Online Appendix

A Preliminaries

We first introduce relevant concepts on limit experiments and asymptotic representations. We refer the reader to van der Vaart (1991, 1998) for comprehensive treatments.

An *experiment* is a measurable space equipped with a family of probability measures. Let $\mathcal{E}_n = (\mathcal{X}^n, \mathcal{B}_n, P_{n,h} : h \in \mathbb{R}^K)$ where $\mathcal{B}_n$ denotes the Borel σ-algebra on $\mathcal{X}^n$. Each $\mathcal{E}_n$ corresponds to observing a random sample $X_1, \ldots, X_n$ from $P_{\mu_0 + h/\sqrt{n}}$ with parameter $h \in \mathbb{R}^K$. Also let $\mathcal{E} = (\mathbb{R}^K, \mathcal{B}_K, N(h, I_{\mu_0}^{-1}) : h \in \mathbb{R}^K)$ where $\mathcal{B}_K$ denotes the Borel σ-algebra on $\mathbb{R}^K$. This experiment corresponds to observing a single $Z \sim N(h, I_{\mu_0}^{-1})$ with unknown h. We suppress dependence of $\mathcal{E}_n$, $\mathcal{E}$, $P_{n,h}$, and N_h on μ_0 to simplify notation.

The sequence of experiments $\{\mathcal{E}_n\}$ is said to converge to the *limit experiment* $\mathcal{E}$ if for all finite subsets H of $\mathbb{R}^K$ and every $h_0 \in \mathbb{R}^K$, we have

$$\left(\frac{dP_{n,h}}{dP_{n,h_0}}(X^n)\right)_{h\in H} \overset{P_{n,h_0}}{\rightsquigarrow} \left(\frac{dN(h, I_{\mu_0}^{-1})}{dN(h_0, I_{\mu_0}^{-1})}(Z)\right)_{h\in H},$$

where $\overset{P_{n,h}}{\rightsquigarrow}$ denotes convergence in distribution along $\{P_{n,h}\}$. This convergence holds under Assumption 2(i)-(iii) (van der Vaart, 1998, Example 7.15 and Corollary 9.5).

Because $\mathcal{E}$ is dominated,[16] we can invoke an asymptotic representation theorem of van der Vaart (1991). For each $\{\delta_n\} \in \mathbb{D}$ there exists a *matching decision* $\delta_\infty(Z, U)$ given by a measurable map $\delta_\infty : \mathbb{R}^K \times [0,1] \to \mathcal{D}$, the observation $Z \sim N(h, I_{\mu_0}^{-1})$, and an independent randomization $U \sim \text{Uniform}[0,1]$, with the property that for each $h \in \mathbb{R}^K$,

$$\lim_{n\to\infty} P_{n,h}(\delta_n(X^n) = d) = \mathbb{P}_h(\delta_\infty(Z, U) = d), \quad d \in \mathcal{D},$$

where $\mathbb{P}_h$ is the product of the $N(h, I_{\mu_0}^{-1})$ and $U[0,1]$ distributions. Although $\delta_\infty(Z, U)$ depends on the randomization U in general, under our regularity conditions this randomization is unnecessary and any optimal decision depends only on Z.

[16] In particular, each $N(h, I_{\mu_0}^{-1})$ distribution has density with respect to Lebesgue measure.

B Proof of Main Results

B.1 Proofs for Section 4

We first state and prove a lemma which is used to prove Theorems 1 and 2. The first part shows that if $\pi \in \Pi$, then $\{\delta_n^*(\cdot\,;\pi)\} \in \mathbb{D}$ and its matching decision is independent of the prior π and nonrandomized. Thus, we drop dependence of the matching decision on π and U and simply write $\delta_\infty^*(Z)$. In the second part, we derive expressions for criteria (11) and (12) in terms of the limit experiment. Finally, in the third part, we show that our optimality criterion (12) is minimized for matching decisions corresponding to $\{\delta_n^*(\cdot\,;\pi)\}$ for $\pi \in \Pi$.

Recall that $\dot{R}_{d,\mu_0}$ denotes the directional derivative of $R(d,\mu)$ at μ_0, $Z \sim N(h, I_{\mu_0}^{-1})$ represents the asymptotic behavior of an efficient estimator of μ, $Z^* \sim N(0, I_{\mu_0}^{-1})$ independently of Z represents asymptotic uncertainty about the local parameter h, and $\mathbb{E}^*$ denotes expectation with respect to Z^*.

Lemma 1 *Suppose that Assumptions 1 and 2 hold. Fix any $\mu_0 \in \mathcal{M}$ for which there are no first-order ties.*

(i) If $\pi \in \Pi$, then $\{\delta_n^*(\cdot\,;\pi)\} \in \mathbb{D}$ and the matching decision of $\{\delta_n^*(\cdot\,;\pi)\}$ is $\delta_\infty^*(Z)$ with

$$\delta_\infty^*(z) = \operatorname*{argmin}_{d \in \underline{\mathcal{D}}_{\mu_0}} \mathbb{E}^*\left[\dot{R}_{d,\mu_0}[Z^* + z]\right]$$

for almost every $z \in \mathbb{R}^K$.

