Supplementary Material on "Asymptotically Efficient Estimation of Weighted Average Derivatives with an Interval Censored Variable"

Hiroaki Kaido
Boston University

In this supplementary material, we include the proofs of results stated in the main text. The contents of the supplemental appendix are organized as follows. Appendix A contains notations and definitions used throughout the appendix. Appendix B contains the proof of Theorems 2.1 and 2.2 and Corollaries 2.1 and 2.2. Appendix C contains the proof of Theorem 3.1 and auxiliary lemmas. Appendix D contains the proof of Theorem 4.1. Appendix E then reports the Monte Carlo results.

Appendix A: Notation and Definitions

Let $\Pi : \mathcal{X} \to \mathcal{Z}$ be the projection map pointwise defined by $x = (y_L, y_U, z) \mapsto z$. Let $\nu = \Pi_{\#}\mu$ be the pushforward measure of μ on $\mathcal{Z}$. We then denote the marginal density of P with respect to ν by $\phi_0^2(z)$. By the disintegration theorem, there exists a family $\{\mu_z : z \in \mathcal{Z}\}$ of probability measures on $\mathcal{X}$. Throughout, we assume that μ_z is absolutely continuous with respect to some σ-finite measure λ for all $z \in \mathcal{Z}$. We then denote the conditional density function of P with respect to λ by $v_0^2(y_L, y_U|z)$.

For any $1 \le p \le \infty$, we let $\|\cdot\|_{L^p_\pi}$ be the usual L^p-norm with respect to a measure π, where $\|\cdot\|_{L^\infty_\pi}$ denotes the essential supremum.

Appendix B: Proof of Theorems 2.1 and 2.2

Proof of Theorem 2.1. We first show that the identified set can be written as

$$\Theta_0(P) = \{\theta : \theta = E[m(Z)l(Z)],\ P(m_L(Z) \le m(Z) \le m_U(Z)) = 1\}. \tag{B.1}$$

For this, we note that, by Assumptions 2.1-2.3 and arguing as in the proof of Theorem 1 in Stoker (1986), we have

$$E[w(Z)\nabla_z m(Z)] = E[m(Z)l(Z)]. \tag{B.2}$$

Further, the distribution of Y first-order stochastically dominates that of Y_L. Similarly, the distribution of Y_U first-order stochastically dominates that of Y. Since q is nondecreasing by the convexity of ϱ, it then follows that, for each $u \in \mathbb{R}$,

$$E[q(Y_L - u)|Z] \le E[q(Y - u)|Z] \le E[q(Y_U - u)|Z],\ P - a.s. \tag{B.3}$$

Eq. (3) then follows by (B.3), Assumption 2.3 (iii), and the hypothesis that $E[q(Y - u)|Z = z] = 0$ has a unique solution at $u = m(z)$ on D.

For the convexity of $\Theta_0(P)$, observe that for any $\theta_1, \theta_2 \in \Theta_0(P)$, there exist $m_1, m_2 : \mathcal{Z} \to \mathbb{R}$ such that $\theta_j = E[m_j(Z)l(Z)]$ and $m_L(Z) \le m_j(Z) \le m_U(Z), P - a.s.$ for $j = 1, 2$. Let $\alpha \in [0, 1]$ and let $\theta_\alpha \equiv \alpha\theta_1 + (1-\alpha)\theta_2$. Then,

$$\theta_\alpha = E[m_\alpha(Z)l(Z)],$$

where $m_\alpha \equiv \alpha m_1 + (1-\alpha)m_2$. Since $m_L(Z) \le m_\alpha(Z) \le m_U(Z), P - a.s.$, it follows that $\theta_\alpha \in \Theta_0(P)$. Therefore, $\Theta_0(P)$ is convex.

We show compactness of $\Theta_0(P)$ by showing $\Theta_0(P)$ is bounded and closed. By Assumption 2.3 (i)-(ii), for any $\theta \in \Theta_0(P)$,

$$|\theta^{(j)}| \le \sup_{z\in\mathcal{Z}} |m(z)| E[|l^{(j)}(Z)|] \le \sup_{x\in D} |x| E[|l^{(j)}(Z)|] < \infty, \quad \text{for } j = 1, \cdots, \ell. \tag{B.4}$$

Hence, $\Theta_0(P)$ is bounded. To see that $\Theta_0(P)$ is closed, consider the following maximization problem:

$$\text{maximize} \quad E[m(Z)p'l(Z)], \tag{B.5}$$
$$\text{s.t.} \quad m_L(Z) \le m(Z) \le m_U(Z), P - a.s. \tag{B.6}$$

Arguing as in the proof of Proposition 2 in Bontemps, Magnac, and Maurin (2012), the objective function is maximized by setting $m(z) = m_L(z)$ when $p'l(z) \le 0$ and setting $m(z) = m_U(z)$ otherwise. This and (B.2) give the support function of $\Theta_0(P)$ in (13) and also shows that, for each $p \in \mathbb{S}$, there exists $m_p(z) \equiv 1\{p'l(z) \le 0\}m_L(z) + 1\{p'l(z) > 0\}m_U(z)$ such that $\upsilon(p, \Theta_0(P)) = \langle p, \theta^*(p)\rangle$, where $\theta^*(p) = E[m_p(Z)l(Z)]$. Since m_p satisfies $m_L(Z) \le m_p(Z) \le m_U(Z), P - a.s.$, we have $\theta^*(p) \in \Theta_0(P)$. By Proposition 8.29 (a) in Rockafellar and Wets (2005), the boundary of $\Theta_0(P)$ is $\{\tilde{\theta} : \langle p, \tilde{\theta}\rangle = \upsilon(p, \Theta_0(P)), p \in \mathbb{S}^\ell\}$. Therefore, $\Theta_0(P)$ contains its boundary, and hence it is closed.

For the strict convexity of $\Theta_0(P)$, we show it through the differentiability of the support function. The proof is similar to that of Lemma A.8 in Beresteanu and Molinari (2008) and Lemma 23 in Bontemps, Magnac, and Maurin (2012). To this end, we extend the support function and define $s(p, \Theta_0(P))$ as in (13) for each $p \in \mathbb{R}^\ell \setminus \{0\}$.

For each $z \in \mathcal{Z}$, let $\xi(z) \equiv (m_L(z) - m_U(z))l(z)$. For each $p \in \mathbb{R}^\ell \setminus \{0\}$, let $\zeta(p) \equiv E[1\{p'\xi(Z) \ge 0\}p'\xi(Z)]$. Then, since $m_L(Z) - m_U(Z) \le 0$ almost surely, it holds that $\upsilon(p, \Theta_0(P)) = \zeta(p) + E[m_U(Z)p'l(Z)]$ for all $p \in \mathbb{R}^\ell \setminus \{0\}$. For any $p, q \in \mathbb{R}^\ell \setminus \{0\}$, it then follows by the Cauchy-Shwarz inequality that

$$\begin{aligned} &|\zeta(q) - \zeta(p) - (q-p)'E[\xi(Z)1\{p'\xi(Z) \ge 0\}]| \\ &\qquad = |E[(1\{q'\xi(Z) \ge 0\} - 1\{p'\xi(Z) \ge 0\})q'\xi(Z)]| \le \|1\{q'\xi \ge 0\} - 1\{p'\xi \ge 0\}\|_{L^2_P} \|q'\xi\|_{L^2_P}. \end{aligned} \tag{B.7}$$

By Assumptions 2.1 (i), the distribution of $\xi(Z)$ does not assign a positive measure to any proper subspace of $\mathbb{R}^\ell$ with dimension $\ell - 1$, which ensures $P(p'\xi(Z) = 0) = 0$. Thus, for any sequence $\{q_n\}$ such that $q_n \to p$, it follows that $1\{q_n'\xi(Z) \ge 0\} \stackrel{a.s.}{\to} 1\{p'\xi(Z) \ge 0\}$ as $n \to \infty$. Note that $1\{p'\xi(Z)\}$ is bounded for all p. Thus, the function class $\{1^2\{p'\xi(\cdot)\} : p \in \mathbb{R}^\ell \setminus \{0\}\}$ is uniformly integrable. These results ensure that $\|1\{q'\xi \ge 0\} - 1\{p'\xi \ge 0\}\|_{L^2_P} \to 0$ as $q \to p$. This and $\|q'\xi\|_{L^2_P} < \infty$ imply that the right hand side of (B.7) is $o(1)$. Hence, ζ is differentiable at every point on $\mathbb{R}^\ell \setminus \{0\}$ with the derivative $E[\xi(Z)1\{p'\xi(Z) \ge 0\}]$. Note that, for each $z \in \mathcal{Z}$, $p \mapsto m_U(z)p'l(z)$ is differentiable with respect to p and $m_U l$ is integrable with respect to P by Assumption 2.3 (i)-(ii). This ensures that $p \mapsto E[m_U(Z)p'l(Z)]$ is differentiable with respect to

p at every $p \in \mathbb{R} \setminus \{0\}$. Therefore, the map $p \mapsto \upsilon(p, \Theta_0(P))$ is differentiable for all $p \in \mathbb{R}^\ell \setminus \{0\}$. By Corollary 1.7.3 in Schneider (1993), the support set $H(p, \Theta_0(P)) \equiv \{\theta : \langle p, \theta \rangle = \upsilon(p, \Theta_0(P))\} \cap \Theta_0(P)$ for each p then contains only one point, which ensures the strict convexity of $\Theta_0(P)$.

To see that $\Theta_0(P)$ is sharp, take any $\theta \in \Theta_0(P)$. Then, by convexity, there exist $p, q \in \mathbb{S}^\ell$ and $\alpha \in [0,1]$ such that $\theta = \alpha\theta^*(p) + (1-\alpha)\theta^*(q)$, which further implies

$$\theta = E[(\alpha m_p(Z) + (1-\alpha)m_q(Z))l(Z)] = E[w(Z)\nabla_z m_{\alpha,p,q}(Z)], \tag{B.8}$$

where $m_{\alpha,p,q} \equiv \alpha m_p + (1-\alpha)m_q$, and the last equality follows from integration by parts and Assumptions 2.1 (i) and 2.3 (iv) ensuring the almost everywhere differentiability of $m_{\alpha,p,q}$. Since $m_{\alpha,p,q}$ satisfies (3) in place of m with $m_{\alpha,p,q}$ and $m_{\alpha,p,q}$ is almost everywhere differentiable, θ is the weighted average derivative of a regression function consistent with some data generating process. Hence, $\Theta_0(P)$ is sharp. ■

Proof of Corollary 2.1. By Assumption 2.4, the weighted average derivative and index coefficients are related to each other by

$$\theta^{(j)} = E[w(Z)M'(Z'\beta)]\beta^{(j)},\ j = 1, \cdots, \ell. \tag{B.9}$$

By Theorem 2.1 and letting $p = \iota_j$ and $-\iota_j$ respectively, we obtain bounds $\theta_L^{(j)}, \theta_U^{(j)}$ on each component $\theta^{(j)}$ of the average derivative vector as in (16)-(17). Due to the scale normalization $\beta^{(1)} = 1$ and (B.9), we then obtain

$$\theta_L^{(1)} \le E[w(Z)M'(Z'\beta)] \le \theta_U^{(1)}. \tag{B.10}$$

Furthermore, by Assumption 2.4 and $w(z) \ge 0$ for all z, one may tighten these bounds as

$$\theta_{L,+}^{(1)} \le E[w(Z)M'(Z'\beta)] \le \theta_U^{(1)}, \tag{B.11}$$

where $\theta_{L,+}^{(1)} = \max\{\theta_L^{(1)}, 0\}$. By (B.9), (B.11), and $\theta_U^{(1)}, \theta_{L,+}^{(1)}$ being non-negative, it follows that

$$\frac{\theta_L^{(j)}}{\theta_U^{(1)}} \le \beta^{(j)} \le \frac{\theta_U^{(j)}}{\theta_{L,+}^{(1)}},\ j = 2, \cdots, \ell. \tag{B.12}$$

Intersecting these bounds with the a priori bounds $[\underline{\beta}^{(j)}, \overline{\beta}^{(j)}]$ yields the conclusion of the corollary. ■

Proof of Theorem 2.2. We first show (24). By the first order condition for (20) and Assumption 2.5 (ii), $E[q(Y - g(Z,V))|Z,V] = E[q(Y - g(Z,V))|\tilde{Z}, V] = 0$, $P-a.s.$ By Assumption 2.5 (i) and the monotonicity of q, for $v_L \le v \le v_U$, we have

$$\int q(y - g(z, v_U))dP(y|\tilde{z}, v) \le 0 \le \int q(y - g(z, v_L))dP(y|\tilde{z}, v),\ P - a.s. \tag{B.13}$$

Taking expectations with respect to V, we obtain

$$\int q(y - g(z, v_U))dP(y|\tilde{z}) \le 0 \le \int q(y - g(z, v_L))dP(y|\tilde{z}),\ P - a.s. \tag{B.14}$$

Further, by Assumption 2.6 (ii),

$$\int q(y - \gamma(z, v_L, v_U))dP(y|\tilde{z}) = 0. \tag{B.15}$$

By (B.14)-(B.15) and the monotonicity of q, we then have

$$g(z, v_L) \leq \gamma(z, v_L, v_U) \leq g(z, v_U). \tag{B.16}$$

Let $\Xi_L(v) \equiv \{(v_L, v_U) : v_L \leq v_U \leq v\}$ and $\Xi_U(v) \equiv \{(v_L, v_U) : v \leq v_L \leq v_U\}$. To prove the lower bound on $g(z, v)$, take any $v_U \leq v$. Then by Assumption 2.5 (i) and (B.16), we have $\gamma(z, v_L, v_U) \leq g(z, v)$ for any $(v_L, v_U) \in \Xi(v)$. Hence, it follows that $g_L(z, v) \equiv \sup_{(v_L, v_U)\in\Xi_L(v)} \gamma(z, v_L, v_U) \leq g(z, v)$. Note that $g_L(z, v)$ is weakly increasing in v by construction and differentiable in z with a bounded derivative by Assumption 2.6 (iii). Hence, it is consistent with Assumption 2.5 (i). Thus, the bound is sharp. A similar argument gives the upper bound. Hence, (24) holds. This and integration by parts imply that the sharp identified set can be written as

$$\Theta_{0,v}(P) = \{\theta : \theta = E[g(Z, v)l(Z)],\ P(g_L(Z, v) \leq g(Z, v) \leq g_U(Z, v)) = 1\}. \tag{B.17}$$

The rest of the proof is then similar to that of Theorem 2.1. It is therefore omitted. ■

Proof of Corollary 2.2. By Assumption 2.7, the weighted average derivative and index coefficients are related to each other by

$$\theta_v^{(j)} = E[w(Z)G'(Z'\beta, v)]\beta^{(j)},\ j = 1, \cdots, \ell. \tag{B.18}$$

By Theorem 2.2 and letting $p = \iota_j$ and $-\iota_j$ respectively, we obtain bounds $\theta_L^{(j)}(v), \theta_U^{(j)}(v)$ on each component $\theta^{(j)}(v)$ of the average derivative vector as in (27)-(28). Mimic the argument of the proof of Corollary 2.2. We then obtain

$$\frac{\theta_L^{(j)}(v)}{\theta_U^{(1)}(v)} \leq \beta^{(j)} \leq \frac{\theta_U^{(j)}(v)}{\theta_{L,+}^{(1)}(v)},\ \forall v \in \mathcal{V},\ j = 2, \cdots, \ell, \tag{B.19}$$

where $\theta_{L,+}^{(1)}(v) = \max\{\theta_L^{(1)}(v), 0\}$. Hence, by intersecting these bounds across $v \in \mathcal{V}$, we obtain the following bounds:

$$\sup_{v\in\mathcal{V}} \frac{\theta_L^{(j)}(v)}{\theta_U^{(1)}(v)} \leq \beta^{(j)} \leq \inf_{v\in\mathcal{V}} \frac{\theta_U^{(j)}(v)}{\theta_{L,+}^{(1)}(v)},\ j = 2, \cdots, \ell. \tag{B.20}$$

Finally, intersecting these bounds with the a priori bounds $[\underline{\beta}^{(j)}, \overline{\beta}^{(j)}]$ yields the conclusion of the corollary. ■

Appendix C: Proof of Theorem 3.1

This Appendix contains the proof of Theorem 3.1 and auxiliary lemmas needed to establish the main result.

Below, we adopt the framework of Bickel, Klassen, Ritov, and Wellner (1993). To characterize the efficiency bound, it proves useful to study a parametric submodel of $\mathbf{P}$ defined in (32). We define a parametric submodel through a curve in L^2_μ. Let $h_0 \equiv \sqrt{dP/d\mu}$. Let $\tilde{v} : \mathcal{X} \to \mathbb{R}$ and $\tilde{\phi} : \mathcal{Z} \to \mathbb{R}$ be bounded functions that are continuously differentiable in z with bounded derivatives. We then define

$$\bar{v}(x) \equiv \tilde{v}(x) - E[\tilde{v}(X)|Z = z], \quad \text{and} \quad \bar{\phi}(z) \equiv \tilde{\phi}(z) - E[\tilde{\phi}(Z)], \tag{C.1}$$

where expectations are with respect to $P \in \mathbf{P}$. For each $\eta \in \mathbb{R}$, define $v_\eta : \mathcal{X} \to \mathbb{R}$ and $\phi_\eta : \mathcal{Z} \to \mathbb{R}$ by

$$v_\eta^2(y_L, y_U|z) = v_0^2(y_L, y_U|z)(1 + 2\eta\bar{v}(x)), \quad \text{and} \quad \phi_\eta^2(z) = \phi_0^2(z)(1 + 2\eta\bar{\phi}(z)). \tag{C.2}$$

We then let h_η^2 be defined pointwise by

$$h_\eta^2(x) \equiv v_\eta^2(y_L, y_U|z)\phi_\eta^2(z). \tag{C.3}$$

It is straightforward to show that $\eta \mapsto h_\eta^2$ is a curve in L_μ^2 with the Fréchet derivative $\dot{h}_0 = \dot{v}_0\phi_0 + v_0\dot{\phi}_0$, where $\dot{v}_0(y_L, y_U|z) \equiv \bar{v}(x)v_0(y_L, y_U|z)$ and $\dot{\phi}_0(z) = \bar{\phi}(z)\phi_0(z)$. We also note that for any η and η_0 in a neighborhood of 0, it holds that

$$v_\eta^2(y_L, y_U|z) = v_{\eta_0}^2(y_L, y_U|z)(1 + 2(\eta - \eta_0)\bar{v}_{\eta_0}(x)), \quad \text{and} \quad \phi_\eta^2(z) = \phi_{\eta_0}^2(z)(1 + 2\eta\bar{\phi}_{\eta_0}(z)). \tag{C.4}$$

where $\bar{v}_{\eta_0} = \bar{v}v_0^2/v_{\eta_0}^2$ and $\bar{\phi}_{\eta_0} = \bar{\phi}\phi_0^2/\phi_{\eta_0}^2$. We then define $\dot{v}_{\eta_0}(y_L, y_U|z) = \bar{v}_{\eta_0}(x)v_{\eta_0}(y_L, y_U|z)$ and $\dot{\phi}_{\eta_0}(z) = \bar{\phi}_{\eta_0}(z)\phi_{\eta_0}(z)$.

