$ and $\{0\}$, which implies

$$0 \in \text{int}\bigl(\text{co}\bigl(\text{supp}_Q\bigl(g(X, \theta_0(P))\bigr)\bigr)\bigr)$$

(relative to the topology on $\mathbf{R}^m$). We therefore conclude $\ddot{\mathbf{M}}_0(Q, \theta_0(P)) \neq \emptyset$ as desired.

To complete the proof, we verify that (A37) holds uniformly in $P \in \mathbf{P}_0$ for $\eta > 0$ sufficiently small. By Lemma A.9, for every $P \in \mathbf{P}_0$, there exists an $\bar{\eta}(P) > 0$ and an open neighborhood in the weak topology $N(P)$ in $\mathbf{P}_0$ such that for all $0 \leq \eta \leq \bar{\eta}(P)$, we have

$$\begin{aligned} &\inf_{Q \in N(P)} \inf_{\gamma \in \Gamma(\eta, Q)} \sup_{\lambda_0, \lambda_1 \geq 0} \lambda_0 + \eta(\lambda_1 - 1) \\ &\quad - e^{\lambda_0 - 1} \int \bigl(1 + \gamma' g(x, \theta_0(Q))\bigr)^{\lambda_1} I\bigl\{1 + \gamma' g(x, \theta_0(Q)) > 0\bigr\}\, dQ \geq 0. \end{aligned}$$

By Theorem 15.11 in Aliprantis and Border (2006), $\mathbf{M}$ is compact under the weak topology, and hence since $\mathbf{P}_0 \subset \mathbf{M}$ is closed by Lemma A.4, it is compact as well. Consequently, as

$$\mathbf{P}_0 = \bigcup_{P \in \mathbf{P}_0} N(P)$$

and $N(P)$ are open for all $P \in \mathbf{P}_0$, compactness implies the existence of a finite subcover such that

$$\text{(A102)} \qquad \mathbf{P}_0 = \bigcup_{i=1}^{k} N(P_i).$$

To conclude, let

$$\omega_2 = \min\{\bar{\eta}(P_1), \dots, \bar{\eta}(P_k)\} \tag{A103}$$

and notice that by construction, $\omega_2 > 0$ and, in addition, for all $0 \le \eta \le \omega_2$,

$$\begin{aligned} &\inf_{P\in\mathbf{P}_0} \inf_{\gamma\in\Gamma(\eta,P)} \sup_{\lambda_0,\lambda_1\ge 0} \lambda_0 + \eta(\lambda_1 - 1) \\ &\qquad - e^{\lambda_0 - 1}\int \big(1 + \gamma' g(x, \theta_0(P))\big)^{\lambda_1} I\{1 + \gamma' g(x, \theta_0) > 0\}\, dP \ge 0. \end{aligned} \tag{A104}$$

Letting $\bar{\eta} = \min\{\omega_1, \omega_2\}$ implies that the conditions of Lemma A.3 are satisfied for all $P \in \mathbf{P}_0$ and $0 \le \eta \le \bar{\eta}$, which establishes claim (a) of the theorem.

The proof of part (b) closely follows arguments in Kitamura (2001) and Dembo and Zeitouni (1998). Define the set of probability measures

$$\mathbf{R}(\eta) = \Big\{Q \in \mathbf{M} : \inf_{P\in\mathbf{M}_0\setminus\mathbf{P}_0} I(Q|P) \ge \eta\Big\}. \tag{A105}$$

We first aim to show that the proposition

$$\Lambda_0(\eta) \cap \mathbf{R}(\eta) \subseteq \Omega_{0,n} \cap \mathbf{R}(\eta) \tag{A106}$$

holds for all $n > n_0$ and n_0 sufficiently large. Suppose by way of contradiction that there exists an infinite sequence of probability measures $\{\xi_n\}_{n=1}^{\infty}$ such that $\xi_n \in \Lambda_0(\eta) \cap \mathbf{R}(\eta)$ and $\xi_n \in \Omega_{1,n} \cap \mathbf{R}(\eta)$. Since $\mathbf{M}$ is compact under the weak topology by Theorem 15.11 in Aliprantis and Border (2006), there exists a subsequence $\{\xi_{n_k}\}_{k=1}^{\infty}$ such that

$$\xi_{n_k} \to \xi \tag{A107}$$

for some $\xi \in \mathbf{M}$. Hence, there exists a k_0 such that for all $k \ge k_0$, it follows that $\xi_{n_k} \in B(\xi, \delta/2)$ and, therefore, $B(\xi, \delta/2) \subset \Omega^{\delta}_{1,n_k}$. Hence, by Sanov's theorem (see Theorem 6.2.10 in Dembo and Zeitouni (1998)) and various inclusion restrictions, it follows that