(ii) For any $\{\delta_n\} \in \mathbb{D}$,

$$\mathcal{R}(\{\delta_n\};\mu_0,h) = \sum_{d \in \underline{\mathcal{D}}_{\mu_0}} \left(\dot{R}_{d,\mu_0}[h] - \min_{i \in \underline{\mathcal{D}}_{\mu_0}} \dot{R}_{i,\mu_0}[h]\right) \mathbb{P}_h(\delta_\infty(Z,U) = d),$$

$$\mathcal{R}(\{\delta_n\};\mu_0) = \int \left(\sum_{d \in \underline{\mathcal{D}}_{\mu_0}} \left(\dot{R}_{d,\mu_0}[h] - \min_{i \in \underline{\mathcal{D}}_{\mu_0}} \dot{R}_{i,\mu_0}[h]\right) \mathbb{P}_h(\delta_\infty(Z,U) = d) \right) dh, \quad (29)$$

where $\delta_\infty(Z,U)$ is the matching decision of $\{\delta_n\}$.

(iii) The matching decision $\delta_\infty^*(Z)$ in (i) minimizes the right-hand side of (29) over all matching decisions $\delta_\infty(Z,U)$ of sequences $\{\delta_n\} \in \mathbb{D}$.

Proof of Lemma 1. **Part (i):** Let A_n denote the event $\delta_n^*(X^n;\pi) \in \underline{\mathcal{D}}_{\mu_0}$. By Lemma 11, we have $\lim_{n\to\infty} \sqrt{n}\, P_{n,h}(A_n^c) = 0$ for all $h \in \mathbb{R}^K$. Hence, condition (ii) in Definition 1 holds. In particular, if $\underline{\mathcal{D}}_{\mu_0} \subsetneq \mathcal{D}$, then

$$\lim_{n\to\infty} P_{n,h}(\delta_n^*(X^n;\pi) = d) = 0 \quad \text{all } h \in \mathbb{R}^K \text{ and } d \notin \underline{\mathcal{D}}_{\mu_0}. \quad (30)$$

Hence, if $\underline{\mathcal{D}}_{\mu_0} = \{d\}$ we have $\lim_{n\to\infty} P_{n,h}(\delta_n^*(X^n;\pi) = d) = 1$.

It remains to characterize $\lim_{n\to\infty} P_{n,h}(\delta_n^*(X^n;\pi) = d)$ for $d \in \underline{\mathcal{D}}_{\mu_0}$ when $|\underline{\mathcal{D}}_{\mu_0}| \geq 2$. We require a rule to break ties when $\operatorname{argmin}_{d\in\mathcal{D}} \bar{R}_n(d)$ is not a singleton. We shall take the smallest index d among the set of minimizers, though it will be seen below that the matching decision doesn't depend on the tie-breaking rule. Thus, any (possibly randomized) tie-breaking rule would lead to the same matching decision. Under this tie-breaking rule, we have

$$\mathbb{I}[\delta_n^*(X^n;\pi) = d] = \mathbb{I}\left[\bar{R}_n(d) < \min_{i\in\underline{\mathcal{D}}_{\mu_0}:i<d} \bar{R}_n(i), \text{ and } \bar{R}_n(d) \leq \min_{i\in\underline{\mathcal{D}}_{\mu_0}:i\geq d} \bar{R}_n(i)\right] \text{ on } A_n,$$

where the minimum over an empty set is $+\infty$. Since $|\Pr(A) - \Pr(B)| \leq 2\Pr(C^c)$ if $A = B$ holds on C, we obtain

$$\left|P_{n,h}(\delta_n^*(X^n;\pi) = d) - P_{n,h}\left(\bar{R}_n(d) < \min_{i\in\underline{\mathcal{D}}_{\mu_0}:i<d} \bar{R}_n(i), \text{ and } \bar{R}_n(d) \leq \min_{i\in\underline{\mathcal{D}}_{\mu_0}:i\geq d} \bar{R}_n(i)\right)\right| \leq 2P_{n,h}(A_n^c) \to 0.$$

As $R(d,\mu_0) = R(d',\mu_0)$ for all $d, d' \in \underline{\mathcal{D}}_{\mu_0}$, we have

$$\begin{aligned}
&P_{n,h}\left(\bar{R}_n(d) < \min_{i\in\underline{\mathcal{D}}_{\mu_0}:i<d} \bar{R}_n(i), \text{ and } \bar{R}_n(d) \leq \min_{i\in\underline{\mathcal{D}}_{\mu_0}:i\geq d} \bar{R}_n(i)\right) \\
&= P_{n,h}\Bigg(\int \sqrt{n}(R(d,\mu) - R(d,\mu_0))\, d\pi_n(\mu) < \min_{i\in\underline{\mathcal{D}}_{\mu_0}:i<d} \int \sqrt{n}(R(i,\mu) - R(i,\mu_0))\, d\pi_n(\mu), \\
&\quad \text{and } \int \sqrt{n}(R(d,\mu) - R(d,\mu_0))\, d\pi_n(\mu) \leq \min_{i\in\underline{\mathcal{D}}_{\mu_0}:i\geq d} \int \sqrt{n}(R(i,\mu) - R(i,\mu_0))\, d\pi_n(\mu)\Bigg).
\end{aligned}$$

Now by Proposition 2, we have

$$\left(\int \sqrt{n}(R(d,\mu) - R(d,\mu_0))\, d\pi_n(\mu)\right)_{d\in\underline{\mathcal{D}}_{\mu_0}} \overset{P_{n,h}}{\rightsquigarrow} \left(\mathbb{E}^*\left[\dot{R}_{d,\mu_0}\left[Z^* + Z\right]\middle| Z\right]\right)_{d\in\underline{\mathcal{D}}_{\mu_0}}.$$