We further introduce notation for population objects along this curve. Let $f_\eta(z) \equiv \phi_\eta^2(z)$ and $l_\eta \equiv -\nabla_z w(z) - w(z)\nabla_z f_\eta(z)/f_\eta(z)$. Lemma C.1 will show that there exists a neighborhood N of 0 such that the equations $\int q(y_L - \tilde{m})v_\eta^2(y_L, y_U|z)d\lambda(y_L, y_U) = 0$ and $\int q(y_U - \tilde{m})v_\eta^2(y_L, y_U|z)d\lambda(y_L, y_U) = 0$ have unique solutions on D for all $\eta \in N$. We denote these solutions by $m_{L,\eta}$ and $m_{U,\eta}$ respectively. We then let $m_{p,\eta}(z) \equiv \Gamma(m_{L,\eta}(z), m_{U,\eta}(z), p'l_\eta(z))$. Further, we define

$$r_{j,\eta}(z) \equiv -\frac{d}{d\tilde{m}}E_\eta\left[q(y_j - \tilde{m})|Z = z\right]\big|_{\tilde{m}=m_{j,\eta}(z)},\ j = L, U, \tag{C.5}$$

where the expectation is taken with respect to P_η. Finally, define

$$\zeta_{p,\eta} \equiv \Gamma(r_{L,\eta}^{-1}(z)q(y_L - m_{L,\eta}(z)), r_{U,\eta}^{-1}(z)q(y_U - m_{U,\eta}(z)), p'l_\eta(z)). \tag{C.6}$$

Given these definitions, we give an outline of the general structure of the proof. The proof of Theorem 3.1 proceeds by verifying the conditions of Theorem 5.2.1 in Bickel, Klassen, Ritov, and Wellner (1993), which requires (i) the characterization of the tangent space at P, which we accomplish in Theorem C.1 and (ii) the pathwise weak differentiability of the map $Q \mapsto v(\cdot, \Theta_0(Q))$, which is established by Theorem C.2.

Tangent Space (Theorem C.1)

Step 1: Lemmas C.1-C.5 show that for some neighborhood N of 0, Assumptions 2.3 and 3.2 hold with P_η in place of P for all $\eta \in N$, where $\sqrt{dP_\eta/d\mu} = h_\eta$ defined in (C.2)-(C.3). This means that the restrictions on P in Assumptions 2.3 and 3.2 do not restrict the neighborhood in such a way that affects the tangent space derived in the next step. In this step, we exploit the fact that $\mathcal{Z}$ is determined by the dominating measure μ instead of each distribution in the model.

Step 2: Theorem C.1 then establishes that the tangent space $\dot{\mathbf{S}}$ equals $\mathbf{T} \equiv \{h \in L_\mu^2 : \int h(x)s(x)d\mu(x) = 0\}$ by showing that (i) $\dot{\mathbf{S}} \subseteq \mathbf{T}$ generally and (ii) due to Step 1, $\{P_\eta, \eta \in N\}$ is a regular parametric submodel of $\mathbf{P}$ whose tangent set $\dot{\mathbf{U}} \subset \dot{\mathbf{S}}$ is dense in $\mathbf{T}$ implying $\mathbf{T} \subseteq \dot{\mathbf{S}}$.

Differentiability (Theorem C.2)

Step 1: Lemmas C.1 and C.7 explicitly characterize the pathwise derivatives of $m_{j,\eta}, j = L, U$ along the curve $\eta \mapsto h_\eta$ defined in (C.2)-(C.3).

Step 2: Based on Step 1 and Lemma C.8, Lemma C.9 then characterizes the pathwise derivative of the support function $\upsilon(p, \Theta_0(P_\eta))$ at a point p along the curve $\eta \mapsto h_\eta$ defined in (C.2)-(C.3). Lemmas C.10 and C.11 further show that this pathwise derivative is uniformly bounded and continuous in $(p, \eta) \in \mathbb{S}^\ell \times N$.

Step 3: Based on Step 2, Theorem C.2 first characterizes the pathwise weak derivative of $\rho(P_\eta) = \upsilon(\cdot, \Theta_0(P_\eta))$ on the tangent space of the curve $\eta \mapsto h_\eta$ and further extends it to $\dot{\mathbf{S}}$.

Lemma C.1. *Let $\eta \mapsto h_\eta$ be a curve in L^2_μ defined in* (C.2)-(C.3)*. Suppose Assumption 2.1 holds. Suppose $P \in \mathbf{P}$. Then, there exists a neighborhood N of 0 such that (i) $\int q(y_j - \tilde{m})v_\eta^2(y_L, y_U|z)d\lambda(y_L, y_U) = 0$ has a unique solution at $\tilde{m} = m_{j,\eta}(z)$ on D for $j = L, U$ and for all $\eta \in N$; (ii) For each $(z, \eta) \in \mathcal{Z} \times N$, $m_{\eta,L}$ and $m_{\eta,U}$ are continuously differentiable a.e. on the interior of $\mathcal{Z} \times N$ with bounded derivative. In particular, it holds that*

$$\frac{\partial}{\partial\eta} m_{L,\eta}(z)\Big|_{\eta=\eta_0} = 2r^{-1}_{L,\eta_0}(z) \int q(y_L - m_{L,\eta_0}(z))\dot{v}_{\eta_0}(y_L, y_U|z)v_{\eta_0}(y_L, y_U|z)d\lambda(y_L, y_U) \tag{C.7}$$

$$\frac{\partial}{\partial\eta} m_{U,\eta}(z)\Big|_{\eta=\eta_0} = 2r^{-1}_{U,\eta_0}(z) \int q(y_U - m_{U,\eta_0}(z))\dot{v}_{\eta_0}(y_L, y_U|z)v_{\eta_0}(y_L, y_U|z)d\lambda(y_L, y_U), \tag{C.8}$$

for all $\eta_0 \in N$.

Proof of Lemma C.1. The proof builds on the proof of Theorem 3.1 in Newey and Stoker (1993). By Eq. (C.2), it follows that

$$\int q(y_L - \tilde{m})v_\eta^2(y_L, y_U|z)d\lambda(y_L, y_U) = \int q(y_L - \tilde{m})v_0^2(y_L, y_U|z)d\lambda(y_L, y_U) + 2\eta \int q(y_L - \tilde{m})\bar{v}(x)v_0^2(y_L, y_U|z)d\lambda(y_L, y_U). \tag{C.9}$$

Since $P \in \mathbf{P}$, Assumption 3.2 and Lemma C.2 in Newey (1991) imply that the map $(z, \tilde{m}) \mapsto \int q(y_j - \tilde{m})\bar{v}(x)v_0^2\,(y_L, y_U|z)d\lambda(y_L, y_U)$ is continuously differentiable in on $\mathcal{Z} \times D$ for $j = L, U$. Hence, by (C.9), there is a neighborhood N' of 0 such that the map $(z, \tilde{m}, \eta) \mapsto \int q(y_L - \tilde{m})v_\eta^2(y_L, y_U|z)d\lambda(y_L, y_U)$ is continuously differentiable on $\mathcal{Z} \times D \times N'$ with bounded derivatives. By continuity, we may take a neighborhood N of 0 small enough so that $\int q(y_L - \tilde{m})v_\eta^2(y_L, y_U|z)d\lambda(y_L, y_U) = 0$ admits a unique solution $m_{L,\eta}(z)$ for all $\eta \in N$. A similar argument can be made for $m_{U,\eta}$.

By the implicit function theorem, there is a neighborhood of $(z, 0)$ on which $\nabla_z m_{j,\eta_0}$ and $\frac{\partial}{\partial\eta} m_{j,\eta}(z)|_{\eta=\eta_0}$ exist and are continuous in their arguments on that neighborhood for $j = L, U$. By the compactness of $\mathcal{Z}$, N can be chosen small enough so that $\nabla_z m_{\eta,L}$ and $\frac{\partial}{\partial\eta} m_{j,\eta}(z)|_{\eta=\eta_0}$ are continuous and bounded on $\mathcal{Z} \times N$ and

$$\frac{\partial}{\partial\eta} m_{j,\eta}(z)\Big|_{\eta=\eta_0} = 2r^{-1}_{j,\eta_0}(z) \int q(y_j - m_{j,\eta_0}(z))\dot{v}_{\eta_0}(y_L, y_U|z)v_{\eta_0}(y_L, y_U|z)d\lambda(y_L, y_U),\ j = L, U. \tag{C.10}$$

This completes the proof of the lemma. ■

Lemma C.2. *Let $\eta \mapsto h_\eta$ be a curve in L^2_μ defined in* (C.2)-(C.3)*. Suppose Assumption 2.1 (i) holds. Suppose further that $P \in \mathbf{P}$. Then, there exists a neighborhood N of 0 such that the conditional support of (Y_L, Y_U) given Z is in $D^o \times D^o$, $w(z)f_\eta(z) = 0$ on $\partial\mathcal{Z}$, $P_\eta - a.s.$, $\nabla_z f_\eta / f_\eta(z)$ is continuous a.e., and $\int \|l_\eta(z)\|^2 \phi_\eta^2(z)d\nu(z) < \infty$ for all $\eta \in N$.*

Proof of Lemma C.2. By (C.2), $\{(y_L, y_U) : v_0^2(y_L, y_U|z) = 0\} \subseteq \{(y_L, y_U) : v_\eta^2(y_L, y_U|z) = 0\}$ for all $z \in \mathcal{Z}$ and $\eta \in \mathbb{R}$. This implies $\{(y_L, y_U) : v_\eta^2(y_L, y_U|z) > 0\} \subseteq \{(y_L, y_U) : v_0^2(y_L, y_U|z) > 0\} \subset D^o \times D^o$ for all $z \in \mathcal{Z}$ and $\eta \in \mathbb{R}$, where the last inclusion holds by Assumption 2.3. This establishes the first claim. Similarly, the second claim follows immediately from Eq. (C.2) and Assumption 2.3 (ii).

For the third claim, using Eq. (C.2), we write

$$\frac{\nabla_z f_\eta(z)}{f_\eta(z)} = \frac{\nabla_z f(z)}{f(z)} + \frac{2\eta \nabla_z \bar{\phi}(z)}{1 + 2\eta\bar{\phi}(z)}. \tag{C.11}$$

By Assumption 2.3 (ii), (C.11), and $\bar{\phi}$ being bounded and continuously differentiable in z, $(\eta, z) \mapsto \nabla_z f_\eta(z)/f_\eta(z)$ is continuous. This and Assumption 2.2 in turn imply that the map $(\eta, z) \mapsto \|l_\eta(z)\|^2$ is continuous. Hence, by Assumption 2.1 (i), it achieves a finite maximum on $N \times \mathcal{Z}$ for some N small enough. Therefore, $\int \|l_\eta(z)\|\phi_\eta^2(z)d\nu(z) < \infty$ for all $\eta \in N$. ■

Lemma C.3. *Let $\eta \mapsto h_\eta$ be a curve in L^2_μ defined in* (C.2)-(C.3). *Suppose further that $P \in \mathbf{P}$. Then, there exists a neighborhood N of 0 such that $|r_{L,\eta}(z)| > \bar{\epsilon}$ and $|r_{U,\eta}(z)| > \bar{\epsilon}$, for all $z \in \mathcal{Z}$ and $\eta \in N$.*

Proof of Lemma C.3. By (C.2) and (C.5), $r_{L,\eta}$ can be written as

$$\begin{aligned} r_{L,\eta}(z) &\equiv -\frac{d}{d\tilde{m}} E_\eta\left[q(y_L - \tilde{m})|Z = z\right]\big|_{\tilde{m}=m_{L,\eta}(z)} \\ &= -\frac{d}{d\tilde{m}}\Big(E\left[q(y_L - \tilde{m})|Z = z\right] + 2\eta \int q(y_L - \tilde{m})\bar{v}(x)v_0^2(y_L, y_U|z)d\lambda(y_L, y_U)\Big)\big|_{\tilde{m}=m_{L,\eta}(z)} \\ &= r_L(z) - 2\eta\frac{d}{d\tilde{m}}\int q(y_L - \tilde{m})\bar{v}(x)v_0^2(y_L, y_U|z)d\lambda(y_L, y_U)\big|_{\tilde{m}=m_{L,\eta}(z)}. \end{aligned} \tag{C.12}$$

Since $\bar{v}$ is bounded and continuously differentiable, the second term on the right hand side of (C.12) is well-defined and is bounded because Assumption 3.2 (iii) holds for $P \in \mathbf{P}$. By Assumption 3.2 (i) and Eq. (C.12), we may take a neighborhood N of 0 small enough so that $|r_{L,\eta}(z)| > \bar{\epsilon}$ for all $\eta \in N$ and $z \in \mathcal{Z}$. A similar argument can be made for $r_{U,\eta}(z)$. Thus, the claim of the lemma follows. ■

Lemma C.4. *Let $\eta \mapsto h_\eta$ be a curve in L^2_μ defined in* (C.2)-(C.3). *Suppose further that $P \in \mathbf{P}$. Then, there exists a neighborhood N of 0 such that (i) for any $\varphi : \mathcal{X} \to \mathbb{R}$ that is bounded and continuously differentiable in z with bounded derivatives, $\int \varphi(x)v_\eta^2(y_L, y_U|z)d\lambda(y_L, y_U)$ is continuously differentiable in z on $\mathcal{Z}$ with bounded derivatives; (ii) $\int q(y_L - \tilde{m})\varphi(x)v_\eta^2(y_L, y_U|z)d\lambda(y_L, y_U)$ and $\int q(y_U - \tilde{m})\varphi(x)v_\eta^2(y_L, y_U|z)d\lambda(y_L, y_U)$ are continuously differentiable in $(z, \tilde{m})$ on $\mathcal{Z} \times D$ with bounded derivatives for all $\eta \in N$.*

Proof of Lemma C.4. Let $\varphi : \mathcal{X} \to \mathbb{R}$ be bounded and continuously differentiable in z with bounded derivatives. By (C.2), we may write

$$\int \varphi(x)v_\eta^2(y_L, y_U)d\lambda(y_L, y_U) = \int \varphi(x)v_0^2(y_L, y_U|z)d\lambda(y_L, y_U) + 2\eta\int \varphi(x)\bar{v}(x)v_0^2(y_L, y_U|z)d\lambda(y_L, y_U) \tag{C.13}$$

Note that $\bar{v}$ is bounded and continuously differentiable in z with bounded derivatives. Thus, by Assumption 3.2 (ii), $\int \varphi(x)v_\eta^2(y_L, y_U)d\lambda(y_L, y_U)$ is bounded and continuously differentiable in z with bounded derivatives. Similarly, we may write

$$\int q(y_L - \tilde{m})\varphi(x)v_\eta^2(y_L, y_U|z)d\lambda(y_L, y_U)$$
$$= \int q(y_L - \tilde{m})\varphi(x)v_0^2(y_L, y_U|z)d\lambda(y_L, y_U) + 2\eta \int q(y_L - \tilde{m})\varphi(x)\bar{v}(x)v_0^2(y_L, y_U|z)d\lambda(y_L, y_U). \quad \text{(C.14)}$$

By Assumption 3.2 (iii), $\int q(y_L - \tilde{m})\varphi(x)v_0^2(y_L, y_U)d\lambda(y_L, y_U)$ is bounded and continuously differentiable in z with bounded derivatives. Further, since $\bar{v}$ is bounded and continuously differentiable in z with bounded derivatives, again by Assumption 3.2 (iii), the same is true for the second term on the right hand side of (C.14). The argument for $\int q(y_U - \tilde{m})\varphi(x)v_\eta^2(y_L, y_U)d\lambda(y_L, y_U)$ is similar. Thus, the claim of the lemma follows. ■

Lemma C.5. *Let $\eta \mapsto h_\eta$ be a curve in L_μ^2 defined in* (C.2)-(C.3). *Suppose that $P \in \mathbf{P}$. Then, there exists a neighborhood N of 0 such that $m_{L,\eta}$ and $m_{U,\eta}$ are continuously differentiable a.e. on $\mathcal{Z}$ with bounded derivatives. Further, the maps $(z, \eta) \mapsto \nabla_z m_{L,\eta}(z)$ and $(z, \eta) \mapsto \nabla_z m_{U,\eta}(z)$ are continuous on $\mathcal{Z} \times N$.*

Proof of Lemma C.5. We show the claims of the lemma for $m_{L,\eta}$. By $P \in \mathbf{P}$, Assumption 3.2 (iii) holds, which implies that the maps $(z, \tilde{m}) \mapsto \int q(y_L - \tilde{m})v_0^2(y_L, y_U|z)d\lambda(y_L, y_U)$ and $(z, \tilde{m}) \mapsto \int q(y_L - \tilde{m})\bar{v}(x)v_0^2(y_L, y_U|z)d\lambda(y_L, y_U)$ are continuously differentiable on $\mathcal{Z} \times D$. By (C.14) (with $\varphi(x) = 1$), it then follows that $(z, \tilde{m}, \eta) \mapsto \int_{\mathcal{X}} q(y_L - \tilde{m})v_\eta^2(y_L, y_U|z)d\lambda(y_L, y_U)$ is continuously differentiable on $\mathcal{Z} \times D \times N$ for some N that contains 0 in its interior. Following the argument based on the implicit function theorem in the proof of Theorem 3.1 in Newey and Stoker (1993), it then follows that $\nabla_z m_{L,\eta}$ exists and is continuous on $\mathcal{Z} \times N$. By the compactness of $\mathcal{Z}$, N can be chosen small enough so that $\nabla_z m_{L,\eta}$ is bounded on $\mathcal{Z} \times N$. The argument for $m_{U,\eta}$ is similar. Hence it is omitted. ■

Theorem C.1. *Let Assumptions 2.1-2.2 and 3.1 hold and $P \in \mathbf{P}$. Then, the tangent space of $\mathbf{S} = \{s \in L_\mu^2 : s = \sqrt{dP/d\mu}, P \in \mathbf{P}\}$ at $s \equiv \sqrt{dP/d\mu}$ is given by $\dot{\mathbf{S}} = \{h \in L_\mu^2 : \int h(x)s(x)d\mu(x) = 0.\}$*

Proof of Theorem C.1. Let $\mathbf{T} \equiv \{h \in L_\mu^2 : \int h(x)s(x)d\mu(x) = 0\}$. $\dot{\mathbf{S}} \subseteq \mathbf{T}$ holds by Proposition 3.2.3 in Bickel, Klassen, Ritov, and Wellner (1993).