$$\begin{aligned} \sup_{P\in\mathbf{P}_0} \limsup_{n\to\infty} \frac{1}{n} \log P^n\{\hat{P}_n \in \Omega^{\delta}_{1,n}\} &\ge \sup_{P\in\mathbf{P}_0} \liminf_{n_k\to\infty} \frac{1}{n_k} \log P^{n_k}\big\{\hat{P}_{n_k} \in \Omega^{\delta}_{1,n_k}\big\} \\ &\ge \sup_{P\in\mathbf{P}_0} \liminf_{n\to\infty} \frac{1}{n} \log P^n\{\hat{P}_n \in B(\xi, \delta/2)\} \\ &\ge \sup_{P\in\mathbf{P}_0} - \inf_{Q\in B(\xi,\delta/2)} I(Q|P) \\ &\ge \sup_{P\in\mathbf{P}_0} -I\big(\xi_{n_{k_0}}|P\big). \end{aligned} \tag{A108}$$

Since $\xi_{n_{k_0}} \in \Lambda_0(\eta) \cap \mathbf{R}(\eta)$, it must be that

$$\inf_{P\in\mathbf{M}_0} I(\xi_{n_{k_0}}|P) \leq \inf_{P\in\mathbf{M}_0(\xi_{n_{k_0}})} I(\xi_{n_{k_0}}|P) < \eta \tag{A109}$$

by virtue of $\mathbf{M}_0(\xi_{n_{k_0}}) \subseteq \mathbf{M}_0$ and $\xi_{n_{k_0}} \in \Lambda_0(\eta)$. Furthermore, since $\xi_{n_{k_0}} \in \mathbf{R}(\eta)$, we have

$$\inf_{P\in\mathbf{M}_0\setminus\mathbf{P}_0} I(\xi_{n_{k_0}}|P) \geq \eta. \tag{A110}$$

Hence, combining (A109), (A110), and $\mathbf{P}_0 \subset \mathbf{M}_0$, we conclude

$$\inf_{P\in\mathbf{P}_0} I(\xi_{n_{k_0}}|P) < \eta. \tag{A111}$$

Therefore, it follows from results (A108) and (A111) that

$$\sup_{P\in\mathbf{P}_0} \limsup_{n\to\infty} \frac{1}{n} \log P^n\{\hat{P}_n \in \Omega^\delta_{1,n}\} > -\eta, \tag{A112}$$

which contradicts the assumptions on $(\Omega_{0,n}, \Omega_{1,n})$ and hence (A106) must be true. Therefore,

$$\limsup_{n\to\infty} \frac{1}{n} \log Q^n\{\hat{P}_n \in \Lambda_0(\eta) \cap R(\eta)\} \leq \limsup_{n\to\infty} \frac{1}{n} Q^n\{\hat{P}_n \in \Omega_{0,n}\}. \tag{A113}$$

To conclude the proof of claim (b) of Theorem 3.1, we aim to establish that

$$\limsup_{n\to\infty} \frac{1}{n} \log Q^n\{\hat{P}_n \in \Lambda_0(\eta)\} = \limsup_{n\to\infty} \frac{1}{n} \log Q^n\{\hat{P}_n \in \Lambda_0(\eta) \cap R(\eta)\} \tag{A114}$$

for all $Q \in \mathbf{P}$ satisfying (18) in the main text, which together with inequality (A113) yields the desired result. Toward this end, let $R^c(\eta) = \mathbf{M} \setminus R(\eta)$ and note that since $I(Q|P) \geq 2d^2_{PL}(Q, P)$, it follows from (A105) that

$$R^c(\eta) \subseteq \Big\{Q \in \mathbf{M} : \inf_{P\in\mathbf{M}_0\setminus\mathbf{P}_0} 2d^2_{PL}(Q, P) \leq \eta\Big\}. \tag{A115}$$

We conclude from (A115) that $R^c(\eta) \subseteq \overline{\mathbf{M}_0 \setminus \mathbf{P}_0}^{\sqrt{\eta/2}}$. Thus appealing to Sanov's theorem, we obtain

$$\begin{aligned}
&\limsup_{n\to\infty} \frac{1}{n} \log Q^n\{\hat{P}_n \in \Lambda_0(\eta) \cap R^c(\eta)\} \\
&\quad \leq \limsup_{n\to\infty} \frac{1}{n} \log Q^n\{\hat{P}_n \in \overline{\Lambda_0(\eta) \cap R^c(\eta)}\}
\end{aligned} \tag{A116}$$

$$\le -\inf_{P\in\Lambda_0(\eta)\cap R^c(\eta)} I(Q|P)$$
$$\le -\inf_{P\in\overline{\mathbf{M}_0\setminus\mathbf{P}_0}^{\sqrt{\eta/2}}} I(Q|P).$$