Define $U_d = \{(u_i)_{i\in\underline{\mathcal{D}}_{\mu_0}} : u_d < \min_{i\in\underline{\mathcal{D}}_{\mu_0}:i<d} u_i, \text{ and } u_d \leq \min_{i\in\underline{\mathcal{D}}_{\mu_0}:i\leq d} u_i\} \subset \mathbb{R}^{|\underline{\mathcal{D}}_{\mu_0}|}$ for $d \in \underline{\mathcal{D}}_{\mu_0}$. These sets form a partition of $\mathbb{R}^{|\underline{\mathcal{D}}_{\mu_0}|}$. Evidently, $\operatorname{int}(U_d) \supseteq \{(u_i)_{i\in\underline{\mathcal{D}}_{\mu_0}} : u_d < \min_{i\in\underline{\mathcal{D}}_{\mu_0}:i\neq d} u_i\}$. Thus, boundary points of U_d are points at which $\operatorname{argmin}_{i\in\underline{\mathcal{D}}_{\mu_0}} u_i$ is not a singleton. Since $\operatorname{argmin}_{d\in\underline{\mathcal{D}}_{\mu_0}} \mathbb{E}^*[\dot{R}_{d,\mu_0}[Z^* + z]]$ is a singleton for almost every z, it follows that

$$\mathbb{P}_h\left(\left(\mathbb{E}^*\left[\dot{R}_{d,\mu_0}\left[Z^* + Z\right]\middle| Z\right]\right)_{d\in\underline{\mathcal{D}}_{\mu_0}} \in \partial U_d\right) = 0.$$

Hence, by the Portmanteau theorem and the preceding four displays, for all $d \in \underline{\mathcal{D}}_{\mu_0}$ we have

$$\lim_{n\to\infty} P_{n,h}(\delta_n^*(X^n;\pi) = d) = \mathbb{P}_h\Big(\Big(\mathbb{E}^*[\dot{R}_{d,\mu_0}[Z^* + Z]|Z] < \min_{i\in\underline{\mathcal{D}}_{\mu_0}: i<d} \mathbb{E}^*[\dot{R}_{i,\mu_0}[Z^* + Z]|Z]\Big), \text{ and } \Big(\mathbb{E}^*[\dot{R}_{d,\mu_0}[Z^* + Z]|Z] \le \min_{i\in\underline{\mathcal{D}}_{\mu_0}: i\ge d} \mathbb{E}^*[\dot{R}_{i,\mu_0}[Z^* + Z]|Z]\Big)\Big). \quad (31)$$

Thus, we have shown that $\{\delta_n^*(\,\cdot\,;\pi)\} \in \mathbb{D}$.

We can now invoke Theorem 3.1 of van der Vaart (1991) to guarantee existence of the matching form $\delta_\infty^*(Z, U; \pi)$ of $\{\delta_n^*(\,\cdot\,;\pi)\}$. By (30), we see that $\delta_\infty^*(z, u; \pi)$ must take values in $\underline{\mathcal{D}}_{\mu_0}$ for almost every (z, u). Moreover, as $\operatorname{argmin}_{d\in\underline{\mathcal{D}}_{\mu_0}} \mathbb{E}^*[\dot{R}_{d,\mu_0}[Z^* + z]]$ is a singleton for almost every z, we may restate (31) as

$$\lim_{n\to\infty} P_{n,h}(\delta_n^*(X^n;\pi) = d) = \mathbb{P}_h\Big(d = \operatorname*{argmin}_{i\in\underline{\mathcal{D}}_{\mu_0}} \mathbb{E}^*[\dot{R}_{i,\mu_0}[Z^* + Z]|Z]\Big).$$

Hence, we may take $\delta_\infty^*(z, u; \pi) = \operatorname{argmin}_{i\in\underline{\mathcal{D}}_{\mu_0}} \mathbb{E}^*[\dot{R}_{i,\mu_0}[Z^* + z]]$ for almost every (z, u). As $\delta_\infty^*(z, u; \pi)$ doesn't depend on u or π, we simply write $\delta_\infty^*(z)$.

Part (ii): By Assumption 1(i), $\min_{d\in\mathcal{D}} R(d, \mu_0 + h/\sqrt{n}) = \min_{d\in\underline{\mathcal{D}}_{\mu_0}} R(d, \mu_0 + h/\sqrt{n})$ holds for all n sufficiently large. Hence, for all n sufficiently large,

$$\begin{aligned}
&\sqrt{n}\left(\mathbb{E}_{n,h}\left[R\left(\delta_n(X^n), \mu_0 + h/\sqrt{n}\right)\right] - \min_{i\in\mathcal{D}} R\left(i, \mu_0 + h/\sqrt{n}\right)\right) \\
&\quad = \sum_{d\in\underline{\mathcal{D}}_{\mu_0}} P_{n,h}\left(\delta_n(X^n) = d\right)\sqrt{n}\Big(R(d, \mu_0 + h/\sqrt{n}) - \min_{i\in\underline{\mathcal{D}}_{\mu_0}} R(i, \mu_0 + h/\sqrt{n})\Big) \\
&\qquad + \sum_{d\notin\underline{\mathcal{D}}_{\mu_0}} \sqrt{n}\, P_{n,h}\left(\delta_n(X^n) = d\right)\Big(R(d, \mu_0 + h/\sqrt{n}) - \min_{i\in\underline{\mathcal{D}}_{\mu_0}} R(i, \mu_0 + h/\sqrt{n})\Big),
\end{aligned}$$

where the second sum is zero if $\mathcal{D} = \underline{\mathcal{D}}_{\mu_0}$; otherwise, it converges to zero as $n \to \infty$ by Definition 1(ii) and Assumption 1(i). For the first sum, Theorem 3.1 of van der Vaart (1991) implies

$$\lim_{n\to\infty} P_{n,h}(\delta_n(X^n) = d) = \mathbb{P}_h(\delta_\infty(Z, U) = d), \quad d \in \underline{\mathcal{D}}_{\mu_0},$$