For the converse: $\mathbf{T} \subseteq \dot{\mathbf{S}}$, it suffices to show that a dense subset of $\mathbf{T}$ is contained in $\dot{\mathbf{S}}$. For this, let $\dot{\mathbf{U}} \equiv \{\dot{h}_0 \in L_\mu^2 : \int \dot{h}_0(x)s(x)d\mu(x) = 0\}$ denote the tangent space of the curve defined in (C.2)-(C.3). By Lemmas C.1-C.5, there is a neighborhood N of 0 for which Assumptions 2.3 and 3.2 hold for all $\eta \in N$. Therefore, $\eta \mapsto h_\eta^2, \eta \in N$ is a regular parametric submodel of $\mathbf{P}$ whose Fréchet derivative at $\eta = 0$ is given by $\dot{h}_0$. Hence, $\dot{\mathbf{U}} \subseteq \dot{\mathbf{S}}$. Further, by Lemma C.7 in Newey (1991) and the argument used in the proof of Theorem 3.1 in Newey and Stoker (1993), $\dot{\mathbf{U}}$ is dense in $\mathbf{T}$. Thus, $\mathbf{T} \subseteq \dot{\mathbf{S}}$. ■

Lemma C.6. *Let $\eta \mapsto h_\eta$ be a curve in L_μ^2 defined in* (C.2)-(C.3). *Suppose Assumption 2.1 holds. Suppose further that $P \in \mathbf{P}$. Then, there is a compact set D' and a neighborhood N of 0 such that D' contains the support of $Y_L - m_{L,\eta_0}(Z)$ and $Y_U - m_{U,\eta_0}(Z)$ in its interior for all $\eta_0 \in N$.*

Proof of Lemma C.6. By Lemma C.2, there exists a neighborhood N' of 0 such that the supports of Y_L and Y_U are contained in the interior of D under P_η for all η in N'. Similarly, by Lemma C.1, there is a neighborhood N'' of 0 such that $m_{L,\eta}(Z), m_{U,\eta}(Z)$ are well defined for all $\eta \in N''$ and their supports are contained in the interior of D respectively. Without loss of generality, let $D = [a, b]$ for some $-\infty < a <$

$b < \infty$ and let $N = N' \cap N''$. Then, the support of $Y_L - m_{L,\eta}(Z)$ is contained in $D' \equiv [a-b, b-a]$ for all $\eta \in N$. A similar argument ensures that the support of $Y_U - m_{U,\eta}(Z)$ is contained in D'. This completes the proof. ■

Lemma C.7. *Let $\eta \mapsto h_\eta$ be a curve in L^2_μ defined in* (C.2)-(C.3)*. Suppose Assumption 2.1 holds. Suppose further that $P \in \mathbf{P}$. Then there is a neighborhood N of 0 such that (i) the functions $(z, \eta_0) \mapsto \frac{\partial}{\partial\eta} m_{L,\eta}(z)\big|_{\eta=\eta_0}$ and $(z,\eta_0) \mapsto \frac{\partial}{\partial\eta} m_{U,\eta}(z)\big|_{\eta=\eta_0}$ are bounded on $\mathcal{Z} \times N$; (ii) For each $z \in \mathcal{Z}$, the maps $\eta_0 \mapsto \frac{\partial}{\partial\eta} m_{L,\eta}(z)\big|_{\eta=\eta_0}$ and $\eta_0 \mapsto \frac{\partial}{\partial\eta} m_{U,\eta}(z)\big|_{\eta=\eta_0}$ are continuous on N.*

Proof of Lemma C.7. By Lemmas C.1, C.3 and C.6, Assumption 2.1 and $\bar{v}$ being bounded, it follows that

$$\Big|\frac{\partial}{\partial\eta} m_{j,\eta}(z)\Big|_{\eta=\eta_0}\Big| \le 2\bar{\epsilon}^{-1} \sup_{u \in D'} |q(u)| \times \sup_{x \in \mathcal{X}} \bar{v}(x) < \infty, \; j = L, U. \tag{C.15}$$

Hence, the first claim follows. Now let $\eta_n \in N$ be a sequence such that $\eta_n \to \eta_0$. Then, by the triangle and Cauchy-Schwarz inequalities,

$$\Big|\frac{\partial}{\partial\eta} m_{L,\eta}(z)\Big|_{\eta=\eta_n} - \frac{\partial}{\partial\eta} m_{L,\eta}(z)\Big|_{\eta=\eta_0}\Big| \le 2|r^{-1}_{L,\eta_n}(z) - r^{-1}_{L,\eta_0}(z)| \sup_{u\in D'} |q(u)| \times \|\dot{v}_{\eta_0}\|_{L^2_\lambda} \|v_{\eta_0}\|_{L^2_\lambda} \tag{C.16}$$

$$+ 2\bar{\epsilon}^{-1} \sup_{u \in D'} |q(u)| \big(\|v_{\eta_n} - v_{\eta_0}\|_{L^2_\lambda} \|\dot{v}_{\eta_n}\|_{L^2_\lambda} + \|\dot{v}_{\eta_n} - \dot{v}_{\eta_0}\|_{L^2_\lambda} \|v_{\eta_0}\|_{L^2_\lambda}\big) \tag{C.17}$$

$$+ 2\bar{\epsilon}^{-1} \int |q(y_L - m_{\eta_n}(z)) - q(y_L - m_{\eta_0}(z))| \dot{v}_{\eta_0}(y_L, y_U|z) v_{\eta_0}(y_L, y_U z) d\lambda(y_L, y_U). \tag{C.18}$$

Note that $|r^{-1}_{L,\eta_n}(z) - r^{-1}_{L,\eta_0}(z)| \to 0$, *a.e.* by (C.12). By the continuous Fréchet differentiability of $\eta \mapsto v_\eta$, it follows that $\|v_{\eta_n} - v_{\eta_0}\|_{L^2_\lambda} = o(1)$ and $\|\dot{v}_{\eta_n} - \dot{v}_{\eta_0}\|_{L^2_\lambda} = o(1)$. Further, the term in (C.18) tends to 0 as $\eta_n \to \eta_0$ by the dominated convergence theorem, almost everywhere continuity of q ensured by Assumption 2.1 (ii), and $m_{\eta_n} \to m_{\eta_0}$, *a.e.* by Lemma C.5. Therefore, $|\frac{\partial}{\partial\eta} m_{L,\eta}(z)|_{\eta=\eta_n} - \frac{\partial}{\partial\eta} m_{L,\eta}(z)|_{\eta=\eta_0}| = o(1)$. The continuity of $\frac{\partial}{\partial\eta} m_{U,\eta}(z)|_{\eta=\eta_0}$ can be shown in the same way. This completes the proof. ■

Lemma C.8. *Let $\eta \mapsto h_\eta$ be a curve in L^2_μ defined in* (C.2)-(C.3)*. Suppose Assumption 3.1 holds. Suppose further that $P \in \mathbf{P}$. Then, there is a neighborhood N of 0 such that*

$$\frac{\partial}{\partial\eta} \int p' l_{\eta_0}(z) 1\{p' l_\eta(z) > 0\} (m_{U,\eta_0}(z) - m_{L,\eta_0}(z)) \phi^2_{\eta_0}(z) d\nu(z)\Big|_{\eta=\eta_0} = 0, \tag{C.19}$$

for all $\eta_0 \in N$.

Proof of Lemma C.8. By (C.11), there is a neighborhood N of 0 such that for all $\eta_0 \in N$,

$$\frac{\partial p' l_\eta(z)}{\partial\eta}\Big|_{\eta=\eta_0} = \frac{4p' \nabla_z \bar{\phi}(z)(1 + \eta_0 \bar{\phi}(z))}{(1 + 2\eta_0 \bar{\phi}(z))^2}. \tag{C.20}$$

Hence, by compactness of $\mathbb{S}^\ell \times \mathcal{Z}$, the map $(p, z, \eta_0) \mapsto \frac{\partial p' l_\eta(z)}{\partial\eta}|_{\eta=\eta_0}$ is uniformly bounded on $\mathbb{S}^\ell \times \mathcal{Z} \times N$ by some constant $M > 0$. Let η_n be a sequence such that $\eta_n \to \eta_0 \in N$. By the mean value theorem, there is $\bar{\eta}_n(z)$ such that

$$\lim_{n\to\infty}\int p'l_{\eta_0}(z)(1\{p'l_{\eta_n}(z)>0\}-1\{p'l_{\eta_0}(z)>0\})(m_{U,\eta_0}(z)-m_{L,\eta_0}(z))\phi^2_{\eta_0}(z)d\nu(z)$$
$$=\lim_{n\to\infty}\int p'l_{\eta_0}(z)(1\{p'l_{\eta_0}(z)>(\eta_0-\eta_n)\frac{\partial p'l_\eta(z)}{\partial\eta}\Big|_{\eta=\bar{\eta}_n(z)}\}-1\{p'l_{\eta_0}(z)>0\})$$
$$\times(m_{U,\eta_0}(z)-m_{L,\eta_0}(z))\phi^2_{\eta_0}(z)d\nu(z)$$
$$\le\lim_{n\to\infty}\int|p'l_{\eta_0}(z)|1\{|p'l_{\eta_0}(z)|\le M|\eta_0-\eta_n|\}|m_{U,\eta_0}(z)-m_{L,\eta_0}(z)|\phi^2_{\eta_0}(z)d\nu(z), \tag{C.21}$$

where the last inequality follows from $\frac{\partial p'l_\eta(z)}{\partial\eta}|_{\eta=\bar{\eta}_n(z)}$ being bounded by M. Therefore, from (C.21), we conclude that

$$\lim_{n\to\infty}\frac{1}{|\eta_n-\eta_0|}\Big|\int p'l_{\eta_0}(z)(1\{p'l_{\eta_n}(z)>0\}-1\{p'l_{\eta_0}(z)>0\})(m_{U,\eta_0}(z)-m_{L,\eta_0})\phi^2_{\eta_0}(z)d\nu(z)\Big|$$
$$\le 2\sup_{y\in D}|y|\times M\times\lim_{n\to\infty}\int 1\{|p'l_{\eta_0}(z)|\le M|\eta_0-\eta_n|\}\phi^2_{\eta_0}(z)d\nu(z)=0, \tag{C.22}$$

where the last equality follows from the monotone convergence theorem and P being in $\mathbf{P}$ ensuring Assumption 3.1 (ii). ■

Lemma C.9. *Let $\eta\mapsto h_\eta$ be a curve in L^2_μ defined in* (C.2)-(C.3). *Suppose Assumptions 2.1-2.2, and 3.1 hold. Suppose further that $P\in\mathbf{P}$. Then, there is a neighborhood N of 0 such that for all $\eta_0\in N$,*

$$\frac{\partial\upsilon(p,\Theta_0(P_\eta))}{\partial\eta}\Big|_{\eta=\eta_0}=2\int\{w(z)p'\nabla_z m_{p,\eta_0}(z)-\upsilon(p,\Theta_0(P_{\eta_0}))+p'l_{\eta_0}(z)\zeta_{p,\eta_0}(y_L,y_U,z)\}\dot{h}_{\eta_0}(x)h_{\eta_0}d\mu(x). \tag{C.23}$$

Proof of Lemma C.9. We first show $\Gamma(\nabla_z m_{L,\eta}(z),\nabla_z m_{U,\eta}(z),p'l_\eta(z))$ is the gradient of $m_{p,\eta}(z)$, $\mu-a.e.$ By Assumption 3.1 (ii), it suffices to show the equality for z such that $p'l_\eta(z)\neq 0$. Write

$$m_{p,\eta}(z+h)-m_{p,\eta}(z)-\Gamma(\nabla_z m_{L,\eta}(z),\nabla_z m_{U,\eta}(z),p'l_\eta(z))'h$$
$$=1\{p'l_\eta(z+h)>0\}[(m_{U,\eta}(z+h)-m_{L,\eta}(z+h))$$
$$-(m_{U,\eta}(z)-m_{L,\eta}(z))-(\nabla_z m_{U,\eta}(z)-\nabla_z m_{L,\eta}(z))'h]$$
$$+(1\{p'l_\eta(z+h)>0\}-1\{p'l_\eta(z)>0\})\times[(m_{U,\eta}(z)-m_{L,\eta}(z))-(\nabla_z m_{U,\eta}(z)-\nabla_z m_{L,\eta}(z))'h]$$
$$+(m_{L,\eta}(z+h)-m_{L,\eta}(z)-\nabla_z m_{L,\eta}(z)'h). \tag{C.24}$$

$\nabla_z m_{U,\eta}$ and $\nabla_z m_{L,\eta}$ being the gradients of $m_{U,\eta}$ and $m_{L,\eta}$, respectively implies that

$$(m_{U,\eta}(z+h)-m_{L,\eta}(z+h))-(m_{U,\eta}(z)-m_{L,\eta}(z))-(\nabla_z m_{U,\eta}(z)-\nabla_z m_{L,\eta}(z))'h=o(\|h\|),$$
$$m_{L,\eta}(z+h)-m_{L,\eta}(z)-\nabla_z m_{L,\eta}(z)'h=o(\|h\|). \tag{C.25}$$

By the continuity of $z\mapsto l_\eta(z)$, ensured by Assumption 2.2, (C.11), and $\bar{\phi}$ being continuously differentiable, there exists $\epsilon>0$ such that $1\{p'l_\eta(z+h)>0\}=1\{p'l_\eta(z)>0\}$ for all h such that $\|h\|<\epsilon$. These results and (C.24) ensure that

$$m_{p,\eta}(z+h)-m_{p,\eta}(z)-\Gamma(\nabla_z m_{L,\eta}(z),\nabla_z m_{U,\eta}(z),p'l_\eta(z))'h=o(\|h\|). \tag{C.26}$$

In what follows, we therefore simply write $\nabla_z m_{p,\eta}(z)=\Gamma(\nabla_z m_{L,\eta}(z),\nabla_z m_{U,\eta}(z),p'l_\eta(z))$.

Next, we show that $\eta\mapsto\nabla_z m_{p,\eta}(z)$ is continuous for almost all $z\in\mathcal{Z}$. By Lemma C.7, $\eta\mapsto\nabla_z m_{L,\eta}$

and $\eta \mapsto \nabla_z m_{U,\eta}$ are continuous. Further, if $\eta_n \to \eta_0$, then

$$\mu(\{x : \lim_{n\to\infty} 1\{p' l_{\eta_n}(z) > 0\} = 1\{p' l_{\eta_0}(z) > 0\}\}) = 1$$

by the continuity of $\eta \mapsto p' l_\eta(z)$ ensured by (C.11) and $\mu(\{x : p' l_{\eta_0}(z) = 0\}) = 0$ by Assumption 3.1 (ii). Hence, $\eta \mapsto \nabla_z m_{p,\eta}(z)$ is continuous *a.e.*

By Theorem 2.1, integration by parts, and (C.4), we may write

$$\upsilon(p, \Theta_0(P_\eta)) = \int w(z) p' \nabla_z m_{p,\eta}(z) \phi^2_{\eta_0}(z) d\nu(z) + 2(\eta - \eta_0) \int w(z) p' \nabla_z m_{p,\eta}(z) \dot{\phi}_{\eta_0}(z) \phi_{\eta_0}(z) d\nu(z) \quad \text{(C.27)}$$

$$= \int p' l_{\eta_0}(z) m_{p,\eta}(z) \phi^2_{\eta_0}(z) d\nu(z) + 2(\eta - \eta_0) \int w(z) p' \nabla_z m_{p,\eta}(z) \dot{\phi}_{\eta_0}(z) \phi_{\eta_0}(z) d\nu(z). \quad \text{(C.28)}$$

This and $\upsilon(p, \Theta_0(P_{\eta_0})) = \int p' l_{\eta_0}(z) m_{p,\eta_0}(z) \phi^2_{\eta_0}(z) d\nu(z)$ by Theorem 2.1 imply

$$\begin{aligned}
&\frac{\partial \upsilon(p, \Theta_0(P_\eta))}{\partial \eta}\Big|_{\eta=\eta_0} \\
&= \lim_{\eta\to\eta_0} \frac{1}{\eta - \eta_0} \int p' l_{\eta_0}(z)(m_{p,\eta}(z) - m_{p,\eta_0}(z)) \phi^2_{\eta_0}(z) d\nu(z) \\
&\quad + 2 \lim_{\eta\to\eta_0} \int w(z) p' \nabla_z m_{p,\eta}(z) \dot{\phi}_{\eta_0}(z) \phi_{\eta_0}(z) d\nu(z), \\
&= \lim_{\eta\to\eta_0} \frac{1}{\eta - \eta_0} \int p' l_{\eta_0}(z)(m_{p,\eta}(z) - m_{p,\eta_0}(z)) \phi^2_{\eta_0}(z) d\nu(z) + 2 \int w(z) p' \nabla_z m_{p,\eta_0}(z) \dot{\phi}_{\eta_0}(z) \phi_{\eta_0}(z) d\nu(z),
\end{aligned} \quad \text{(C.29)}$$

where the second equality follows from w, $\nabla_z m_{p,\eta}$ and $\bar{\phi}_{\eta_0}$ being bounded by Assumption 2.2 and Lemma C.1, which allows us to apply the dominated convergence theorem, and the almost everywhere continuity of $\eta \mapsto \nabla_z m_{p,\eta}(z)$. The first term on the right hand side of (C.29) may be further rewritten as

$$\lim_{\eta\to\eta_0} \frac{1}{\eta - \eta_0} \int p' l_{\eta_0}(z)(m_{p,\eta}(z) - m_{p,\eta_0}) \phi^2_{\eta_0}(z) d\nu(z) \quad \text{(C.30)}$$

$$= \lim_{\eta\to\eta_0} \frac{1}{\eta - \eta_0} \int p' l_{\eta_0}(z) 1\{p' l_{\eta_0}(z) > 0\}(m_{U,\eta}(z) - m_{U,\eta_0}(z)) \phi^2_{\eta_0}(z) d\nu(z) \quad \text{(C.31)}$$

$$+ \lim_{\eta\to\eta_0} \frac{1}{\eta - \eta_0} \int p' l_{\eta_0}(z) 1\{p' l_{\eta_0}(z) \le 0\}(m_{L,\eta}(z) - m_{L,\eta_0}(z)) \phi^2_{\eta_0}(z) d\nu(z) \quad \text{(C.32)}$$

$$+ \lim_{\eta\to\eta_0} \frac{1}{\eta - \eta_0} \int p' l_{\eta_0}(z)(1\{p' l_\eta(z) > 0\} - 1\{p' l_{\eta_0}(z) > 0\})(m_{U,\eta_0}(z) - m_{L,\eta_0}(z)) \phi^2_{\eta_0}(z) d\nu(z). \quad \text{(C.33)}$$