Next let $\ddot{\Lambda}_0(\eta)=\mathbf{M}\setminus\ddot{\Lambda}_1(\eta)$ for $\ddot{\Lambda}_1(\eta)$ as defined in (A8), and notice that since $\mathbf{P}_0\subseteq\ddot{\Lambda}_0(\eta)\subseteq\Lambda_0(\eta)$ and $\ddot{\Lambda}_0(\eta)$ is open as a consequence of Lemma A.2, it follows that

$$\mathbf{P}_0\subseteq\Lambda_0^o(\eta). \tag{A117}$$

Moreover, since from result (A115) it is possible to conclude that

$$\Big\{Q\in\mathbf{M}:\inf_{P\in\mathbf{M}_0\setminus\mathbf{P}_0} 2d_{PL}^2(Q,P)>\eta\Big\}\subseteq R(\eta) \tag{A118}$$

and since the left hand side of (A118) is open, we obtain $\mathbf{P}_0\setminus\overline{\mathbf{M}_0\setminus\mathbf{P}_0}^{\sqrt{\eta/2}}\subset R^o(\eta)$. Thus, by (A117) and (A118),

$$\mathbf{P}_0\setminus\overline{\mathbf{M}_0\setminus\mathbf{P}_0}^{\sqrt{\eta/2}}\subseteq(\Lambda_0(\eta)\cap R(\eta))^o. \tag{A119}$$

Hence, appealing once again to Sanov's theorem and exploiting (A119), we establish

$$\liminf_{n\to\infty}\frac{1}{n}\log Q^n\{\hat{P}_n\in\Lambda_0(\eta)\cap R(\eta)\} \tag{A120}$$
$$\ge\liminf_{n\to\infty}\frac{1}{n}\log Q^n\big\{\hat{P}_n\in(\Lambda_0(\eta)\cap R(\eta))^o\big\}$$
$$\ge-\inf_{P\in(\Lambda_0(\eta)\cap R(\eta))^o} I(Q|P)$$
$$\ge-\inf_{P\in\mathbf{P}_0\setminus\overline{\mathbf{M}_0\setminus\mathbf{P}_0}^{\sqrt{\eta/2}}} I(Q|P).$$

Therefore, combining (A116) and (A120), we obtain that for any Q satisfying (18) in the main text, we have

$$\limsup_{n\to\infty}\frac{1}{n}\log\frac{Q^n\{\hat{P}_n\in\Lambda_0(\eta)\cap R^c(\eta)\}}{Q^n\{\hat{P}_n\in\Lambda_0(\eta)\cap R(\eta)\}}\le 0, \tag{A121}$$

which in turn implies that for any $\kappa>0$, there is an n_0 such that for all $n>n_0$,

$$\frac{Q^n\{\hat{P}_n\in\Lambda_0(\eta)\cap R^c(\eta)\}}{Q^n\{\hat{P}_n\in\Lambda_0(\eta)\cap R(\eta)\}}\le e^{\kappa n}. \tag{A122}$$

Moreover, employing $\Lambda_0(\eta) \cap R(\eta) \subseteq \Lambda_0(\eta)$, result (A122), and simple manipulations yields that

$$\begin{aligned}
0 &\leq \limsup_{n\to\infty} \frac{1}{n} \log Q^n\{\hat{P}_n \in \Lambda_0(\eta)\} \\
&\quad - \limsup_{n\to\infty} \frac{1}{n} \log Q^n\{\hat{P}_n \in \Lambda_0(\eta) \cap R(\eta)\} \\
&\leq \limsup_{n\to\infty} \frac{1}{n} \log\left\{1 + \frac{Q^n\{\hat{P}_n \in \Lambda_0(\eta) \cap R^c(\eta)\}}{Q^n\{\hat{P}_n \in \Lambda_0(\eta) \cap R(\eta)\}}\right\} \\
&\leq \limsup_{n\to\infty} \frac{1}{n} \log\{1 + e^{\kappa n}\} \\
&\leq 2\kappa.
\end{aligned} \tag{A123}$$

Because $\kappa > 0$ was arbitrary, we conclude that (A114) holds, which together with (A113) establishes the theorem. *Q.E.D.*

PROOF OF COROLLARY 3.1: Let $\tilde{\mathbf{P}} = \mathbf{P} \setminus \mathbf{D}_0^{\varepsilon/2}$ and notice that under the assumptions of the corollary, Assumptions 3.3, 3.4, and 3.5 hold with $\tilde{\mathbf{P}}$ in place of $\mathbf{P}$. Therefore, since

$$\left\{P \in \tilde{\mathbf{P}} : E_P[g(X, \theta)] = 0 \text{ for some } \theta \in \Theta\right\} = \mathbf{P}_0 \setminus \mathbf{D}_0^{\varepsilon/2}, \tag{A124}$$

it follows from Theorem 3.1 that there exists an $\bar{\eta}_1(\varepsilon) > 0$ such that for all $\bar{\eta}_1(\varepsilon) \geq \eta > 0$, we have