where $\delta_\infty(Z, U)$ is the matching decision of $\{\delta_n\}$. As $R(d, \mu_0) = R(d', \mu_0)$ for all $d, d' \in \underline{\mathcal{D}}_{\mu_0}$,

we have

$$\begin{aligned}
&\sqrt{n}\Big(R(d,\mu_0+h/\sqrt{n}) - \min_{i\in\underline{\mathcal{D}}_{\mu_0}} R(i,\mu_0+h/\sqrt{n})\Big) \\
&= \sqrt{n}\left(R(d,\mu_0+h/\sqrt{n}) - R(d,\mu_0)\right) - \sqrt{n}\left(\min_{i\in\underline{\mathcal{D}}_{\mu_0}} R(i,\mu_0+h/\sqrt{n}) - \min_{i\in\underline{\mathcal{D}}_{\mu_0}} R(i,\mu_0)\right). \\
&= \dot{R}_{d,\mu_0}[h] - \min_{i\in\underline{\mathcal{D}}_{\mu_0}} \dot{R}_{i,\mu_0}[h],
\end{aligned}$$

where the final line follows from the fact that if $\{f_i\}_{i\in\mathcal{I}}$ is a finite set of functions that are directionally differentiable at μ_0 with $f_i(\mu_0) = f_{i'}(\mu_0)$ for all $i, i' \in \mathcal{I}$, then $\min_{i\in\mathcal{I}} f_i(\mu)$ has directional derivative $\min_{i\in\mathcal{I}} \dot{f}_{i,\mu_0}[h]$ at μ_0.

Part (iii): Let $g_{d,\mu_0}[h] = \dot{R}_{d,\mu_0}[h] - \min_{i\in\underline{\mathcal{D}}_{\mu_0}} \dot{R}_{i,\mu_0}[h]$. By part (ii), we have

$$\begin{aligned}
\mathcal{R}(\{\delta_n\};\mu_0) &= \int \left(\sum_{d\in\underline{\mathcal{D}}_{\mu_0}} g_{d,\mu_0}[h]\mathbb{P}_h(\delta_\infty(Z,U)=d)\right)dh \\
&\propto \sum_{d\in\underline{\mathcal{D}}_{\mu_0}} \int\int\int_0^1 \mathbb{I}[\delta_\infty(z,u)=d] e^{-\frac{1}{2}(z-h)^T I_{\mu_0}(z-h)}\, du\, dz\, (g_{d,\mu_0}[h])\, dh \\
&= \int_0^1 \int \sum_{d\in\underline{\mathcal{D}}_{\mu_0}} \mathbb{I}[\delta_\infty(z,u)=d] \left(\int g_{d,\mu_0}[h] e^{-\frac{1}{2}(z-h)^T I_{\mu_0}(z-h)}\, dh\right) dz\, du.
\end{aligned}$$

Changing the order of integration in the final line above is justified by Tonelli's theorem, which only requires non-negativity of the integrand. Define

$$\mathcal{I}(z) = \operatorname*{argmin}_{i\in\underline{\mathcal{D}}_{\mu_0}} \int \dot{R}_{i,\mu_0}[h] e^{-\frac{1}{2}(z-h)^T I_{\mu_0}(z-h)}\, dh \equiv \operatorname*{argmin}_{i\in\underline{\mathcal{D}}_{\mu_0}} \mathbb{E}^*\left[\dot{R}_{i,\mu_0}[Z^*+z]\right]. \tag{32}$$

Evidently, any δ_∞ with $\delta_\infty(z,u) \in \mathcal{I}(z)$ for almost every (z,u) minimizes the criterion. Since there are no first-order ties at μ_0, the set $\mathcal{I}(z)$ is a singleton for almost every z and the criterion is minimized by $\delta^*_\infty(z)$ derived in part (i). ■

Proof of Theorem 1. By Lemma 1, for any $\{\delta_n\} \in \mathbb{D}$, we have

$$\begin{aligned}
\mathcal{R}(\{\delta_n\};\mu_0) &= \int \left(\sum_{d\in\underline{\mathcal{D}}_{\mu_0}} \left(\dot{R}_{d,\mu_0}[h] - \min_{i\in\underline{\mathcal{D}}_{\mu_0}} \dot{R}_{i,\mu_0}[h]\right) \mathbb{P}_h(\delta_\infty(Z,U)=d)\right)dh \\
&\geq \int \left(\sum_{d\in\underline{\mathcal{D}}_{\mu_0}} \left(\dot{R}_{d,\mu_0}[h] - \min_{i\in\underline{\mathcal{D}}_{\mu_0}} \dot{R}_{i,\mu_0}[h]\right) \mathbb{P}_h(\delta^*_\infty(Z)=d)\right)dh = \mathcal{R}(\{\delta^*_n(\,\cdot\,;\pi)\};\mu_0).
\end{aligned}$$

Result (i) now follows by taking the infimum over $\{\delta_n\} \in \mathbb{D}$ and noting that $\{\delta_n^*(\cdot\,;\pi)\} \in \mathbb{D}$ by Lemma 1(i).

Result (ii) follows from the fact that the matching forms δ_∞^* of $\{\delta_n^*(\cdot\,;\pi)\}$ derived in Lemma 1(i) do not depend on the prior $\pi \in \Pi$.