For (C.31), by the mean value theorem, we have

$$\begin{aligned}
&\lim_{\eta\to\eta_0} \frac{1}{\eta - \eta_0} \int p' l_{\eta_0}(z) 1\{p' l_{\eta_0}(z) > 0\}(m_{U,\eta}(z) - m_{U,\eta_0}) \phi^2_{\eta_0}(z) d\nu(z) \\
&\qquad = \lim_{\eta\to\eta_0} \int p' l_{\eta_0}(z) 1\{p' l_{\eta_0}(z) > 0\} \frac{\partial}{\partial \eta} m_{U,\eta}(z)\Big|_{\eta=\bar{\eta}(z,\eta)} \phi^2_{\eta_0}(z) d\nu(z) \\
&\qquad = \int p' l_{\eta_0}(z) 1\{p' l_{\eta_0}(z) > 0\} \frac{\partial}{\partial \eta} m_{U,\eta}(z)\Big|_{\eta=\eta_0} \phi^2_{\eta_0}(z) d\nu(z),
\end{aligned} \quad \text{(C.34)}$$

where the first equality holds for each p for some $\bar{\eta}(p, \eta)$ a convex combination of η and η_0. The second equality follows from Lemmas C.2 and C.7, $\|p\| = 1$, and Assumption 2.3 (ii) justifying the use of the dominated convergence theorem. Similarly, for (C.32), we have

$$\lim_{\eta\to\eta_0}\frac{1}{\eta-\eta_0}\int p'l_{\eta_0}(z)1\{p'l_{\eta_0}(z)\le 0\}(m_{L,\eta}(z)-m_{L,\eta_0})\phi^2_{\eta_0}(z)d\nu(z)$$
$$=\int p'l_{\eta_0}(z)1\{p'l_{\eta_0}(z)\le 0\}\frac{\partial}{\partial\eta}m_{L,\eta}(z)\Big|_{\eta=\eta_0}\phi^2_{\eta_0}(z)d\nu(z)\ . \quad \text{(C.35)}$$

Hence, by (C.29)-(C.34), integration by parts, and (C.33) being 0 by Lemma C.8, we obtain

$$\frac{\partial\upsilon(p,\Theta_0(P_\eta))}{\partial\eta}\Big|_{\eta=\eta_0}=2\int p'l_{\eta_0}(z)\zeta_{p,\eta_0}(y_L,y_U,z)\dot{v}_{\eta_0}(y_L,y_U|z)v_{\eta_0}(y_L,y_U,z)d\lambda(y_L,y_U)\phi^2_{\eta_0}(z)d\nu(z) \quad \text{(C.36)}$$
$$+2\int w(z)p'\nabla_z m_{p,\eta_0}(z)\dot{\phi}_{\eta_0}(z)\phi_{\eta_0}(z)d\nu(z)\ . \quad \text{(C.37)}$$

Using $\dot{h}_{\eta_0}=\dot{v}_{\eta_0}\phi_{\eta_0}+v_{\eta_0}\dot{\phi}_{\eta_0}$, $\int v^2_{\eta_0}d\lambda=1$, and $\int\dot{h}_{\eta_0}h_{\eta_0}d\mu=0$, we may rewrite this as

$$\frac{\partial\upsilon(p,\Theta_0(P_\eta))}{\partial\eta}\Big|_{\eta=\eta_0}=2\int\{w(z)p'\nabla_z m_{p,\eta_0}(z)-\upsilon(p,\Theta_0(P_{\eta_0}))+p'l_{\eta_0}(z)\zeta_{p,\eta_0}(y_L,y_U,z)\}\dot{h}_{\eta_0}(x)h_{\eta_0}d\mu(x)\ . \quad \text{(C.38)}$$

Therefore, the conclusion of the lemma follows. ■

Lemma C.10. *Let $\eta\mapsto h_\eta$ be a curve in L^2_μ defined in* (C.2)-(C.3)*. Suppose Assumptions 2.1-2.2, and 3.1 hold. Suppose further that $P\in\mathbf{P}$. Then, there is a neighborhood N of $\eta=0$ such that the map $(p,\eta_0)\mapsto\frac{\partial\upsilon(p,\Theta_0(P_\eta))}{\partial\eta}\big|_{\eta=\eta_0}$ is uniformly bounded on $\mathbb{S}^\ell\times N$.*

Proof of Lemma C.10. By Lemma C.9 and the triangle inequality,

$$\Big|\frac{\partial\upsilon(p,\Theta_0(P_\eta))}{\partial\eta}\Big|_{\eta=\eta_0}\Big|=2\Big|\int\{w(z)p'\nabla_z m_{p,\eta_0}(z)-\upsilon(p,\Theta_0(P_{\eta_0}))\}\dot{h}_{\eta_0}(x)h_{\eta_0}d\mu(x)\Big|$$
$$+2\Big|\int p'l_{\eta_0}(z)\zeta_{p,\eta_0}(y_L,y_U,z)\dot{h}_{\eta_0}(x)h_{\eta_0}(x)d\mu(x)\Big|. \quad \text{(C.39)}$$

By Assumption 2.2 and $\|p\|=1$, uniformly on N,

$$\|w(z)p'\nabla_z m_{p,\eta}(z)\|_{L^\infty_\mu}\le\sup_{z\in\mathcal{Z}}|w(z)|\times\|\nabla_z m_{p,\eta}(z)\|_{L^\infty_\mu}$$
$$\le\sup_{z\in\mathcal{Z}}|w(z)|\times(\sup_{z\in\mathcal{Z}}|\nabla_z m_{L,\eta}(z)|+\sup_{z\in\mathcal{Z}}|\nabla_z m_{U,\eta}(z)|)<\infty. \quad \text{(C.40)}$$

where the last inequality follows from Lemma C.1. This ensures that $(p,\eta)\mapsto\upsilon(p,\Theta_0(P_\eta))$ is uniformly bounded on $\mathbb{S}^\ell\times N$. We therefore have

$$\Big|\int\{w(z)p'\nabla_z m_{p,\eta_0}(z)-\upsilon(p,\Theta_0(P_{\eta_0}))\}\dot{h}_{\eta_0}(x)h_{\eta_0}d\mu(x)\Big|$$
$$\le\sup_{(p,\eta_0,z)\in\mathbb{S}^\ell\times N\times\mathcal{Z}}|w(z)p'\nabla_z m_{p,\eta_0}(z)-\upsilon(p,\Theta_0(P_{\eta_0}))|\|\dot{h}_{\eta_0}\|_{L^2_\mu}\|h_{\eta_0}\|_{L^2_\mu}<\infty. \quad \text{(C.41)}$$

Further, by Assumption 2.1 (ii) and Lemmas C.3 and C.6, it follows that

$$\sup_{x\in\mathcal{X}}|\zeta_{p,\eta_0}(x)|\le 2\bar{\epsilon}^{-1}\sup_{u\in D'}|q(u)|<\infty$$

for all $(p,\eta_0)\in\mathbb{S}^\ell\times N$. Therefore, by the triangle and Cauchy-Schwarz inequalities and Lemma C.2, we have

$$\Big|\int p'l_{\eta_0}(z)\zeta_{p,\eta_0}(y_L,y_U,z)\dot h_{\eta_0}(x)h_{\eta_0}(x)d\mu(x)\Big|$$
$$\le 2\bar\epsilon^{-1}\sup_{u\in D'}q(u)\int\|l_{\eta_0}(z)\|\dot h_{\eta_0}(x)h_{\eta_0}(x)d\mu(x)\le 2\bar\epsilon^{-1}\sup_{u\in D'}q(u)\|\dot h_{\eta_0}\|_{L^2_\mu}E_{\eta_0}[\|l_{\eta_0}(z)\|^2]^{1/2}<\infty. \quad (C.42)$$

By (C.39), (C.41), and (C.42), the conclusion of the lemma follows. ■

Lemma C.11. *Let $\eta\mapsto h_\eta$ be a curve in L^2_μ defined in* (C.2)-(C.3). *Suppose Assumptions 2.1-2.2, and 3.1 hold. Suppose further that $P\in\mathbf{P}$. Then, there is a neighborhood N of* 0 *such that the map $(p,\eta_0)\mapsto\frac{\partial\upsilon(p,\Theta_0(P_\eta))}{\partial\eta}\big|_{\eta=\eta_0}$ is continuous on $\mathbb{S}^\ell\times N$.*

Proof of Lemma C.11. Let (p_n,η_n) be a sequence such that $(p_n,\eta_n)\to(p,\eta_0)$. For each $(p,\eta,z)\in\mathbb{S}^\ell\times N\times\mathcal{Z}$, let $\gamma_{p,\eta}(z)\equiv w(z)p'\nabla_z m_{p,\eta}(z)-\upsilon(p,\Theta_0(P_\eta))$. We first show that $(p,\eta)\mapsto\gamma_{p,\eta}(z)$ and $(p,\eta)\mapsto\zeta_{p,\eta}(x)$ are continuous *a.e.* By Lemma C.7, $\eta\mapsto\nabla_z m_{L,\eta}(z)$ and $\eta\mapsto\nabla_z m_{U,\eta}(z)$ are continuous for every $z\in\mathcal{Z}$. Further, if $(p_n,\eta_n)\to(p,\eta_0)$, then

$$\mu(\{x:\lim_{n\to\infty}1\{p_n'l_{\eta_n}(z)>0\}=1\{p'l_{\eta_0}(z)>0\}\})=1$$

by the continuity of $(p,\eta)\mapsto p'l_\eta(z)$ implied by (C.11) and $\mu(\{x:p'l_{\eta_0}(z)=0\})=0$ by Assumption 3.1 (ii). Hence, $(p,\eta)\mapsto w(z)p'\nabla_z m_{p,\eta}(z)$ is continuous *a.e.* Note that by (C.4), $\upsilon(p,\Theta_0(P_\eta))=\int w(z)p'\nabla_z m_{p,\eta}(z)(1+2(\eta-\eta_0)\bar\phi_{\eta_0}(z))\phi^2_{\eta_0}(z)d\nu(z)$. Hence, as $(p_n,\eta_n)\to(p,\eta_0)$, it follows that

$$\lim_{n\to\infty}\upsilon(p_n,\Theta_0(P_{\eta_n}))=\lim_{n\to\infty}\int w(z)p_n'\nabla_z m_{p_n,\eta_n}(z)(1+2(\eta_n-\eta_0)\bar\phi_{\eta_0}(z))\phi^2_{\eta_0}(z)d\nu(z)$$
$$=\int\lim_{n\to\infty}w(z)p_n'\nabla_z m_{p_n,\eta_n}(z)(1+2(\eta_n-\eta_0)\bar\phi_{\eta_0}(z))\phi^2_{\eta_0}(z)d\nu(z)=\upsilon(p,\Theta_0(P_{\eta_0})), \quad (C.43)$$

where the second equality follows from $w(z)p'\nabla_z m_{p,\eta}(z)$ and $\bar\phi(z)$ being bounded on $\mathbb{S}^\ell\times N\times\mathcal{Z}$ and an application of the dominated convergence theorem, while the last equality follows from the continuity of $(p,\eta)\mapsto w(z)p'\nabla_z m_{p,\eta}(z)$ for almost all z. Hence, $(p,\eta)\mapsto\gamma_{p,\eta}(z)$ is continuous *a.e.*

The maps $(p,\eta)\mapsto r^{-1}_{j,\eta}(z)q(y_j-m_{j,\eta}(z)),j=L,U$ are continuous for almost all x by Assumption 2.1 (ii), (C.12), and $\eta\mapsto m_{j,\eta}(z)$ being continuous for almost all z for $j=L,U$ by Lemma C.7. Since $(p,\eta)\mapsto 1\{p'l_\eta(z)>0\}$ is continuous for almost all z as shown above, it then follows that $(p,\eta)\mapsto\zeta_{p,\eta}(x)$ is continuous for almost all x.

Given these results, we show $\frac{\partial\upsilon(p_n,\Theta_0(P_\eta))}{\partial\eta}|_{\eta=\eta_n}-\frac{\partial\upsilon(p,\Theta_0(P_\eta))}{\partial\eta}|_{\eta=\eta_0}\to 0$ as $\eta_n\to\eta_0$. Toward this end, we first note that

$$\frac{\partial\upsilon(p_n,\Theta_0(P_\eta))}{\partial\eta}\Big|_{\eta=\eta_n}-\frac{\partial\upsilon(p,\Theta_0(P_\eta))}{\partial\eta}\Big|_{\eta=\eta_0} \quad (C.44)$$
$$=2\int\gamma_{p_n,\eta_n}(z)\dot h_{\eta_n}(x)h_{\eta_n}d\mu(x)-2\int\gamma_{p,\eta_0}(z)\dot h_{\eta_0}(x)h_{\eta_0}d\mu(x) \quad (C.45)$$
$$+2\int p_n'l_{\eta_n}(z)\zeta_{p_n,\eta_n}(x)\dot h_{\eta_n}(x)h_{\eta_n}d\mu(x)-2\int p'l_{\eta_0}(z)\zeta_{p,\eta_0}(x)\dot h_{\eta_0}(x)h_{\eta_0}d\mu(x). \quad (C.46)$$

By Lemma C.10, the Cauchy-Schwarz and triangle inequalities,

$$|\int \gamma_{p_n,\eta_n}(z)\dot{h}_{\eta_n}(x)h_{\eta_n}d\mu(x) - \int \gamma_{p,\eta_0}(z)\dot{h}_{\eta_0}(x)h_{\eta_0}d\mu(x)| \tag{C.47}$$

$$\leq \sup_{(p,\eta,z)\in\mathbb{S}^\ell\times N\times\mathcal{Z}} |\gamma_{p,\eta}(z)|(\|\dot{h}_{\eta_n}\|_{L^2_\mu}\|h_{\eta_n}-h_{\eta_0}\|_{L^2_\mu} + \|\dot{h}_{\eta_n}-\dot{h}_{\eta_0}\|_{L^2_\mu}\|h_{\eta_0}\|_{L^2_\mu}) \tag{C.48}$$

$$+|\int \{\gamma_{p_n,\eta_n}(z)-\gamma_{p,\eta_0}(z)\}\dot{h}_{\eta_0}(z)h_{\eta_0}(z)\mu(x) = o(1), \tag{C.49}$$

where the last equality follows from $\eta \mapsto h_\eta$ being continuously Fréchet differentiable, $(p,\eta,z)\mapsto\gamma_{p,\eta}(z)$ being bounded on $\mathbb{S}^\ell\times N\times\mathcal{Z}$ as shown in Lemma C.10, the dominated convergence theorem, and $(p,\eta)\mapsto \gamma_{p,\eta}(z)$ being continuous *a.e.*

Further, we may write (C.46) as

$$\int p_n' l_{\eta_n}(z)\zeta_{p_n,\eta_n}(x)\dot{h}_{\eta_n}(x)h_{\eta_n}d\mu(x) - \int p' l_{\eta_0}(z)\zeta_{p,\eta_0}(x)\dot{h}_{\eta_0}(x)h_{\eta_0}d\mu(x) \tag{C.50}$$

$$= \int p_n' l_{\eta_n}(z)\zeta_{p_n,\eta_n}(x)(\dot{h}_{\eta_n}(x)h_{\eta_n}(x) - \dot{h}_{\eta_0}(x)h_{\eta_0}(x))d\nu(x) \tag{C.51}$$

$$+ \int p_n' l_{\eta_n}(z)(\zeta_{p_n,\eta_n}(x)-\zeta_{p,\eta_0}(x))\dot{h}_{\eta_0}(x)h_{\eta_0}(x)d\mu(x) \tag{C.52}$$

$$+ \int (p_n' l_{\eta_n}(z) - p' l_{\eta_0}(z))\zeta_{p,\eta_0}(x)\dot{h}_{\eta_0}(x)h_{\eta_0}(x)d\mu(x). \tag{C.53}$$

By Assumptions 2.2 and 2.3 (ii), (C.11), and $\bar{\phi}$ being continuously differentiable, $(\eta,z)\mapsto\|l_\eta(z)\|$ is continuous on $N\times\mathcal{Z}$. Hence, it achieves a finite maximum on $N\times\mathcal{Z}$. Further, by Lemmas C.3, C.6 and Assumption 2.1 (ii), $\sup_{x\in\mathcal{X}}|\zeta_{p,\eta_0}(x)|\leq 2\bar{\epsilon}^{-1}\sup_{u\in D'}|q(u)|<\infty$ for all $(p,\eta_0)\in\mathbb{S}^\ell\times N$. By Cauchy-Schwarz inequality, and $\|p_n\|\leq 1$ for all n, it then follows that

$$\int p_n' l_{\eta_n}(z)(\dot{h}_{\eta_n}(x)h_{\eta_n}(x) - \dot{h}_{\eta_0}(x)h_{\eta_0}(x))d\nu(x)$$
$$\leq \sup_{(\eta,z)\in N\times\mathcal{Z}}\|l_\eta(z)\| \sup_{(p,\eta)\in\mathbb{S}^\ell\times N}\sup_{x\in\mathcal{X}}|\zeta_{p,\eta}(x)|(\|\dot{h}_{\eta_n}\|_{L^2_\mu}\|h_{\eta_n}-h_{\eta_0}\|_{L^2_\mu}+\|\dot{h}_{\eta_n}-\dot{h}_{\eta_0}\|_{L^2_\mu}\|h_{\eta_0}\|_{L^2_\mu}) = o(1), \tag{C.54}$$

where the last equality follows from $\eta\mapsto h_\eta$ being continuously Fréchet differentiable. Further, by the almost everywhere continuity of $(p,\eta)\mapsto p'l_\eta(\zeta_{p,\eta}-\zeta_{p,\eta_0})$,

$$\lim_{n\to\infty}\int p_n' l_{\eta_n}(z)(\zeta_{p_n,\eta_n}(x)-\zeta_{p,\eta_0}(x))\dot{h}_{\eta_0}(x)h_{\eta_0}(x)d\mu(x)$$
$$= \int \lim_{n\to\infty} p_n' l_{\eta_n}(z)(\zeta_{p_n,\eta_n}(x)-\zeta_{p,\eta_0}(x))\dot{h}_{\eta_0}(x)h_{\eta_0}(x)d\mu(x) = 0. \tag{C.55}$$

where the first equality follows from the dominated convergence theorem. Finally, again by the dominated convergence theorem,

$$\lim_{n\to\infty}\int (p_n' l_{\eta_n}(z) - p' l_{\eta_0}(z))\zeta_{p,\eta_0}(x)\dot{h}_{\eta_0}(x)h_{\eta_0}(x)d\mu(x)$$
$$= \int \lim_{n\to\infty}(p_n' l_{\eta_n}(z) - p' l_{\eta_0}(z))\zeta_{p,\eta_0}(x)\dot{h}_{\eta_0}(x)h_{\eta_0}(x)d\mu(x) = 0. \tag{C.56}$$

By (C.44)-(C.56), we conclude that $\frac{\partial\upsilon(p_n,\Theta_0(P_\eta))}{\partial\eta}|_{\eta=\eta_n} - \frac{\partial\upsilon(p,\Theta_0(P_\eta))}{\partial\eta}|_{\eta=\eta_0}\to 0$ as $\eta_n\to\eta_0$. This establishes the claim of the lemma. ■