$$\begin{aligned}
&\sup_{P \in \mathbf{P}_0 \setminus \mathbf{D}_0^{\varepsilon/2}} \limsup_{n\to\infty} \frac{1}{n} \log P^n\{\hat{P}_n \in \tilde{\Lambda}_1(\eta) \setminus \mathbf{D}_0^{\varepsilon}\} \\
&\quad \leq \sup_{P \in \mathbf{P}_0 \setminus \mathbf{D}_0^{\varepsilon/2}} \limsup_{n\to\infty} \frac{1}{n} \log P^n\{\hat{P}_n \in \Lambda_1(\eta)\} \\
&\quad \leq -\eta.
\end{aligned} \tag{A125}$$

Moreover, by $\tilde{\Lambda}_1(\eta) \setminus \mathbf{D}_0^{\varepsilon} \subseteq (\mathbf{D}_0^{\varepsilon})^c$, Sanov's theorem, and the inequality $I(Q|P) \geq 2d_{PL}^2(Q, P)$, we obtain

$$\begin{aligned}
&\sup_{P \in \mathbf{P}_0 \cap \mathbf{D}_0^{\varepsilon/2}} \limsup_{n\to\infty} \frac{1}{n} \log P^n\{\hat{P}_n \in \tilde{\Lambda}_1(\eta) \setminus \mathbf{D}_0^{\varepsilon}\} \\
&\quad \leq \sup_{P \in \mathbf{P}_0 \cap \mathbf{D}_0^{\varepsilon/2}} \limsup_{n\to\infty} \frac{1}{n} \log P^n\{\hat{P}_n \in (\mathbf{D}_0^{\varepsilon})^c\}
\end{aligned} \tag{A126}$$

$$\le \sup_{P\in\mathbf{P}_0\cap\mathbf{D}_0^{\varepsilon/2}} - \inf_{Q\in(\mathbf{D}_0^{\varepsilon})^c} I(Q|P)$$

$$\le - \inf_{P\in\mathbf{P}_0\cap\mathbf{D}_0^{\varepsilon/2}} \inf_{Q\in(\mathbf{D}_0^{\varepsilon})^c} 2d_{PL}^2(Q, P).$$

Therefore, results (A125), (A126), and setting $\bar{\eta}(\varepsilon) \le \min\{\bar{\eta}_1(\varepsilon), \varepsilon^2/2\}$ establishes part (a) of the corollary. Furthermore, the same arguments as in (A107)–(A112) yield $\tilde{\Lambda}_0(\eta) \cup \mathbf{D}_0^{\varepsilon} \subseteq \Omega_{0,n} \cup \mathbf{D}_0^{\varepsilon}$ for n large enough, which implies $\tilde{\Lambda}_0(\eta) \subseteq \Omega_{0,n}$ for n large enough, thus yielding part (b) of the corollary. *Q.E.D.*

REFERENCES

ALIPRANTIS, C. D., AND K. C. BORDER (2006): *Infinite Dimensional Analysis—A Hitchhiker's Guide*. Berlin: Springer-Verlag. [11,13,24,25]

AUSUBEL, L., AND R. DENECKERE (1993): "A Generalized Theorem of the Maximum," *Economic Theory*, 3, 99–107. [21]

BORWEIN, J. M., AND A. S. LEWIS (1992a): "Partially Finite Convex Programming Part I: Quasi-Relative Interiors and Duality Theory," *Mathematical Programming*, 57, 15–48. [10]

——— (1992b): "Partially Finite Convex Programming Part II: Explicit Lattice Models," *Mathematical Programming*, 57, 49–83. [10]

——— (1993): "Partially Finite Programming in L_1 and the Existence of Maximum Entropy Estimates," *SIAM Journal of Optimization*, 3, 248–267. [3,5,9]

DEMBO, A., AND O. ZEITOUNI (1998): *Large Deviations Techniques and Applications*. New York: Springer. [1,8,25]

KITAMURA, Y. (2001): "Asymptotic Optimality of Empirical Likelihood for Testing Moment Restrictions," *Econometrica*, 69, 1661–1672. [25]

VAN DER VAART, A., AND J. WELLNER (1996): *Weak Convergence and Empirical Processes With Applications to Statistics*. New York: Springer-Verlag. [20]

Dept. of Economics, Yale University, New Haven, CT 06520-8281, U.S.A.; yuichi.kitamura@yale.edu,

Dept. of Economics, University of California–San Diego, 9500 Gilman Drive 0508, La Jolla, CA 92093-0508, U.S.A.; a2santos@ucsd.edu,

and

Dept. of Economics, University of Chicago, Chicago, IL 60637, U.S.A.; amshaikh@uchicago.edu.

Manuscript received August, 2009; final revision received April, 2011.