Finally, if $\{\delta_n\} \in \mathbb{D}$ is asymptotically equivalent to $\{\delta_n^*(\cdot\,;\pi)\}$, then $\{\delta_n\}$ must also have matching form $\delta_\infty^*(Z)$. Hence, by similar arguments to the above, we have $\mathcal{R}(\{\delta_n'\};\mu_0) \geq \mathcal{R}(\{\delta_n\};\mu_0)$ for any $\{\delta_n'\} \in \mathbb{D}$. Taking the infimum over $\{\delta_n'\} \in \mathbb{D}$ yields result (iii). ■

Proof of Theorem 2. First note that asymptotic equivalence cannot fail if $|\underline{\mathcal{D}}_{\mu_0}| = 1$, since if $\underline{\mathcal{D}}_{\mu_0} = \{d\}$, say, then $\lim_{n\to\infty} P_{n,h}(\delta_n(X^n) = d) = 1$ must hold for any $\{\delta_n\} \in \mathbb{D}$.

Now suppose $|\underline{\mathcal{D}}_{\mu_0}| \geq 2$. Let $g_{d,\mu_0}[h] = \dot{R}_{d,\mu_0}[h] - \min_{i\in\underline{\mathcal{D}}_{\mu_0}} \dot{R}_{i,\mu_0}[h]$. By Lemma 1(ii)-(iii), it suffices to show

$$\int \left(\sum_{d\in\underline{\mathcal{D}}_{\mu_0}} g_{d,\mu_0}[h]\mathbb{P}_h(\delta_\infty(Z,U) = d) \right) dh > \int \left(\sum_{d\in\underline{\mathcal{D}}_{\mu_0}} g_{d,\mu_0}[h]\mathbb{P}_h(\delta_\infty^*(Z) = d) \right) dh,$$

where $\delta_\infty(Z,U)$ is the matching form of $\{\delta_n\}$ and $\delta_\infty^*(Z)$ is the matching form of $\{\delta_n^*(\cdot\,;\pi)\}$ from Lemma 1(i). Since both integrands are non-negative, by Tonelli's theorem we may restate this inequality as

$$\int_0^1 \int \sum_{d\in\underline{\mathcal{D}}_{\mu_0}} \mathbb{I}[\delta_\infty(z,u) = d] \left(\int g_{d,\mu_0}[h]\, e^{-\frac{1}{2}(z-h)^T I_{\mu_0}(z-h)}\, dh \right) dz\, du$$
$$> \int_0^1 \int \sum_{d\in\underline{\mathcal{D}}_{\mu_0}} \mathbb{I}[\delta_\infty^*(z) = d] \left(\int g_{d,\mu_0}[h]\, e^{-\frac{1}{2}(z-h)^T I_{\mu_0}(z-h)}\, dh \right) dz\, du.$$

As $\{\delta_n\}$ and $\{\delta_n^*(\cdot\,;\pi)\}$ are not asymptotically equivalent, we have

$$\lim_{n\to\infty} P_{n,h_0}(\delta_n(X^n) = d) \neq \lim_{n\to\infty} P_{n,h_0}(\delta_n^*(X^n;\pi) = d)$$

for some $h_0 \in \mathbb{R}^K$ and some $d \in \mathcal{D}$. As $\{\delta_n\}$ and $\{\delta_n^*(\cdot\,;\pi)\}$ are both elements of $\mathbb{D}$, we can use Theorem 3.1 of van der Vaart (1991) and restate this in terms of their matching forms $\delta_\infty(Z,U)$ and $\delta_\infty^*(Z)$:

$$\mathbb{P}_{h_0}(\delta_\infty(Z,U) = d) \neq \mathbb{P}_{h_0}(\delta_\infty^*(Z) = d).$$

Hence, it must be the case that $\delta_\infty(z,u)$ and $\delta_\infty^*(z)$ disagree on a set S of positive Lebesgue measure. As $\delta_\infty^*(z) \in \mathcal{I}(z)$ for almost every z, and $\mathcal{I}(z)$ is a singleton for almost every z as there are no first-order ties, we must have that $\delta_\infty(z,u) \notin \mathcal{I}(z)$ on a set S' of positive

Lebesgue measure. Now by definition of $\mathcal{I}(z)$, we also have that

$$\sum_{d\in\underline{\mathcal{D}}_{\mu_0}} \mathbb{I}[\delta_\infty(z,u)=d]\left(\int g_{d,\mu_0}[h]\, e^{-\frac{1}{2}(z-h)^T I_{\mu_0}(z-h)}\, dh\right)$$
$$\geq \sum_{d\in\underline{\mathcal{D}}_{\mu_0}} \mathbb{I}[\delta^*_\infty(z)=d]\left(\int g_{d,\mu_0}[h]\, e^{-\frac{1}{2}(z-h)^T I_{\mu_0}(z-h)}\, dh\right),$$

holds for almost every (z,u), with strict inequality for (z,u) in a set of positive measure. Integrating with respect to (z,u) yields the strict inequality above. ■

Proof of Corollary 1. Immediate from Theorem 1 and the argument given in the text preceding the statement of this corollary. ■

Proof of Proposition 1. Let $g_{d,\mu_0}[h] = \dot{R}_{d,\mu_0}[h] - \min_{i\in\underline{\mathcal{D}}_{\mu_0}} \dot{R}_{i,\mu_0}[h]$. Note $g_{d,\mu_0}[h] \geq 0$. By Lemma 1, we have

$$\mathcal{R}(\{\delta^*_n(\,\cdot\,;\pi)\};\mu_0) \leq \sum_{d\in\underline{\mathcal{D}}_{\mu_0}} \int g_{d,\mu_0}[h]\, \mathbb{P}_h\left(\operatorname{argmin}_{i\in\underline{\mathcal{D}}_{\mu_0}} \dot{R}_{i,\mu_0}[Z]=d\right) dh,$$