Theorem C.2. *Suppose Assumptions 2.1-2.3, and 3.2 hold. Then, the mapping $\rho : \mathbf{P} \to \mathcal{C}(\mathbb{S}^\ell)$ pointwise defined by $\rho(h_\eta)(p) \equiv \upsilon(p, \Theta_0(P_\eta))$ for $h_\eta = \sqrt{dP_\eta/d\mu}$ is then pathwise weak differentiable at $h_0 \equiv \sqrt{dP_0/d\mu}$. Moreover, the derivative $\dot\rho : \dot{\mathbf{P}} \to \mathcal{C}(\mathbb{S}^\ell)$ satisfies:*

$$\dot\rho(\dot h_0)(p) = 2\int \{w(z)p'\nabla_z m_p(z) - \upsilon(p, \Theta_0(P_0)) + p'l(z)\zeta_p(x)\}\dot h_0(x)h_0(x)d\mu(x). \tag{C.57}$$

Proof of Theorem C.2. We first show that $\partial\rho(P_\eta)/\partial\eta\big|_{\eta=0}$ is the pathwise weak derivative of ρ. For this, note that $\dot\rho(\dot h_0) \in \mathcal{C}(\mathbb{S}^\ell)$ for all $\dot h_0 \in \dot{\mathbf{S}}$ as implied by Lemmas C.9 and C.11. Linearity of $\dot\rho$ is immediate, while continuity follows by noting that by the Cauchy-Schwarz inequality and $\|p\| = 1$,

$$\begin{aligned}&\sup_{\|\dot h_0\|_{L^2_\mu}=1} \|\dot\rho(\dot h_0)\|_\infty \\ &\quad\le 2\{\sup_{p\in\mathbb{S}^\ell} \|w(z)p'\nabla_z m_p(z) - \upsilon(p, \Theta_0(P))\|_{L^\infty_\mu} + \sup_{z\in\mathcal{Z}} \|l(z)\| \times \sup_{u\in\mathcal{D}'} |q(u)|\}\|\dot h_0\|_{L^2_\mu}\|h_0\|_{L^2_\mu} < \infty,\end{aligned} \tag{C.58}$$

where we exploited (C.40), Assumption 2.1 (ii), and the fact that $z \mapsto \|l(z)\|$ being continuous hence achieves a finite maximum on $\mathcal{Z}$ by Assumptions 2.1 (i), 2.2, and $P \in \mathbf{P}$. Let $\eta \mapsto h_\eta$ be a curve in L^2_μ defined in (C.2)-(C.3). For each $p \in \mathbb{S}^\ell$, by the mean value theorem,

$$\lim_{\eta_0\to 0}\int_{\mathbb{S}^\ell} \frac{\upsilon(p, P_{\eta_0}) - \upsilon(p, P_0)}{\eta_0} dB(p) = \lim_{\eta_0\to 0}\int_{\mathbb{S}^\ell} \frac{\partial\upsilon(p, P_\eta)(p)}{\partial\eta}\Big|_{\eta=\bar\eta(p,\eta_0)} dB(p) \tag{C.59}$$

$$= \int_{\mathbb{S}^\ell} \frac{\partial\upsilon(p, P_\eta)(p)}{\partial\eta}\Big|_{\eta=0} dB(p) = \int_{\mathbb{S}^\ell} \dot\rho(\dot h_0)(\tau)dB(p)\ , \tag{C.60}$$

where the first equality holds at each p for some $\bar\eta(p, \eta_0)$ a convex combination of η_0 and 0. The second equality in turn follows by Lemma C.10 justifying the use of the dominated convergence theorem, while the final equality follows by Lemma C.11 and the definition of $\dot\rho : \dot{\mathbf{P}} \to \mathcal{C}(\mathbb{S}^\ell)$. Eqs. (C.59)-(C.60) hold for any $\dot h_0$ in the tangent space $\dot{\mathbf{U}}$ of the curve defined in (C.2)-(C.3). As discussed in the proof of Theorem C.1, $\dot{\mathbf{U}}$ is dense in $\dot{\mathbf{S}}$. Since $\dot\rho$ is continuous, Eqs. (C.59)-(C.60) then hold for any $\dot h_0 \in \dot{\mathbf{P}}$. This completes the proof. ■

Proof of Theorem 3.1. Let $\mathbf{B} \equiv \mathcal{C}(\mathbb{S}^\ell)$ and let $\mathbf{B}^*$ be the set of finite Borel measures on $\mathbb{S}^\ell$, which is the norm dual of $\mathbf{B}$ by Corollary 14.15 in Aliprantis and Border (2006). By Theorem C.2, ρ has pathwise weak derivative $\dot\rho$. For each $B \in \mathbf{B}^*$, define

$$\dot\rho^T(B)(x) \equiv \int_{\mathbb{S}^\ell} 2\{w(z)p'\nabla_z m_p(z) - \upsilon(p, \Theta_0(P)) + p'l(z)\zeta_p(x)\}h_0(x)dB(p). \tag{C.61}$$

We show that (i) $\dot\rho^T$ is well defined for any $B \in \mathbf{B}^*$, (ii) $\dot\rho^T(B) \in \dot{\mathbf{S}}$ and finally (iii) $\dot\rho^T$ is the adjoint operator of $\dot\rho$.

We first note that $(p, z) \mapsto p'l(z)$ is continuous in z for each p by Assumption 2.2 and measurable in p for each z. Thus, $(p, x) \mapsto p'l(z)$ is jointly measurable by Lemma 4.51 in Aliprantis and Border (2006). This implies the joint measurability of $(p, x) \mapsto 1\{p'l(z) > 0\}$. A similar argument also ensures the joint measurability of $p'\nabla_z m_L(z)$ and $p'\nabla_z m_U(z)$. By the joint measurability of $(p, x) \mapsto (w(z), 1\{p'l(z) > 0\}, p'\nabla_z m_L(z), p'\nabla_z m_U(z))$ and Assumption 2.2, $(p, x) \mapsto w(z)p'\nabla_z m_p(z)$ is jointly measurable. By the proof of Theorem 2.1, $\upsilon(p, \Theta_0(p))$ is differentiable in p and is therefore continuous, implying $(p, x) \mapsto \upsilon(p, \Theta_0(p))$ is jointly measurable. Further, r_L and r_U are measurable by $P \in \mathbf{P}$ satisfying Assumption 3.2 (iii). $q(y_L - m_L(z)), q(y_U - m_U(z))$ are measurable by Assumption 2.1 and

$P \in \mathbf{P}$ satisfying Assumption 2.3 (iv). Hence, $(p,x) \mapsto \zeta_p(x)$ is jointly measurable. Therefore, the map $(p,x) \mapsto (w(z)p'\nabla_z m_p(z), \upsilon(p, \Theta_0(p), p'l(z), \zeta_p(x), h_0(x))'$ is jointly measurable by Lemma 4.49 in Aliprantis and Border (2006). Hence, the map

$$(p,x) \mapsto 2\{w(z)p'\nabla_z m_p(z) - \upsilon(p, \Theta_0(P_0)) + p'l(z)\zeta_p(x)\} \tag{C.62}$$

is jointly measurable by the measurability of the composite map.

Moreover, for $|B|$ the total variation of the measure B, by (C.40), we have

$$\begin{aligned}\int \Big(\int_{\mathbb{S}^\ell} 2\{w(z)p'\nabla_z m_p(z) - \upsilon(p, \Theta_0(P_0))\} h_0(z) dB(p)\Big)^2 d\mu(x) \\ \le 16 \times \sup_{p \in \mathbb{S}^\ell} \|w(z)p'\nabla_z m_{p,\eta}(z)\|^2_{L^\infty_\mu} \times |B|^2 < \infty.\end{aligned} \tag{C.63}$$

Further,

$$\int \Big(\int_{\mathbb{S}^\ell} 2p'l(z)\zeta_p(x)h_0(x)) dB(p)\Big)^2 d\mu(x) \le 16 \int |p'l(z)|^2 h_0(x)^2 d\mu(x) \times \bar{\epsilon}^{-2} \times \sup_{u \in D'} |q(u)|^2 \times |B|^2 < \infty, \tag{C.64}$$

by Assumption 2.1 and $P \in \mathbf{P}$ satisfying Assumptions 2.3 and 3.2. Therefore, $\dot{\rho}^T(B) \in L^2_\mu$ for each $B \in \mathbf{B}^*$.

By Fubini's theorem and Assumption 2.1 (iv), we have

$$\begin{aligned}\int \int_{\mathbb{S}^\ell} 2\{w(z)p'\nabla_z m_p(z) - \upsilon(p, \Theta_0(P_0)) + p'l(z)\zeta_p(x)\} h_0(x) dB(p) h_0(x) d\mu(x) \\ = \int_{\mathbb{S}^\ell} \int 2\{w(z)p'\nabla_z m_p(z) - \upsilon(p, \Theta_0(P_0)) + p'l(z)\zeta_p(x)\} h_0^2(x) d\mu(x) dB(p) = 0,\end{aligned} \tag{C.65}$$

where we exploited $\upsilon(p, \Theta_0(P_0)) = E[w(Z)p'\nabla_z m_p(Z)]$ and $E[\zeta_p(x)|Z = z] = 1\{p'l(z) \le 0\}E[q(Y_L - m_L(Z))|Z = z] + 1\{p'l(z) > 0\}E[q(Y_U - m_U(Z))|Z = z] = 0$, $P - a.s.$ Thus, by Theorem C.1 and (C.61), $\dot{\rho}^T(B) \in \dot{\mathbf{S}}$ for all $B \in \mathbf{B}^*$. Further, for any $\dot{h}_0 \in \dot{\mathbf{S}}$, again by interchanging the order of integration

$$\int_{\mathbb{S}^\ell} \dot{\rho}(\dot{h}_0)(p) dB(p) = \int_{\mathcal{X}} \dot{h}_0(x) \dot{\rho}^T(B)(x) d\mu(x), \tag{C.66}$$

which ensures that $\dot{\rho}^T : \mathbf{B}^* \to \dot{\mathbf{P}}$ is the adjoint of $\dot{\rho} : \dot{\mathbf{P}} \to \mathbf{B}$.

Since $\dot{\mathbf{S}}$ is linear by Theorem C.1, Theorem C.2 and Theorem 5.2.1 in Bickel, Klassen, Ritov, and Wellner (1993) establishes that

$$\begin{aligned}\mathrm{Cov}\Big(\int_{\mathbb{S}^\ell} \mathbb{G}(p) dB_1(p), \int_{\mathbb{S}^\ell} \mathbb{G}(p) dB_2(p)\Big) &= \frac{1}{4} \int_{\mathcal{X}} \dot{\rho}^T(B_1)(x) \dot{\rho}^T(B_2)(x) d\mu(x) \\ &= \int_{\mathbb{S}^\ell} \int_{\mathbb{S}^\ell} E[\{w(z)p'\nabla_z m_p(z) - \upsilon(p, \Theta_0(P_0)) + p'l(z)\zeta_p(x)\} \\ &\quad \times \{w(z)q'\nabla_z m_q(z) - \upsilon(q, \Theta_0(P_0)) + q'l(z)\zeta_q(x)\}] dB_1(p) dB_2(q),\end{aligned}$$

for any $B_1, B_2 \in \mathbf{B}^*$ by Fubini's theorem. Letting B_1 and B_2 be the degenerate measures at p and q in

(C.67), we obtain

$$
\begin{aligned}
&\mathrm{Cov}(\mathbb{G}(p), \mathbb{G}(q)) \\
&\quad = E[\{w(z)p'\nabla_z m_p(z) - \upsilon(p, \Theta_0(P_0)) + p'l(z)\zeta_p(x)\}\{w(z)q'\nabla_z m_q(z) - \upsilon(q, \Theta_0(P_0)) + q'l(z)\zeta_q(x)\}].
\end{aligned} \tag{C.67}
$$

Therefore, the efficient influence function ψ_p is as given in (33). This establishes the claim of the theorem. ■

Appendix D: Proof of Theorem 4.1.

In this appendix, we establish Theorem 4.1. Throughout, let $Y_{p,i} \equiv 1\{p'l(Z_i) \leq 0\}Y_{L,i} + 1\{p'l(Z_i) > 0\}Y_{U,i}$, and let $\bar{\upsilon}_n(p) \equiv \frac{1}{n}\sum_{i=1}^n p'\hat{l}_{i,h}(Z_i)Y_{p,i}$. The proof of Theorem 4.1 proceeds by decomposing $\sqrt{n}(\hat{\upsilon}_n(p) - \upsilon(p, \Theta_0(P_0)))$ as follows:

$$
\begin{aligned}
&\sqrt{n}(\hat{\upsilon}_n(p) - \upsilon(p, \Theta_0(P))) \\
&\qquad = \sqrt{n}(\hat{\upsilon}_n(p) - \bar{\upsilon}_n(p)) + \sqrt{n}(\bar{\upsilon}_n(p) - E[\bar{\upsilon}_n(p)]) + \sqrt{n}(E[\bar{\upsilon}_n(p)] - \upsilon(p, \Theta_0(P))) \\
&\qquad \equiv G_{1n}(p) + G_{2n}(p) + G_{3n}(p).
\end{aligned} \tag{D.1}
$$

G_{1n} is the difference between $\hat{\upsilon}_n$ and the infeasible trimmed estimator $\bar{\upsilon}_n$, which requires the knowledge of $Y_{p,i}$. G_{2n} represents the infeasible estimator centered at its expected value, and G_{3n} is the asymptotic bias of $\bar{\upsilon}_n$. The auxiliary lemmas are then used to show the following results:

Step 1: Lemma D.2 shows $G_{1n} = o_p(1)$ uniformly in $p \in \mathbb{S}^\ell$, while Lemma D.5 shows $G_{3n} = o(1)$ uniformly in $p \in \mathbb{S}^\ell$.

Step 2: Using the result of Lemma D.1, Lemmas D.3 and D.4 then establish that $G_{2n} = \frac{1}{\sqrt{n}}\sum_{i=1}^n \psi_p(X_i) + o_p(1)$ uniformly in $p \in \mathbb{S}^\ell$, and Lemma D.6 establishes that $\{\psi_p : p \in \mathbb{S}^\ell\}$ is a P-Donsker class.

Step 3: Combining Steps 1-2 and (D.1) gives the main claim of Theorem 4.1.

Before proceeding further, we introduce one more piece of notation. For each $p \in \mathbb{S}^\ell$, define

$$
p_n^{(1)}(x, x'; p) \equiv \frac{-1}{2}\Big(\frac{1}{h}\Big)^{\ell+1} p'\nabla_z K\Big(\frac{z - z'}{h}\Big)(g_p^{(1)}(x) - g_p^{(1)}(x')), \quad g_p^{(1)}(x) \equiv \frac{w(z)}{f(z)} y_{p,i}\tau_{n,i} \tag{D.2}
$$

$$
p_n^{(2)}(x, x'; p) \equiv \frac{1}{2}\Big(\frac{1}{h}\Big)^{\ell} K\Big(\frac{z - z'}{h}\Big)(g_p^{(2)}(x) + g_p^{(2)}(x')), \quad g_p^{(2)}(x) \equiv w(Z_i)\frac{p'\nabla_z f(z)}{f(z)^2} y_{p,i}\tau_{n,i}. \tag{D.3}
$$

For each $k \in \{1, 2\}$, we then define $r_n^{(k)}(x_i; p) = E[p_n^{(k)}(X_i, X_j; p)|X_i = x_i]$.

Lemma D.1. *Suppose Assumptions 2.2, 2.3, 4.2 and 4.3 hold. For each $k \in \{1, 2\}$ and $n \in \mathbb{N}$, let $\mathcal{H}_n^{(k)} \equiv \{\tilde{p}_n^{(k)}/b_n : \mathcal{X} \times \mathcal{X} \to \mathbb{R} : \tilde{p}_n^{(k)}(x, x'; p), p \in \mathbb{B}^\ell\}$ and $\mathcal{G}_n^{(k)} \equiv \{\tilde{q}_n^{(k)}/b_n : \mathcal{X} \times \mathcal{X} \to \mathbb{R} : \tilde{q}_n(x, x'; p) = \tilde{p}_n^{(k)}(x, x'; p) - \tilde{r}_n^{(k)}(x, p) - \tilde{r}_n^{(k)}(x'; p) - E[\tilde{r}_n^{(k)}(x; p)], p \in \mathbb{B}^\ell\}$, where $\tilde{r}_n^{(k)}(x_i; p) \equiv E[\tilde{p}_n(X_i, X_j; p)|X_i = x_i]$ and $\mathbb{B}^\ell \equiv \{p \in \mathbb{R}^\ell : \|p\| \leq 1\}$ is the unit ball in $\mathbb{R}^\ell$. Then, $\mathcal{H}_n^{(k)}$ and $\mathcal{G}_n^{(k)}$ are Euclidean in the sense of Pakes and Pollard (1989) and Sherman (1994a) with envelope functions $H^{(k)} : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ and $G^{(k)} : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ such that $E[H^{(k)}(X_i, X_j)^2] < \infty$ and $E[G^{(k)}(X_i, X_j)^2] < \infty$ for $k = 1, 2$.*

Proof of Lemma D.1. For any (fixed) function $g : \mathcal{X}\times\mathcal{X} \to \mathbb{R}^\ell$, let $\mathcal{F}_{g,n} \equiv \{\phi/b_n : \mathcal{X}\times\mathcal{X} \to \mathbb{R} : \phi(x,x,p) = p'g(x,x')\frac{\tau_n(z)}{f(z)}, p \in \mathbb{B}^\ell\}$. By the Cauchy-Schwarz inequality and Assumption 4.3, for any $p,q \in \mathbb{B}^\ell$, we then have $|\phi(x,x',p) - \phi(x,x',p')|/b_n \le \|g(x,x')\|\|p-q\|$. Hence, by Lemma 2.13 in Pakes and Pollard (1989), $\mathcal{F}_{g,n}$ is Euclidean with the envelope function $F_g(x,x') = g(x,x')'p_0 + M\|g(x,x')\|$ for some $p_0 \in \mathbb{B}^\ell$, where $M = 2\sqrt{\ell}\sup_{p\in\mathbb{B}^\ell}\|p-p_0\|$, which can be further bounded from above by $4\sqrt{\ell}$. Hence, we may take the envelope function as $F_g(x,x') = (1+4\sqrt{\ell})\|g(x,x')\|$.

By Lemma 2.4 in Pakes and Pollard (1989), the class of sets $\{x \in \mathcal{X} : p'l(z) > 0\}$ is a VC-class, which in turn implies that the function classes $\mathcal{F}_{\phi_1} \equiv \{\phi : \mathcal{X}\times\mathcal{X} \to \mathbb{R} : \phi(x,x',p) = 1\{p'l(z) > 0\}\}$ and $\mathcal{F}_{\phi_2} \equiv \{\phi : \mathcal{X}\times\mathcal{X} \to \mathbb{R} : \phi(x,x',p) = 1\{p'l(z') > 0\}\}$ are Euclidean, where $x = (y_L, y_U, z)$ and $x' = (y_L', y_U', z')$ with envelope function $F_{\phi_j}(x,x') = 1, j = 1,2$.

Note that we may write

$$\tilde{p}_n^{(1)}(x,x';p) = -p'\nabla_z K\left(\frac{z-z'}{h}\right)\Big[\frac{w(z)}{f(z)}\big((y_U-y_L)1\{p'l(z)>0\}+y_L\big)\tau_n(z) - \frac{w(z')}{f(z')}\big((y_U'-y_L')1\{p'l(z')>0\}+y_L'\big)\tau_n(z)\Big]. \quad \text{(D.4)}$$

Hence, $\mathcal{H}_n^{(1)}$ can be written as the combination of classes of functions: $\mathcal{H}_n^{(1)} = \mathcal{F}_{g_1,n}\cdot\mathcal{F}_{\phi_1} + \mathcal{F}_{g_2,n} + \mathcal{F}_{g_3,n}\cdot\mathcal{F}_{\phi_2} + \mathcal{F}_{g_4,n}$, where

$$g_1(x,x') = -\nabla_z K\Big(\frac{z-z'}{h}\Big)w(z)(y_U-y_L),\quad g_2(x,x') = -\nabla_z K\Big(\frac{z-z'}{h}\Big)w(z)y_L$$
$$g_3(x,x') = \nabla_z K\Big(\frac{z-z'}{h}\Big)w(z')(y_U'-y_L'),\quad g_4(x,x') = \nabla_z K\Big(\frac{z-z'}{h}\Big)w(z')y_L'.$$

By Lemma 2.14 in Pakes and Pollard (1989) and $\mathcal{F}_{\phi_1}$ and $\mathcal{F}_{\phi_2}$ having constant envelope functions, $\mathcal{H}_n^{(1)}$ is Euclidean with the envelope function $F_{g_1}+F_{g_2}+F_{g_3}+F_{g_4}$. Hence, $\mathcal{H}_n^{(1)}$ is Euclidean with the envelope function $H^{(1)}(x,x') \equiv 8(1+4\sqrt{\ell})\sup_x|g_p^{(1)}(x)|\sup_{h>0}\|\nabla_z K(\frac{z-z'}{h})\|$, where $g_p^{(1)}$ is bounded by Assumptions 2.2 and 2.3. By Assumption 4.2, $E[\sup_{h>0}\|\nabla_z K(\frac{Z_i-Z_j}{h})\|^2] < \infty$, which in turn implies $E[H^{(1)}(X_i,X_j)^2] < \infty$. This shows the claim of the lemma for $\mathcal{H}_n^{(1)}$. Showing $\mathcal{H}_n^{(2)}, \mathcal{G}_n^{(k)} k = 1,2$ are Euclidean is similar. Hence, the rest of the proof is omitted. ■

Lemma D.2. *Suppose Assumptions 2.3, 3.1, 4.2, and 4.3 hold. Suppose further that $\tilde{h} \to 0$ and $n\tilde{h}^{\ell+2} \to \infty$. Then, uniformly in $p \in \mathbb{S}^\ell$, $\hat{\upsilon}_n(p) - \bar{\upsilon}_n(p) = o_p(n^{-1/2})$.*

Proof of Lemma D.2. By $\|p\| = 1$, the Cauchy-Schwarz inequality, and Assumption 2.3,

$$E[\sup_{p\in\mathbb{S}^\ell}|u_{i,n}(p)|^2] \le 2\sup_{y\in D}|y|E[\|\hat{l}_{i,h}(Z_i)\|^2\sup_{p\in\mathbb{S}^\ell}|1\{p'\hat{l}_{i,\tilde{h}}(Z_i) > 0\} - 1\{p'l(Z_i) > 0|^2]$$
$$\le 2\sup_{y\in D}|y|E[\|\hat{l}_{i,h}(Z_i)\|^4]^{1/2}P(\mathrm{sgn}(p'\hat{l}_{i,\tilde{h}}(Z_i)) \ne \mathrm{sgn}(p'l(Z_i)), \exists p \in \mathbb{S}^\ell), \quad \text{(D.5)}$$

where $E[\|\hat{l}_{i,h}(Z_i)\|^4] < \infty$ under our choice of h and the trimming sequence. Hence, for the desired result, it suffices to show that $P(\mathrm{sgn}(p'\hat{l}_{i,\tilde{h}}(Z_i)) \ne \mathrm{sgn}(p'l(Z_i)), \exists p \in \mathbb{S}^\ell) = o(n^{-1})$. By Assumption 3.1, it follows that

$$
\begin{aligned}
&P(\text{sgn}(p'\hat{l}_{i,\tilde{h}}(Z_i)) \neq \text{sgn}(p'l(Z_i)), \exists p \in \mathbb{S}^\ell) \\
&\qquad \le P(p'\hat{l}_{i,\tilde{h}}(Z_i) > 0 \text{ and } p'l(Z_i) < 0, \exists p \in \mathbb{S}^\ell) + P(p'\hat{l}_{i,\tilde{h}}(Z_i) < 0 \text{ and } p'l(Z_i) > 0, \exists p \in \mathbb{S}^\ell). \quad \text{(D.6)}
\end{aligned}
$$

Without loss of generality, suppose that $p'\hat{l}_{i,\tilde{h}}(Z_i) > 0$ and $p'l(Z_i) < 0$ for some $p \in \mathbb{S}^\ell$. Then, there must exist $\epsilon > 0$ such that $\sup_{p\in\mathbb{S}^\ell} |p'\hat{l}_{i,\tilde{h}}(Z_i) - E[p'\hat{l}_{i,\tilde{h}}(Z_i)] + E[p'\hat{l}_{i,\tilde{h}}(Z_i))] + p'l(Z_i)| > \epsilon$. This is also true if $p'\hat{l}_{i,\tilde{h}}(Z_i)) < 0$ and $p'l(Z_i) > 0$. Therefore, by the triangle inequality and the law of iterated expectations, we may write

$$
\begin{aligned}
&P(\text{sgn}(p'\hat{l}_{i,\tilde{h}}(Z_i)) \neq \text{sgn}(p'l(Z_i)), \exists p \in \mathbb{S}^\ell) \\
&\le 2\Big\{E\Big[P\big(\sup_{p\in\mathbb{S}^\ell} |p'\hat{l}_{i,\tilde{h}}(Z_i) - E[p'\hat{l}_{i,\tilde{h}}(Z_i)|Z_i])| > \epsilon/2|Z_i\big) + P\big(\sup_{p\in\mathbb{S}^\ell} |E[p'\hat{l}_{i,\tilde{h}}(Z_i)|Z_i] - p'l(Z_i)| > \epsilon/2|Z_i\big)\Big]\Big\}, \\
&\hfill \text{(D.7)}
\end{aligned}
$$

where the second term in (D.7) vanishes for all n sufficiently large because the bias satisfies $\|E[\hat{l}_{i,\tilde{h}}(Z_i)|Z_i] - l(Z_i)\| \to 0$ with probability 1 as $\tilde{h} \to 0$. Hence, we focus on controlling the first term in (D.7) below.

Let $\bar{M} \equiv \sup_{z\in\mathcal{Z}} \|\nabla_z K(z)\|$ and define

$$
W_n(p) \equiv \frac{1}{(n-1)\tilde{h}^{(\ell+1)}} \sum_{j=1, j\neq i}^{n} p'\Big\{\nabla_z K\Big(\frac{z - Z_j}{\tilde{h}}\Big) - E[\nabla_z K\Big(\frac{z - Z_j}{\tilde{h}}\Big)]\Big\} \quad \text{(D.8)}
$$

$$
\bar{\sigma}^2 \equiv E\Big[\sup_{p\in\mathbb{S}^\ell}\Big(\frac{1}{(n-1)\tilde{h}^{(\ell+1)}} \sum_{j=1, j\neq i}^{n} p'\Big\{\nabla_z K\Big(\frac{z - Z_j}{\tilde{h}}\Big) - E[\nabla_z K\Big(\frac{z - Z_j}{\tilde{h}}\Big)]\Big\}\Big)^2\Big]. \quad \text{(D.9)}
$$

Note that, arguing as in (D.21)-(D.23), $p'\hat{l}_{i,\tilde{h}}(z) = -p'\nabla_z w(z) - w(z)p'\hat{f}_{i,\tilde{h}}(z)/\hat{f}_{i,\tilde{h}}(z) = -p'\nabla_z w(z) - w(z)p'\hat{f}_{i,\tilde{h}}(z) \times O(b_n)$. Hence,

$$
p'\hat{l}_{i,\tilde{h}}(z) - E[p'\hat{l}_{i,\tilde{h}}(Z_i)|Z_i = z] \le CW_n(p)b_n, \quad \text{(D.10)}
$$

for some $C > 0$. Below, let $a_n \equiv C/b_n$. Define $\mathcal{W} \equiv \{f : \mathcal{X} \to \mathbb{R} : f(z_j) = p'\{\nabla_z K(\frac{z-z_j}{\tilde{h}}) - E[\nabla_z K(\frac{z-Z_j}{\tilde{h}})]/\tilde{h}^{(\ell+1)}\}, p \in \mathbb{S}^\ell\}$. Then by $\mathbb{S}^\ell$ being finite dimensional and Lemma 2.6.15 in van der Vaart and Wellner (1996), $\mathcal{W}$ is a VC-subgraph class, which in turn implies that $\sup_Q N(\epsilon, \mathcal{W}, L_2(Q)) \le (\frac{K}{\epsilon})^V$ for all $0 < \epsilon < K$ for some positive constants V and K by Lemma 2.6.7 in van der Vaart and Wellner (1996). Then, by W_n being independent of Z_i and Theorem 2.14.16 in van der Vaart and Wellner (1996), we have

$$
\begin{aligned}
&P(\sup_{p\in\mathbb{S}^\ell} |p'\hat{l}_{i,\tilde{h}}(Z_i) - E[p'\hat{l}_{i,\tilde{h}}(Z_i)|Z_i])| > \epsilon/2|Z_i = z) \le P(\|W_n\|_{\mathcal{W}} > \epsilon a_n) \\
&\qquad \le C\Big(\frac{1}{\bar{\sigma}}\Big)^{2V}\Big(1 \vee \frac{\epsilon a_n}{\bar{\sigma}}\Big)^{3V+1} \exp\Big(-\frac{1}{2}\frac{(\epsilon a_n)^2}{\bar{\sigma}^2 + (3 + \epsilon a_n)/\sqrt{n}}\Big), \quad \text{(D.11)}
\end{aligned}
$$

where C is a constant that depends on V and K. Note that under the imposed conditions on $\tilde{h}$, we have

$$
\frac{(\epsilon a_n)^2}{\bar{\sigma}^2 + (3 + \epsilon a_n)/\sqrt{n}} = \frac{1}{S_{1,n} + S_{2,n}}, \quad \text{(D.12)}
$$

where $S_{1,n} \equiv \bar{\sigma}^2/(\epsilon a_n)^2$ and $S_{2,n} \equiv (3 + \epsilon a_n)/(\epsilon a_n)^2\sqrt{n}$. By (D.9), $\bar{\sigma}^2 = O(1/n\tilde{h}^{\ell+2})$, which together with Assumption 4.3 implies that $S_{1,n} = o(1)$. Similarly, $S_{2,n} = o(1)$ by $a_n\sqrt{n} \to \infty$ by Assumption 4.3. This ensures that, by (D.11), $P(\sup_{p\in\mathbb{S}^\ell}|p'\hat{l}_{i,\tilde{h}}(Z_i)-E[p'\hat{l}_{i,\tilde{h}}(Z_i)|Z_i])|>\epsilon/2|Z_i=z)$ decays exponentially as $n\to\infty$. Hence, combining this with (D.5)-(D.7), we have $E[\sup_{p\in\mathbb{S}^\ell}|u_{i,n}(p)|^2]=o(n^{-1})$ as desired. This establishes the claim of the Lemma. ■

Lemma D.3. *For each $k\in\{1,2\}$, let $U_n^{(k)}(p)\equiv\binom{n}{2}^{-1}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}p_n^{(k)}(X_i,X_j;p)$ and $\hat{U}_n^{(k)}(p)=\frac{2}{n}\sum_{i=1}^{n}r_n^{(k)}(X_i;p)$. Suppose Assumptions 2.1-2.3, 4.2-4.3 hold. Suppose further that $nh^{\ell+2+\delta}\to\infty$ for some $\delta>0$ as $h\to 0$. Then, $\sqrt{n}(\hat{U}_n^{(k)}(p)-U_n^{(k)}(p))=o_p(1)$ uniformly in $p\in\mathbb{S}^\ell$ for $k=1,2$.*

Proof of Lemma D.3. Following the same argument as in the proof of Lemma 3.1 in Powell, Stock, and Stoker (1989), we may write

$$\hat{U}_n^{(k)}(p)-U_n^{(k)}(p)=\binom{n}{2}^{-1}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}q_n^{(k)}(X_i,X_j;p),\tag{D.13}$$

where $q_n^{(k)}(x_i,x_j;p)=p_n^{(k)}(x_i,x_j;p)-r_n^{(k)}(x_i,p)-r_n^{(k)}(x_j;p)-E[r_n^{(k)}(X_i;p)]$. Recall that $\tilde{q}_n^{(k)}=h^{(\ell+1)}q_n^{(k)}$.

Below we analyze the case $k=1$. By the definition of $p_n^{(1)}$ and Assumption 4.3, we may then obtain the following bound:

$$\begin{aligned}E[\sup_{p\in\mathbb{S}^\ell}|\tilde{q}_n^{(1)}(X_i,X_j;p)/b_n|^2]&\le 16E[\sup_{p\in\mathbb{S}^\ell}|\tilde{p}_n^{(1)}(X_i,X_j;p)/b_n|^2]\\&\lesssim(\sup_{(z,y)\in\mathcal{Z}\times D}|w(z)y|)^2E\Big[\|\nabla_zK((Z_i-Z_j)/h)\|^2\Big]\lesssim h^\ell\int\|\nabla_zK(u)\|^2f(z_i)f(z_i+hu)dz_idu=O(h^\ell),\end{aligned}\tag{D.14}$$

where the second inequality follows from Assumption 2.1-2.3, $\|p\|=1$ for all p, and the Cauchy-Schwarz inequality and $\tau_n(z)/f(z)\le b_n$, while the third inequality uses the change of variables from (z_i,z_j) to $(z_i,u=(z_i-z_j)/h)$ with Jacobian $h^{-\ell}$. By Lemma D.1, $\mathcal{G}_n^{(1)}$ is Euclidean. By Theorem 3 in Sherman (1994b) applied with $\delta_n=1$, $\gamma_n^2=h^\ell$, and $k=2$, it then follows that for some $0<\alpha<1$, which can be made arbitrarily close to 1, we have

$$\begin{aligned}\binom{n}{2}^{-1}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}q_n^{(1)}(X_i,X_j;p)&=h^{-(\ell+1)}b_n\binom{n}{2}^{-1}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\tilde{q}_n^{(1)}(X_i,X_j;p)/b_n\\&=O(h^{-(\ell+1)}b_n)O_p(h^{\alpha\ell/2}/n)=O_p(h^{\ell(\frac{\alpha}{2}-1)-1}b_n/n)\end{aligned}\tag{D.15}$$

uniformly over $\mathbb{B}^\ell$. Since α can be made arbitrarily close to 1, there is $\delta>0$ such that $h^{\ell(\frac{\alpha}{2}-1)-1}=O(h^{-\frac{\ell}{2}-1-\delta})=o(\sqrt{n})O(n^{-\frac{\delta}{2}})$, where the last equality follows from the assumption that $nh^{\ell+2+\delta}\to\infty$. This together with (D.15) and $O(b_nn^{-\delta/2})=o(1)$ by Assumption 4.3 implies $\binom{n}{2}^{-1}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}q_n^{(1)}(X_i,X_j;p)=o(n^{-1/2})$ uniformly over $\mathbb{B}^\ell$. By $\mathbb{S}^\ell\subset\mathbb{B}^\ell$ and (D.13), this establishes the claim of the lemma for $k=1$.

For $k=2$, note that

$$\begin{aligned}E[\sup_{p\in\mathbb{S}^\ell}|\tilde{q}_n^{(2)}(X_i,X_j;p)/b_n|^2]&\le 16E[\sup_{p\in\mathbb{S}^\ell}|\tilde{p}_n^{(2)}(X_i,X_j;p)/b_n|^2]\\&\lesssim(\sup_{(z,y)\in\mathcal{Z}\times D}|w(z)p'\nabla_zf(z)y|)^2E\Big[|K((Z_i-Z_j)/h)|^2\Big]\\&\lesssim h^\ell\int|K(u)|^2f(z_i)f(z_i+hu)dz_idu=O(h^\ell),\end{aligned}\tag{D.16}$$

where the second inequality follows from Assumption 2.1 and $\tau_n/f^2(z) = \tau_n^2/f^2(z) \le b_n^2$ by Assumption 4.3, while the third inequality uses the change of variables from (z_i, z_j) to $(z_i, u = (z_i - z_j)/h)$ with Jacobian $h^{-\ell}$. Mimic the argument for $k = 1$ to obtain $\binom{n}{2}^{-1}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} q_n^{(2)}(X_i, X_j; p) = o(n^{-1/2})$. By $\mathbb{S}^\ell \subset \mathbb{B}^\ell$ and (D.13), this establishes the claim of the lemma for $k = 2$. ■

Lemma D.4. *Suppose Assumptions 2.1-2.3, 3.1, and 4.1-4.3 hold. Suppose further that $nh^{\ell+2+\delta} \to \infty$ for some $\delta > 0$, and $nh^{2J} \to 0$ as $h \to 0$. Then, uniformly in $p \in \mathbb{S}^\ell$, $\sqrt{n}(\bar{v}_n(p) - E[\bar{v}_n(p)]) = \frac{1}{\sqrt{n}}\sum_{i=1}^n \psi_p(X_i) + o_p(1)$.*

Proof of Lemma D.4. We start with the observation that

$$\bar{v}_n(p) = -\frac{1}{n}\sum_{i=1}^n p'\nabla_z w(Z_i)Y_{p,i} - \frac{1}{n}\sum_{i=1}^n w(Z_i)\frac{p'\nabla_z \hat{f}_{i,h}(Z_i)}{\hat{f}_{i,h}(Z_i)}Y_{p,i}\tau_{n,i}. \tag{D.17}$$