where interchanging summation and integration is justified by Tonelli's theorem. Since $h\in\mathbb{R}$, we split the integral into an integral over $[0,\infty)$ and an integral over $(-\infty,0)$. Suppose $h\geq 0$. For each $d\in\underline{\mathcal{D}}_{\mu_0}$, note $g_{d,\mu_0}[h] = h g_{d,\mu_0}[1]$ by positive homogeneity. Suppose $g_{d,\mu_0}[1]>0$ for some $d\in\underline{\mathcal{D}}_{\mu_0}$, else the integral is zero. Then $d = \operatorname{argmin}_{i\in\underline{\mathcal{D}}_{\mu_0}} \dot{R}_{i,\mu_0}[z]$ implies $g_{d,\mu_0}[z]=0$. And since $g_{d,\mu_0}[1]>0$ and $g_{d,\mu_0}[\,\cdot\,]$ is positively homogeneous, $g_{d,\mu_0}[z]=0$ in turn implies $z\leq 0$. Hence,

$$\mathbb{P}_h(\operatorname{argmin}_{i\in\underline{\mathcal{D}}_{\mu_0}} \dot{R}_{i,\mu_0}[Z]=d) \leq \mathbb{P}_h(Z\leq 0) = \Phi(-h/\sigma),$$

where $\sigma = I_{\mu_0}^{-1/2} > 0$, and Φ is the CDF of the $N(0,1)$ distribution. Then

$$\int_0^\infty g_{d,\mu_0}[h]\, \mathbb{P}_h\left(\operatorname{argmin}_{i\in\underline{\mathcal{D}}_{\mu_0}} \dot{R}_{i,\mu_0}[Z]=d\right) dh \leq g_d[1]\int_0^\infty h\,\Phi(-h/\sigma)\, dh < \infty.$$

One may similarly show the integral over $(-\infty,0)$ is finite. ■

We first present a preliminary result before turning to the proof of Theorem 3.

Lemma 2 *Suppose that Assumption 1 and 2 hold. Then $\{\delta_n\}\in\mathbb{D}$ is σ-optimal at μ_0 if its*

matching decision is $\delta^*_\sigma(Z,U)$ *with*

$$\delta^*_\sigma(z,u) \in \operatorname*{argmin}_{d\in\underline{\mathcal{D}}_{\mu_0}} \mathbb{E}^*_\sigma\left[\dot{R}_{d,\mu_0}[Z^*_\sigma + (I+(\sigma I_{\mu_0})^{-1})^{-1}z]\right]$$

for almost every $(z,u)\in\mathbb{R}^K\times[0,1]$*, where* $\mathbb{E}^*_\sigma$ *denotes expectation with respect to* $Z^*_\sigma \sim N(0,(I_{\mu_0}+\sigma^{-1}I)^{-1})$.

Proof of Lemma 2. By Lemma 1(ii), for any $\{\delta_n\}\in\mathbb{D}$ we have

$$\mathcal{R}_\sigma(\{\delta_n\};\mu_0) \propto \int \sum_{d\in\underline{\mathcal{D}}_{\mu_0}} g_{d,\mu_0}[h]\,\mathbb{P}_h(\delta_\infty(Z,U)=d)e^{-\frac{1}{2\sigma}\|h\|^2}\,dh,$$

where $g_{d,\mu_0}[h] = \dot{R}_{d,\mu_0}[h] - \min_{i\in\underline{\mathcal{D}}_{\mu_0}}\dot{R}_{i,\mu_0}[h]$ and $\delta_\infty(Z,U)$ is the matching decision of $\{\delta_n\}$. By Tonelli's theorem, we have

$$\mathcal{R}_\sigma(\{\delta_n\};\mu_0) \propto \int_0^1\int \sum_{d\in\underline{\mathcal{D}}_{\mu_0}} \mathbb{I}[\delta_\infty(z,u)=d]\left(\int g_{d,\mu_0}[h]e^{-\frac{1}{2}(z-h)^T I_{\mu_0}(z-h)-\frac{1}{2\sigma}\|h\|^2}\,dh\right)dz\,du.$$

Therefore, $\mathcal{R}_\sigma(\{\delta_n\};\mu_0)$ is minimized when

$$\delta_\infty(z,u)\in\operatorname*{argmin}_{i\in\underline{\mathcal{D}}_{\mu_0}}\int \dot{R}_{i,\mu_0}[h]e^{-\frac{1}{2}(z-h)^T I_{\mu_0}(z-h)-\frac{1}{2\sigma}\|h\|^2}\,dh$$

for almost every (z,u). The result follows by completing the square. ■

Proof of Theorem 3. Lemma 1(i) implies

$$\lim_{n\to\infty} P_{n,h}\left(\delta^*_n(X^n;\pi)=d\right) = \mathbb{P}_h(\delta^*(Z)=d),$$

for all $d\in\mathcal{D}$ and $h\in\mathbb{R}^K$, where

$$\delta^*(z) = \mathcal{I}(z) := \operatorname{argmin}_{i\in\underline{\mathcal{D}}_{\mu_0}}\mathbb{E}^*\left[\dot{R}_{i,\mu_0}[Z^*+z]\right]$$

for almost every z, since there are no first-order ties. Similarly, Lemma 2 implies

$$\lim_{n\to\infty} P_{n,h}\left(\delta^*_{n,\sigma}(X^n)=d\right) = \mathbb{P}_h(\delta^*_\sigma(Z,U)=d),$$

for all $d \in \mathcal{D}$ and $h \in \mathbb{R}^K$, where

$$\delta^*_\sigma(z,u) \in \mathcal{I}_\sigma(z) := \text{argmin}_{i \in \mathcal{D}_{\mu_0}} \mathbb{E}^*_\sigma \left[\dot{R}_{i,\mu_0}[Z^*_\sigma + (I + (\sigma I_{\mu_0})^{-1})^{-1} z] \right]$$

for almost every $(z,u) \in \mathbb{R}^K \times [0,1]$.