By a second-order Taylor expansion of $\frac{p'\nabla_z \hat{f}_{i,h}(z)}{\hat{f}_{i,h}(z)}$ around $\frac{p'\nabla_z f(z)}{f(z)}$, the second term in (D.17) can be written as

$$\begin{aligned}
&-\frac{1}{n}\sum_{i=1}^n w(Z_i)\frac{p'\nabla_z \hat{f}_{i,h}(Z_i)}{\hat{f}_{i,h}(Z_i)}Y_{p,i}\tau_{n,i} \\
&\quad= -\frac{1}{n}\sum_{i=1}^n w(Z_i)\frac{p'\nabla_z f(Z_i)}{f(Z_i)}Y_{p,i}\tau_{n,i} - \frac{1}{n}\sum_{i=1}^n w(Z_i)\frac{p'(\nabla_z \hat{f}_{i,h}(Z_i) - \nabla_z f(Z_i))}{f(Z_i)}Y_{p,i}\tau_{n,i} \\
&\qquad + \frac{1}{n}\sum_{i=1}^n w(Z_i)\frac{p'\nabla_z f(Z_i)(\hat{f}_{i,h}(Z_i) - f(Z_i))}{f(Z_i)^2}Y_{p,i}\tau_{n,i} + R_n \\
&\quad= -\frac{1}{n}\sum_{i=1}^n w(Z_i)\frac{p'\nabla_z f(Z_i)}{f(Z_i)}Y_{p,i}\tau_{n,i} - \frac{1}{n}\sum_{i=1}^n w(Z_i)\frac{p'\nabla_z \hat{f}_{i,h}(Z_i)}{f(Z_i)}Y_{p,i}\tau_{n,i} \\
&\qquad + \frac{1}{n}\sum_{i=1}^n w(Z_i)\frac{p'\nabla_z f(Z_i)\hat{f}_{i,h}(Z_i)}{f(Z_i)^2}Y_{p,i}\tau_{n,i} + R_n \\
&\quad= H_{1,n} + H_{2,n} + H_{3,n} + R_n,
\end{aligned} \tag{D.18}$$

where

$$\begin{aligned}
&H_{1,n} = -\frac{1}{n}\sum_{i=1}^n w(Z_i)\frac{p'\nabla_z f(Z_i)}{f(Z_i)}Y_{p,i}\tau_{n,i}, \qquad H_{2,n} = -\frac{1}{n}\sum_{i=1}^n w(Z_i)\frac{p'\nabla_z \hat{f}_{i,h}(Z_i)}{f(Z_i)}Y_{p,i}\tau_{n,i} \\
&H_{3,n} = \frac{1}{n}\sum_{i=1}^n w(Z_i)\frac{p'\nabla_z f(Z_i)\hat{f}_{i,h}(Z_i)}{f(Z_i)^2}Y_{p,i}\tau_{n,i},
\end{aligned} \tag{D.19}$$

and R_n is a remainder term that contains quadratic terms in the expansion. By (D.17), (D.18), and the law of iterated expectations, one may therefore write

$$\begin{aligned}
\sqrt{n}(\bar{v}_n(p) - E[\bar{v}_n(p)]) &= -\frac{1}{\sqrt{n}}\sum_{i=1}^n (p'\nabla_z w(Z_i)Y_{p,i} - E[p'\nabla_z w(Z_i)m_{p,i}(Z_i)]) \\
&\quad + \sum_{j=1}^3 \sqrt{n}(H_{j,n} - E[H_{j,n}]) + \sqrt{n}(R_n - E[R_n]).
\end{aligned} \tag{D.20}$$

The remainder term involves the second derivatives of $\frac{p'\nabla_z \hat{f}_{i,h}(z)}{\hat{f}_{i,h}(z)}$ evaluated at $\frac{p'\nabla_z \tilde{f}_{i,h}}{\tilde{f}_{i,h}}$, where $\tilde{f}_{i,h}$ lies between $\hat{f}_{i,h}$ and $f_{i,h}$. For example, one component of $\sqrt{n}R_n$ can be bounded by

$$\begin{aligned}
&\frac{1}{\sqrt{n}}\sum_{i=1}^{n}|-\frac{\tau_{n,i}}{2\tilde{f}_{i,h}(Z_i)}p'(\nabla_z\hat{f}_{i,h}(Z_i)-\nabla_z f(Z_i))(\hat{f}_{i,h}(Z_i)-f(Z_i))| \\
&\qquad\qquad\lesssim \sqrt{n}\sup_{z\in\mathcal{Z}}\frac{\tau_{n,i}}{f(z)}\Big|1-(1-\frac{\tilde{f}_{i,h}(z)}{f(z)})\Big|^{-1}\sup_{z\in\mathcal{Z}}\|\nabla_z\hat{f}_{i,h}(z)-\nabla_z f(z)\|\sup_{z\in\mathcal{Z}}|\hat{f}_{i,h}(z)-f(z))| \\
&\qquad\qquad= \sqrt{n}O(b_n)O_p\big((\frac{\ln n}{nh^{\ell+2}})^{1/2}+h^J\big)O_p\big((\frac{\ln n}{nh^{\ell}})^{1/2}+h^J\big)=o_p(1),
\end{aligned}\tag{D.21}$$

uniformly in $p\in\mathbb{S}^\ell$, where the first equality follows from the geometric expansion (as in Lemma 6A in Sherman (1994b)):

$$\frac{\tau_{n,i}}{f(z)}\Big(1-(1-\frac{\tilde{f}_{i,h}(z)}{f(z)})\Big)^{-1}=\frac{\tau_{n,i}}{f(z)}\Big(1+(1-\frac{\tilde{f}_{i,h}(z)}{f(z)})+\cdots\Big)=\frac{\tau_{n,i}}{f(z)}\Big(f(z)+\frac{1}{f(z)}(f(z)-\tilde{f}_{i,h}(z))+\cdots\Big),\tag{D.22}$$

which can be bounded by the following:

$$f(z)b_n+\frac{\tau_n(z)^2}{f^2(z)}|f(z)-\tilde{f}_{i,h}(z)|+\cdots=O(b_n)+O(b_n^2)O_p\big((\frac{\ln n}{nh^\ell})^{1/2}+h^J\big)+\cdots=O(b_n),\tag{D.23}$$

where we used Assumption 4.3, $\tau_n(z)\in\{0,1\}$, and the uniform convergence rate of $\tilde{f}_{i,h}$, which follows from Theorem 6 in Hansen (2008). Applying Theorem 6 in Hansen (2008) again, the penultimate equality in (D.21) follows. The last equality in (D.21) follows from the assumption $nh^{\ell+2+\delta}\to\infty$, $nh^{2J}\to 0$, $J>(\ell+2)/2$, and Assumption 4.3. Other components of R_n can be shown to be $o_p(1)$ similarly. Hence, it follows that

$$\sqrt{n}(R_n-E[R_n])=o_p(1).\tag{D.24}$$

Below, we investigate $H_{j,n}, j=1,2,3$. We first note that

$$\sqrt{n}(H_{1,n}-E[H_{1,n}])=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(\psi_{p,1}(X_i)-E[\psi_{p,1}(X_i)])+o_p(1),\tag{D.25}$$

where $\psi_{p,1}(x)=-w(z)\frac{p'\nabla_z f(z)}{f(z)}y_p$. This follows from the following argument. Notice that

$$H_{1,n}=-\frac{1}{n}\sum_{i=1}^{n}w(Z_i)\frac{p'\nabla_z f(Z_i)}{f(Z_i)}Y_{p,i}\tau_{n,i}=\frac{1}{n}\sum_{i=1}^{n}\psi_{p,1}(x)\tau_{n,i}.\tag{D.26}$$

and hence

$$E\Big[\Big(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(\psi_{p,1}(X_i)-\psi_{p,1}(X_i)\tau_{n,i})\Big)^2\Big]=o(1),\tag{D.27}$$

by Assumption 4.3 and arguing as in the proof of Theorem A (page 10) in Lewbel (2000). Furthermore, by the Cauchy-Schwarz inequality,

$$|E[\psi_{p,1}(X_i)]-E[\psi_{p,1}(X_i)\tau_{n,i}]|\le\|\psi_{p,1}(X_i)\|_{L_P^2}\|1-\tau_{n,i}\|_{L_P^2}=o_p(n^{-1/2}).\tag{D.28}$$

Eq. (D.25) then follows from (D.27)-(D.28).

Next, noting that ∇K is an odd function,

$$\begin{aligned}
H_{2,n} &= -\frac{1}{n}\sum_{i=1}^{n} w(Z_i)\frac{p'\nabla_z \hat{f}_{i,h}(Z_i)}{f(Z_i)} Y_{p,i}\tau_{n,i} \\
&= \frac{-1}{n(n-1)h^{\ell+1}}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n} w(Z_i)\frac{p'\nabla_z K(Z_i - Z_j/h)}{f(Z_i)} Y_{p,i}\tau_{n,i} \\
&= \frac{-1}{2}\binom{n}{2}^{-1}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\Big(\frac{1}{h}\Big)^{\ell+1} p'\nabla_z K\Big(\frac{Z_i - Z_j}{h}\Big)(g_p^{(1)}(X_i) - g_p^{(1)}(X_j)) \\
&= \binom{n}{2}^{-1}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} p_n^{(1)}(X_i, X_j; p), \qquad \text{(D.29)}
\end{aligned}$$

where $g_p^{(1)}(X_i) = \frac{W(Z_i)}{f(Z_i)} Y_{p,i}\tau_{n,i}$ and $p_n^{(1)}(X_i, X_j; p) = \frac{-1}{2}\Big(\frac{1}{h}\Big)^{\ell+1} p'\nabla_z K\Big(\frac{Z_i - Z_j}{h}\Big)(g_p^{(1)}(X_i) - g_p^{(1)}(X_j))$. By Lemma D.3

$$\sqrt{n}(\binom{n}{2}^{-1}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} p_n^{(1)}(X_i, X_j; p) - E[p_n^{(1)}(X_i, X_j; p)]) = \frac{2}{\sqrt{n}}(\sum_{i=1}^{n} r_n^{(1)}(X_i; p) - E[r_n^{(1)}(X_i; p)]) + o_p(1), \qquad \text{(D.30)}$$

where $r_n^{(1)}(x; p) \equiv E[p_n^{(1)}(X_i, X_j; p)|X_i = x]$.

Arguing as in Eq. (3.15) in PSS and by the law of iterated expectations, we may then write

$$\begin{aligned}
r_n^{(1)}(x; p) &= \frac{-1}{2}\int \Big(\frac{1}{h}\Big)^{\ell+1} p'\nabla_z K\Big(\frac{Z_i - z}{h}\Big)(g_p^{(1)}(X_i) - \frac{w(z)}{f(z)} m_p(z)) f(z) dz \\
&= \frac{1}{2}\int \Big(\frac{1}{h}\Big) p'\nabla_z K(u)(g_p^{(1)}(X_i) - \frac{wm_p}{f}(Z_i + hu)) f(Z_i + hu) du \\
&= \frac{-1}{2}\int K(u) g_p^{(1)}(X_i) p'\nabla_z f(Z_i + hu) du \\
&\quad + \frac{1}{2}\int K(u) p'\nabla_z (wm_p)(Z_i + hu) du \\
&= \frac{1}{2}\{-w(Z_i)\frac{p'\nabla_z f(Z_i)}{f(Z_i)} Y_{p,i}\tau_{n,i} + p'\nabla_z w(Z_i) m_p(Z_i) + w(Z_i) p'\nabla_z m_p(Z_i)\} + t_n^{(1)}(X_i; p), \qquad \text{(D.31)}
\end{aligned}$$

where we used the change of variables, integration by parts and the assumption that $K(u) = 0$ on the boundary of $\mathcal{S}_K$. The remainder term $t_n^{(1)}$ is given by

$$\begin{aligned}
t_n^{(1)}(X_i; p) = \frac{-1}{2}\int K(u) g_p^{(1)}(X_i) p'(\nabla_z f(Z_i + hu) - \nabla_z f(Z_i)) du \\
+ \frac{1}{2}\int K(u) p'(\nabla_z (wm_p)(Z_i + hu) - \nabla_z (wm_p)(Z_i)) du. \qquad \text{(D.32)}
\end{aligned}$$

By Assumptions 2.1 and 4.1, uniformly in $p \in \mathbb{S}^\ell$,

$$|g_p^{(1)}(x)| \|\nabla_z f(z + hu) - \nabla_z f(z)\| \lesssim \sup_{y\in\mathcal{D}} |y| \times b_n \times w(z) M_1(z) \|hu\|. \qquad \text{(D.33)}$$

By Assumptions 3.1, 4.1 and $m_p(z) = 1\{p'l(z) \leq 0\} m_L(z) + 1\{p'l(z) \leq 0\} m_U(z)$, we have uniformly in $p \in \mathbb{S}^\ell$,

$$\|\nabla_z (wm_p)(z + hu) - \nabla_z (wm_p)(z)\| \leq 2M_2(z)\|hu\|. \qquad \text{(D.34)}$$

Assumption 4.1 and (D.33)-(D.34) then imply

$$E[\sup_{p\in\mathbb{S}^\ell}|t_n^{(1)}(X;p)|^2] \lesssim |h|^2E[\{b_n\sup_{y\in D}|y||w(Z)M_1(Z)|+2|M_2(Z)|\}^2](\int\|u\||K(u)|du)^2 = O(h^2b_n^2).$$

By Assumption 4.3 and $h\to 0$ at a polynomial rate (by $nh^{2J}\to 0$.), this in turn implies $\frac{1}{\sqrt{n}}\sum_{i=1}^n t_n^{(1)}(X,\cdot)-E[t_n^{(1)}(X,\cdot)]$ converges in probability to 0 uniformly in $p\in\mathbb{S}^\ell$. Using this result together with (D.29)-(D.31), and arguing as in (D.26)-(D.28) to control the effect of the asymptotic trimming, we obtain

$$\sqrt{n}(H_{2,n}-E[H_{2,n}]) = \frac{1}{\sqrt{n}}\sum_{i=1}^n(\psi_{p,2}(X_i)-E[\psi_{p,2}(X_i)])+o_p(1), \tag{D.35}$$

where $\psi_{p,2}(x) = -w(z)\frac{p'\nabla_z f(z)}{f(z)}y_{p,i}+p'\nabla_z w(z)m_p(z)+w(z)p'\nabla_z m_p(z)$.

Now we turn to $H_{3,n}$ in (D.18). Noting that K is an even function, we have

$$\begin{aligned}
\frac{1}{n}\sum_{i=1}^n w(Z_i)\frac{p'\nabla_z f(Z_i)\hat{f}_{i,h}(Z_i)}{f(Z_i)^2}Y_{p,i}\tau_{n,i} &= \frac{1}{n(n-1)h^\ell}\sum_{i=1}^n\sum_{j=1,j\neq i}^n w(Z_i)\frac{p'\nabla_z f(Z_i)K(Z_i-Z_j/h)}{f(Z_i)^2}Y_{p,i}\tau_{n,i}\\
&= \frac{1}{2}\binom{n}{2}^{-1}\sum_{i=1}^{n-1}\sum_{j=i+1}^n\Big(\frac{1}{h}\Big)^\ell K\Big(\frac{Z_i-Z_j}{h}\Big)(g_p^{(2)}(X_i)+g_p^{(2)}(X_j))\\
&= \binom{n}{2}^{-1}\sum_{i=1}^{n-1}\sum_{j=i+1}^n p_n^{(2)}(X_i,X_j;p),
\end{aligned}\tag{D.36}$$

where $g_p^{(2)}(X_i) = w(Z_i)\frac{p'\nabla_z f(Z_i)}{f(Z_i)^2}Y_{p,i}\tau_{n,i}$ and $p_n^{(2)}(X_i,X_j;p) = \frac{1}{2}\Big(\frac{1}{h}\Big)^\ell K\Big(\frac{Z_i-Z_j}{h}\Big)(g_p^{(2)}(X_i)+g_p^{(2)}(X_j))$. By Lemma D.3, (D.30) holds while replacing $p_n^{(1)},r_n^{(1)}$ with $p_n^{(2)},r_n^{(2)}$. Arguing similarly to (D.31), we may then write

$$\begin{aligned}
r_n^{(2)}(x;p) &= \frac{1}{2}\int\Big(\frac{1}{h}\Big)^\ell K\Big(\frac{Z_i-z}{h}\Big)(g_p^{(2)}(X_i)+\frac{w(z)p'\nabla_z f(z)}{f(z)^2}m_p(z))f(z)dz\\
&= \frac{1}{2}\int K(u)(g_p^{(2)}(X_i)+\frac{wp'\nabla_z f m_p}{f^2}(Z_i+hu))f(Z_i+hu)du\\
&= \frac{1}{2}\int K(u)g_p^{(2)}(X_i)f(Z_i+hu)du\\
&\quad+\frac{1}{2}\int K(u)\frac{wp'\nabla_z f}{f}m_p(Z_i+hu)du\\
&= \frac{1}{2}\{w(Z_i)\frac{p'\nabla_z f(Z_i)}{f(Z_i)}Y_{p,i}\tau_{n,i}+w(Z_i)\frac{p'\nabla_z f(Z_i)}{f(Z_i)}m_p(Z_i)\}+t_n^{(2)}(X_i;p),
\end{aligned}\tag{D.37}$$

where the remainder term $t_n^{(2)}(X_i;p)$ is given by

$$\begin{aligned}
t_n^{(2)}(X_i;p) = \frac{1}{2}\int K(u)g_p^{(2)}(X_i)(f(Z_i+hu)-f(Z_i))du&\\
+\frac{1}{2}\int K(u)\Big(\big(\frac{wp'\nabla_z f m_p}{f}\big)(Z_i+hu)-\big(\frac{wp'\nabla_z f m_p}{f}\big)(Z_i)\Big)du.&
\end{aligned}\tag{D.38}$$

By Assumptions 2.1 and 4.1, uniformly in $p\in\mathbb{S}^\ell$,

$$|g_p^{(2)}(x)||f(z+hu)-f(z)| \lesssim \sup_{y\in\mathcal{D}}|y|\times b_n^2\times w(z)\|\nabla_z f(z)\|M_3(z)\|hu\|. \tag{D.39}$$

By Assumptions 3.1, 4.1 and $m_p(z) = 1\{p'l(z)\le 0\}m_L(z)+1\{p'l(z)\le 0\}m_U(z)$, we have uniformly in

$p \in \mathbb{S}^\ell$,

$$|\frac{wp'\nabla_z f m_p}{f}(z+hu) - \frac{wp'\nabla_z f m_p}{f}(z)| \leq 2M_4(z)\|hu\|. \tag{D.40}$$

Assumption 4.1 and (D.39)-(D.40) then imply

$$E[\sup_{p\in\mathbb{S}^\ell} |t_n^{(2)}(X;p)|^2] \lesssim |h|^2 E[\{b_n^2 \sup_{y\in\mathcal{D}} |y| \times w(Z)\|\nabla_z f(Z)\| M_3(Z) + 2M_4(z)\}^2](\int \|u\||K(u)|du)^2 = O(h^2 b_n^4).$$

By Assumption 4.3 and $h \to 0$ at a polynomial rate, this in turn implies $\frac{1}{\sqrt{n}}\sum_{i=1}^n t_n^{(2)}(X,\cdot) - E[t_n^{(2)}(X,\cdot)]$ converges in probability to 0 uniformly in $p \in \mathbb{S}^\ell$. This result, (D.36), (D.30) with replacing $p_n^{(1)}, r_n^{(1)}$ with $p_n^{(2)}, r_n^{(2)}$, and (D.37) imply

$$\sqrt{n}(H_{3,n} - E[H_{3,n}]) = \frac{1}{\sqrt{n}}\sum_{i=1}^n (\psi_{p,3}(X_i) - E[\psi_{p,3}(X_i)]) + o_p(1), \tag{D.41}$$

where $\psi_{p,3}(x) = w(z)\frac{p'\nabla_z f(z)}{f(z)} y_p + w(z)\frac{p'\nabla_z f(z)}{f(z)} m_p(z)$, and the effect of the asymptotic trimming is controlled by arguing as in (D.26)-(D.28).