For almost every z, we have by dominated convergence that

$$\lim_{\sigma \to \infty} \mathbb{E}^*_\sigma \left[\dot{R}_{i,\mu_0}[Z^*_\sigma + (I + (\sigma I_{\mu_0})^{-1})^{-1} z] \right] = \mathbb{E}^* \left[\dot{R}_{i,\mu_0}[Z^* + z] \right].$$

Hence, $\lim_{\sigma \to \infty} \mathcal{I}_\sigma(z) = \mathcal{I}(z)$ for almost every z, and so $\lim_{\sigma \to \infty} \delta^*_\sigma(z,u) = \delta^*(z)$ for almost every (z,u). Hence, by dominated convergence and Theorem 3.1 of van der Vaart (1991),

$$\begin{aligned} \lim_{n \to \infty} P_{n,h}(\delta^*_n(X^n;\pi) = d) &= \mathbb{P}_h(\delta^*(Z) = d) \\ &= \lim_{\sigma \to \infty} \mathbb{P}_h(\delta^*_\sigma(Z,U) = d) = \lim_{\sigma \to \infty} \lim_{n \to \infty} P_{n,h}\left(\delta^*_{n,\sigma}(X^n) = d\right), \end{aligned}$$

as required. ■

B.2 Proofs for Section 5

Once we fix (μ_0, η_0), the least favorable model $\{P_{\beta(t)} : t \in \mathcal{M}_{(\mu_0,\eta_0)}\}$ is a parametric model. By Assumption 3(i)-(iii) and Example 7.15 and Corollary 9.5 of van der Vaart (1998), the sequence of experiments $\{\mathcal{E}_n\}$, with $\mathcal{E}_n = (\mathcal{X}^n, \mathcal{B}_n, P_{n,h} : h \in \mathbb{R}^K)$ and $P_{n,h}$ denoting the distribution of X^n under $P_{\beta(\mu_0 + h/\sqrt{n})}$, converges to a limit experiment $\mathcal{E} = (\mathbb{R}^K, \mathcal{B}_K, N(h, I^{-1}_{(\mu_0,\eta_0)}) : h \in \mathbb{R}^K)$. Because $\mathcal{E}$ is dominated, the asymptotic representation theorem of van der Vaart (1991) implies that each $\{\delta_n\} \in \mathbb{D}$ is matched by a $\delta_\infty(Z,U)$ given by a measurable map $\delta_\infty : \mathbb{R}^K \times [0,1] \to \mathcal{D}$, the observation $Z \sim N(h, I^{-1}_{(\mu_0,\eta_0)})$, and an independent randomization $U \sim \text{Uniform}[0,1]$, with the property that for each $h \in \mathbb{R}^K$,

$$\lim_{n \to \infty} P_{n,h}(\delta_n(X^n) = d) = \mathbb{P}_h(\delta_\infty(Z,U) = d), \quad d \in \mathcal{D},$$

where $\mathbb{P}_h$ is the product of the $N(h, I^{-1}_{(\mu_0,\eta_0)})$ and $U[0,1]$ distributions

Once these modifications are made, the proofs of Theorems 4 and 5 follow identical arguments to the proofs of Theorems 1 and 2, using Lemma 12 in place of Lemma 11 and Proposition 3 in place of Proposition 2.

C Asymptotic Distribution of the Posterior Mean of Directionally Differentiable Functions

In this section we derive the asymptotic distribution of the posterior mean of directionally differentiable functions. Let $f : \mathcal{M} \to \mathbb{R}^k$ be a directionally differentiable function of μ. The first main result is Proposition 2, which presents the asymptotic distribution of the centered and scaled posterior mean

$$\bar{f}_n := \int \sqrt{n}(f(\mu) - f(\mu_0))d\pi_n(\mu)$$

along sequences of local parameters $P_{n,h}$. In Proposition 3 we then extend this result to semiparametric models. Both results are of independent interest. The most closely related work is Kitagawa, Montiel Olea, Payne, and Velez (2020), which characterizes the asymptotic behavior of the posterior distribution of directionally differentiable functions. We instead characterize the (frequentist) asymptotic distribution of its posterior mean. For both results, we assume the following:

Assumption 4 *(i) f is integrable with respect to the prior π for μ;*
(ii) f is directionally differentiable at μ_0.

C.1 Parametric Models

We first consider parametric models as described in Section 4. We present some regularity conditions and preliminary lemmas before giving the main result.

Recall the notation from Section 4. Define $Z_{n,\mu_0} = \frac{1}{\sqrt{n}} \sum_{i=1}^n I_{\mu_0}^{-1} \dot{\ell}_{\mu_0}(X_i)$. Let $\tilde{\pi}_n$ denote the posterior induced by $\pi_n(\mu)$ under a change of variables from μ to $h = \sqrt{n}(\mu - \mu_0) - Z_{n,\mu_0}$. Let P^* denote the distribution of $Z^* \sim N(0, I_{\mu_0}^{-1})$ and let $\|\cdot\|_{TV}$ denote total variation distance.