Note that by (D.17), (D.25), (D.35), and (D.41), and an integration by parts, $\psi_p(x) - E[\psi_p(X)] = -p'\nabla_z w(z) y_p - E[w(Z)p'\nabla_z m_p(Z)] + \sum_{j=1}^3(\psi_{p,j}(x) - E[\psi_{p,j}(Z)])$. By (D.20), (D.24), and $E[\psi_p(X)] = 0$, the conclusion of the lemma then follows. ■

Lemma D.5. *Suppose Assumptions 4.1-4.2 hold. Suppose that $nh^{2J} \to 0$. Then, uniformly in $p \in \mathbb{S}^\ell$, $E[\bar{\upsilon}_n(p)] - \upsilon(p, \Theta_0(P_0)) = o(n^{-1/2})$.*

Proof of Lemma D.5. Note that by (D.20) and (D.24), and the definition of $H_{1,n}$,

$$\begin{aligned} E[\bar{\upsilon}_n(p)] &= E[-p'\nabla_z w(Z) m_p(Z)] - E[w(Z)\frac{p'\nabla_z f(Z)}{f(Z)} m_P(Z)] + \sum_{j=2}^3 E[H_{j,n}] + E[R_n] \\ &= \upsilon(p, \Theta_0(P_0)) + \sum_{j=2}^3 E[H_{j,n}] + o(n^{-1/2}). \end{aligned} \tag{D.42}$$

Hence, for the conclusion of the lemma, it suffices to show that $\sum_{j=2}^3 E[H_{j,n}] = o(n^{-1/2})$. Further, by Assumption 4.3 and the argument in (D.28), the presence of the trimming function does not affect the analysis, and hence we omit $\tau_{n,i}$ below.

The rest of the proof is based on the proof of Theorem 3.2 in Powell, Stock, and Stoker (1989). Hence, we briefly sketch the argument. By (D.29), the law of iterated expectations, and arguing as in (3.19) in Powell, Stock, and Stoker (1989), we obtain

$$\begin{aligned} E[H_{2,n}] &= -\int\int \left(\frac{1}{h}\right)^{\ell+1} p'\nabla_z K\Big(\frac{z-z'}{h}\Big) g_p^{(1)}(x) f(z) f(z') dz dz' \\ &= \frac{1}{h}\int\int p'\nabla_z K(u)\tilde{g}_p^{(1)}(z) f(z) f(z+hu) dz du = -\int\int K(u)\tilde{g}_p^{(1)}(z) f(z) p'\nabla_z f(z+hu) dz du, \end{aligned} \tag{D.43}$$

where $\tilde{g}_p^{(1)}(z) \equiv \frac{w(z)}{f(z)} m_p(z)$. By Assumptions 4.1, 4.2, and Young's version of Taylor's theorem, for each $p \in \mathbb{S}^\ell$, we then obtain the expansion:

$$\sqrt{n}E[H_{2,n}] = b_1(p)\sqrt{n}h + b_2(p)\sqrt{n}h^2 + \cdots + b_{J-1}(p)\sqrt{n}h^{J-1} + O(\sqrt{n}h^J), \tag{D.44}$$

where b_k is given by

$$b_k(p) = \frac{-1}{k!}\sum_{j_1,\cdots,j_k}\int u^{j_1}\cdots u^{j_k}K(u)du \times \int \tilde{g}_p^{(1)}(z)\sum_{i=1}^{\ell}p^{(i)}\frac{\partial^{k+1}f(z)}{\partial z_{j_1}\cdots\partial z_{j_k}\partial z_i}f(z)dz,\ p\in\mathbb{S}^\ell,\ k=1,\cdots,J, \tag{D.45}$$

which shows that the map $p\mapsto b_k(p)$ is continuous on $\mathbb{S}^\ell$ for $k=1,2,\cdots,J$. This implies that the expansion in (D.44) is valid uniformly over the compact set $\mathbb{S}^\ell$. By Assumption 4.2 (v) and (D.45), $b_k(p)=0$ for all $k\le J$ but $b_k\neq 0$ for $k=J$. By the hypothesis that $nh^{2J}\to 0$, we obtain $\sqrt{n}E[H_{2,n}]=O(\sqrt{n}h^J)=o(1)$.

Similarly, one may write

$$\sqrt{n}E[H_{3,n}] = \int\int\Big(\frac{1}{h}\Big)^\ell K\Big(\frac{z-z'}{h}\Big)g_p^{(2)}(x)f(z)f(z')dzdz' \tag{D.46}$$

$$= \int\int K(u)\tilde{g}_p^{(2)}(z)f(z)p'\nabla_z f(z+hu)dzdu, \tag{D.47}$$

where $\tilde{g}_p^{(2)}(x)\equiv w(z)\frac{p'\nabla_z f(z)}{f(z)^2}m_p(z)$. Mimic the argument for $H_{2,n}$. Then, it follows that $\sqrt{n}E[H_{3,n}]=o(1)$. This establishes the claim of the lemma. ■

Lemma D.6. *Suppose Assumptions 2.1-2.3, and 3.2 hold. Then,* $\mathcal{F}\equiv\{\psi_p:\mathcal{X}\to\mathbb{R}:\psi_p(x)=w(z)p'\nabla_z m_p(z)-\upsilon(p,\Theta_0(P))+p'l(z)\zeta_p(x)\}$ *is Donsker in* $\mathcal{C}(\mathbb{S}^\ell)$.

Proof of Lemma D.6. Let $\mathcal{F}_g\equiv\{f:\mathcal{X}\to\mathbb{R}:f(x)=p'g(x),p\in\mathbb{S}^\ell\}$, where $g:\mathcal{X}\to\mathbb{R}^\ell$ is a known function. Then by $\mathbb{S}^\ell$ being finite dimensional and Lemma 2.6.15 in van der Vaart and Wellner (1996), $\mathcal{F}_g$ is a VC-subgraph class of index $\ell+2$ with an envelope function $F(x)\equiv\|g(x)\|$. Define

$$g_1(x)\equiv w(z)(\nabla_z m_U(z)-\nabla_z m_L(z)),\quad g_2(x)\equiv w(z)\nabla_z m_L(z), \tag{D.48}$$

$$g_3(x)\equiv l(z)\{r_U^{-1}(z)q(y_U-m_U(z))-r_L^{-1}(z)q(y_L-m_L(z))\}, \tag{D.49}$$

$$g_4(x)\equiv l(z)r_L^{-1}(z)q(y_L-m_L(z)),\quad g_5(x)\equiv l(z). \tag{D.50}$$

Then $\mathcal{F}_{g_j}, j=1,\cdots,5$ are VC-subgraph classes. Further, let $\mathcal{F}_\upsilon\equiv\{f:\mathcal{X}\to\mathbb{R}:f(x)=\upsilon(p,\Theta_0(P)),p\in\mathbb{S}^\ell\}$. This is also finite dimensional. Hence, $\mathcal{F}_\upsilon$ is a VC-subgraph class. Finally, let $\mathcal{F}_\phi\equiv\{f:\mathcal{X}\to\mathbb{R}:1\{p'l(z)>0\},p\in\mathbb{S}^\ell\}$. Then, $F_\phi=\phi\circ\mathcal{F}_{g_5}$, where $\phi:\mathbb{R}\to\mathbb{R}$ is the monotone map $\phi(w)=1\{w>0\}$. By Lemma 2.6.18 in van der Vaart and Wellner (1996), $\mathcal{F}_\phi$ is also a VC-subgraph class.

Note that ψ_p can be written as

$$\begin{aligned}\psi_p(x)&=w(z)p'\{1\{p'l(z)>0\}(\nabla_z m_U(z)-\nabla_z m_L(z))+\nabla_z m_L(z)\}-\upsilon(p,\Theta_0(P))\\&+p'l(z)\{1\{p'l(z)>0\}\{r_U^{-1}(z)q(y_U-m_U(z))-r_L^{-1}(z)q(y_L-m_L(z))\}+r_L^{-1}(z)q(y_L-m_L(z))\}.\end{aligned} \tag{D.51}$$

Therefore, $\mathcal{F}=\mathcal{F}_{g_1}\cdot\mathcal{F}_\phi+\mathcal{F}_{g_2}+(-\mathcal{F}_\upsilon)+\mathcal{F}_{g_3}\cdot\mathcal{F}_\phi+\mathcal{F}_{g_4}$, which is again a VC-subgraph class with some index $V(\mathcal{F})$ by Lemma 2.6.18 in van der Vaart and Wellner (1996). By Assumptions 2.1-2.3 and 3.2, we may take $F(x)\equiv\sup_{p\in\mathbb{S}^\ell}\|w(z)p'\nabla_z m_{p,\eta}(z)\|_{L^\infty_\mu}+\|l(z)\|\times\bar{\epsilon}^{-1}\times\sup_{u\in D'}|q(u)|$ as an envelope function such that $E[F(x)^2]<\infty$. Then, by Theorems 2.6.7 and 2.5.1 in van der Vaart and Wellner (1996), $\mathcal{F}$ is a Donsker class. This establishes the claim of the lemma. ■

Proof of Theorem 4.1. For each $p \in \mathbb{S}^\ell$, we have the following decomposition:

$$
\begin{aligned}
&\sqrt{n}(\hat{\upsilon}_n(p) - \upsilon(p, \Theta_0(P_0))) \\
&\qquad = \sqrt{n}(\hat{\upsilon}_n(p) - \bar{\upsilon}_n(p)) + \sqrt{n}(\bar{\upsilon}_n(p) - E[\bar{\upsilon}_n(p)]) + \sqrt{n}(E[\bar{\upsilon}_n(p)] - \upsilon(p, \Theta_0(P_0))) \\
&\qquad \equiv G_{1n}(p) + G_{2n}(p) + G_{3n}(p). \qquad \text{(D.52)}
\end{aligned}
$$

By Lemmas D.2-D.5, uniformly in $p \in \mathbb{S}^\ell$, $G_{1n}(p) = G_{3n}(p) = o_p(1)$, and $G_{2n}(p) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \psi_p(Z_i) + o_p(1)$. This establishes the second claim of the Theorem. By Theorem 3.1, ψ_p is the efficient influence function, and hence regularity of $\{\hat{\upsilon}_n(\cdot)\}$ follows from Lemma D.6 and Theorem 18.1 in Kosorok (2008), which establishes the first claim. The stated convergence in distribution is then immediate from (D.52) and Lemma D.6, while the limiting process having the efficient covariance kernel is a direct result of the characterization of the semiparametric efficiency bound obtained in Theorem 3.1, which establishes the third claim. ■

Appendix E: Figures and Tables

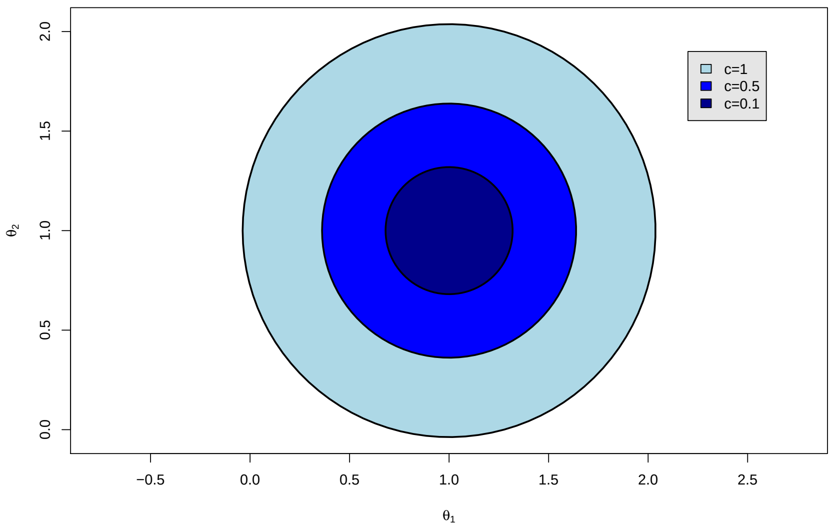

Figure 1: Identified sets for the density weighted average derivatives

Table 1: Risk of $\hat{\upsilon}_n^{IV}$ (Gaussian kernel)

			c=0.1			c=0.5			c=1	
Sample Size	h	R_H	R_{IH}	R_{OH}	R_H	R_{IH}	R_{OH}	R_H	R_{IH}	R_{OH}
n=1000										
	0.4	0.0608	0.0477	0.0600	0.0834	0.0709	0.0673	0.1229	0.1037	0.0801
	0.5	0.0588	0.0468	0.0578	0.0785	0.0749	0.0485	0.1212	0.1185	0.0437
	0.6	0.0572	0.0452	0.0564	0.0809	0.0804	0.0351	0.1305	0.1304	0.0229
	0.7	0.0567	0.0416	0.0563	0.0844	0.0844	0.0263	0.1416	0.1416	0.0086
	0.8	0.0555	0.0386	0.0553	0.0882	0.0882	0.0195	0.1556	0.1556	0.0026
n=500										
	0.4	0.0929	0.0703	0.0919	0.1185	0.0877	0.1072	0.1731	0.1203	0.1437
	0.5	0.0836	0.0684	0.0817	0.1091	0.0979	0.0839	0.1555	0.1414	0.0873
	0.6	0.0799	0.0646	0.0786	0.1038	0.0999	0.0640	0.1555	0.1520	0.0530
	0.7	0.0774	0.0607	0.0762	0.1060	0.1051	0.0512	0.1679	0.1677	0.0297
	0.8	0.0775	0.0592	0.0769	0.1098	0.1096	0.0410	0.1785	0.1785	0.0173
n=250										
	0.4	0.1357	0.0960	0.1349	0.1820	0.1061	0.1770	0.2480	0.1256	0.2339
	0.5	0.1189	0.0941	0.1169	0.1517	0.1231	0.1289	0.2013	0.1638	0.1446
	0.6	0.1133	0.0914	0.1112	0.1413	0.1299	0.1053	0.1954	0.1818	0.1084
	0.7	0.1121	0.0910	0.1098	0.1365	0.1317	0.0890	0.1974	0.1949	0.0725
	0.8	0.1086	0.0864	0.1068	0.1374	0.1360	0.0737	0.2069	0.2061	0.0500

Table 2: Risk of $\hat{\upsilon}_n^{IV}$ (Higher-order kernel)

			c=0.1			c=0.5			c=1	
Sample Size	h	R_H	R_{IH}	R_{OH}	R_H	R_{IH}	R_{OH}	R_H	R_{IH}	R_{OH}
n=1000										
	0.5	0.0722	0.0549	0.0714	0.1267	0.0461	0.1256	0.2038	0.0494	0.2017
	0.6	0.0654	0.0551	0.0637	0.0912	0.0532	0.0872	0.1384	0.0636	0.1312
	0.7	0.0600	0.0511	0.0583	0.0760	0.0631	0.0645	0.1020	0.0835	0.0745
	0.8	0.0564	0.0470	0.0553	0.0759	0.0741	0.0444	0.1093	0.1085	0.0370
	0.9	0.0565	0.0446	0.0559	0.0802	0.0801	0.0313	0.1302	0.1302	0.0134
n=500										
	0.5	0.1104	0.0744	0.1101	0.1867	0.0587	0.1861	0.2887	0.0604	0.2870
	0.6	0.0947	0.0753	0.0930	0.1308	0.0745	0.1267	0.1993	0.0857	0.1914
	0.7	0.0869	0.0737	0.0846	0.1080	0.0843	0.0970	0.1453	0.1085	0.1184
	0.8	0.0802	0.0668	0.0783	0.1019	0.0958	0.0747	0.1373	0.1308	0.0683
	0.9	0.0772	0.0635	0.0758	0.1042	0.1034	0.0564	0.1513	0.1508	0.0372
n=250										
	0.5	0.1788	0.1038	0.1787	0.2832	0.0630	0.2831	0.4316	0.0511	0.4313
	0.6	0.1374	0.1034	0.1359	0.1959	0.0979	0.1925	0.2802	0.1061	0.2716
	0.7	0.1212	0.1001	0.1187	0.1571	0.1147	0.1460	0.2063	0.1401	0.1773
	0.8	0.1143	0.0948	0.1118	0.1385	0.1231	0.1133	0.1811	0.1620	0.1201
	0.9	0.1107	0.0910	0.1085	0.1342	0.1292	0.0903	0.1865	0.1820	0.0819

References

Aliprantis, C. D., and K. C. Border (2006): *Infinite Dimensional Analysis – A Hitchhiker's Guide*. Springer-Verlag, Berlin.

Beresteanu, A., and F. Molinari (2008): "Asymptotic Properties for a Class of Partially Identified Models," *Econometrica*, 76(4), 763–814.

Bickel, P. J., C. A. Klassen, Y. Ritov, and J. A. Wellner (1993): *Efficient and Adaptive Estimation for Semiparametric Models*. Springer, New York.

Bontemps, C., T. Magnac, and E. Maurin (2012): "Set Identified Linear Models," *Econometrica*, 80(3), 1129–1155.

Hansen, B. E. (2008): "Uniform Convergence Rates for Kernel Estimation with Dependent Data," *Econometric Theory*, 24(03).

Kosorok, M. R. (2008): *Introduction to Empirical Processes and Semiparametric Inference*. Springer, New York.

Lewbel, A. (2000): "Asymptotic trimming for bounded density plug-in estimators," Discussion Paper, Boston College.

Newey, W. (1991): "Efficient estimation of Tobit models under conditional symmetry," *Nonparametric and Semiparametric Methods in Econometrics and Statistics*, pp. 291–336.

Newey, W., and T. Stoker (1993): "Efficiency of weighted average derivative estimators and index models," *Econometrica*.

Pakes, A., and D. Pollard (1989): "Simulation and the asymptotics of optimization estimators," *Econometrica*, pp. 1027–1057.

Powell, J., J. Stock, and T. Stoker (1989): "Semiparametric estimation of index coefficients," *Econometrica*, pp. 1403–1430.

Rockafellar, R. T., and R. J.-B. Wets (2005): *Variational Analysis*. Berlin: Springer, 2nd edn.

Schneider, R. (1993): *Convex bodies: the Brunn-Minkowski theory*, vol. 44. Cambridge Univ Pr.

Sherman, R. P. (1994a): "Maximal inequalities for degenerate U-processes with applications to optimization estimators," *The Annals of Statistics*, pp. 439–459.

——— (1994b): "U-processes in the analysis of a generalized semiparametric regression estimator," *Econometric Theory*, 10, 372–372.

Stoker, T. (1986): "Consistent estimation of scaled coefficients," *Econometrica*.

van der Vaart, A. W., and J. A. Wellner (1996): *Weak Convergence and Empirical Processes: with Applications to Statistics*. Springer, New York.