Lemma 3 *Suppose that Assumptions 2(i)-(iii) and (v) hold, Assumption 4(ii) holds, and $\pi \in \Pi$. Then there exists a sequence $\{B_n\} \subset \mathbb{R}_+$ with $B_n \uparrow +\infty$, $B_n = o(\sqrt{n})$, such that for all n sufficiently large,*

$$\sup_{h:\|h\| \leq 2B_n} \left\| \sqrt{n} \left(f(\mu_0 + h/\sqrt{n}) - f(\mu_0) \right) - \dot{f}_{\mu_0}[h] \right\| \leq \frac{1}{B_n},$$

and

$$P_{n,\mu_0} \left(\|\tilde{\pi}_n - P^*\|_{TV} > \frac{1}{B_n^2} \right) \leq \frac{1}{B_n}.$$

Proof of Lemma 3. By Assumption 4(ii) and Proposition 3.3 of Shapiro (1990), for each $a \in \mathbb{N}$ there exists $n_{a,f} \in \mathbb{N}$ such that $n \geq n_{a,f}$ implies

$$\sup_{h:\|h\|\leq 2a} \left\| \sqrt{n}\left(f(\mu_0 + h/\sqrt{n}) - f(\mu_0)\right) - \dot{f}_{\mu_0}[h] \right\| \leq \frac{1}{a}.$$

Under Assumptions 2(i)-(iii) and since $\pi \in \Pi$, we may invoke Theorem 10.1 of van der Vaart (1998), using Assumption 2(v) and Proposition 6.2 of Clarke and Barron (1990) to deduce existence of tests with the desired properties. Further, we may use invariance of total variation under location changes to deduce that for all $a \in \mathbb{N}$ there exists $n_{a,\pi} \in \mathbb{N}$ such that $n \geq n_{a,\pi}$ implies

$$P_{n,\mu_0}\left(\|\tilde{\pi}_n - P^*\|_{TV} > \frac{1}{a^2} \right) \leq \frac{1}{a}.$$

The result follows by setting $B_n = \sup\{a \in \mathbb{N} : n \geq (n_{a,f} \vee n_{a,\pi})\}$ for $n \geq (n_{1,f} \vee n_{1,\pi})$, then letting B_n increase slower if necessary so that $B_n = o(\sqrt{n})$. ■

Let $\overset{P_{n,\mu_0}}{\to}$ denote convergence in probability under $\{P_{n,\mu_0}\}$. Let $\mathbb{E}^*$ denote expectation with respect to $Z^* \sim N(0, I_{\mu_0}^{-1})$ independent of Z_{n,μ_0}.

Lemma 4 *Suppose Assumptions 2 and 4 hold and $\pi \in \Pi$. Then*

$$\left\| \bar{f}_n - \mathbb{E}^*\left[\dot{f}_{\mu_0}\left[Z^* + Z_{n,\mu_0}\right] \middle| Z_{n,\mu_0} \right] \right\| \overset{P_{n,\mu_0}}{\to} 0$$

as $n \to \infty$.

Proof of Lemma 4. Let $\{B_n\}$ be as in Lemma 3. By a change of variables, we have

$$\begin{aligned} \bar{f}_n = &\int_{h:\|h\|\leq B_n} \sqrt{n}\left(f(\mu_0 + n^{-1/2}(Z_{n,\mu_0} + h)) - f(\mu_0)\right) d\tilde{\pi}_n(h) \\ &+ \int_{\mu:\|\sqrt{n}(\mu-\mu_0)-Z_{n,\mu_0}\|>B_n} \sqrt{n}\left(f(\mu) - f(\mu_0)\right) d\pi_n(\mu) =: \mathcal{J}_{1,n} + \mathcal{J}_{2,n}. \end{aligned}$$

Let A_n denote the event $\|Z_{n,\mu_0}\| \leq B_n$ and $\|\tilde{\pi}_n - P^*\|_{TV} \leq B_n^{-2}$. Assumption 2(iii) and Lemma 3 imply $\lim_{n\to\infty} P_{n,\mu_0}(A_n^c) = 0$. Again by Lemma 3, on A_n we have

$$\left\| \mathcal{J}_{1,n} - \int_{h:\|h\|\leq B_n} \dot{f}_{\mu_0}\left[Z_{n,\mu_0} + h\right] d\tilde{\pi}_n(h) \right\| \leq \frac{1}{B_n} \to 0.$$

Note that $\dot{f}_{\mu_0}[\,\cdot\,]$ is continuous under Assumption 4(ii) (Shapiro, 1990, Proposition 3.1) and

hence that

$$C_f := \sup_{u:\|u\|=1} \|\dot{f}_{\mu_0}[u]\| < \infty. \tag{33}$$

Moreover, $\dot{f}_{\mu_0}[\,\cdot\,]$ is positively homogeneous of degree one (Shapiro, 1990, p. 478). It therefore follows by Lemma 3 that on A_n we have

$$\left\| \int_{h:\|h\|\le B_n} \dot{f}_{\mu_0}\left[Z_{n,\mu_0}+h\right] d\tilde{\pi}_n(h) - \int_{h:\|h\|\le B_n} \dot{f}_{\mu_0}\left[Z_{n,\mu_0}+h\right] dP^*(h) \right\| \le 2B_n C_f \|\tilde{\pi}_n - P^*\|_{TV} \le \frac{2C_f}{B_n} \to 0.$$

Now, as $\|Z_{n,\mu_0}\| \le B_n$ holds on A_n, we have by the triangle and Chebyshev inequalities that

$$\begin{aligned}\left\| \int_{h:\|h\|>B_n} \dot{f}_{\mu_0}\left[Z_{n,\mu_0}+h\right] dP^*(h) \right\| &\le C_f \int_{h:\|h\|>B_n} \left(\|Z_{n,\mu_0}\| + \|h\|\right) dP^*(h) \\ &\le C_f \left(B_n P^*(\|Z^*\| > B_n) + \int_{h:\|h\|>B_n} \|h\| dP^*(h) \right) \\ &\le \frac{2C_f \mathrm{tr}(I_{\mu_0}^{-1})}{B_n} \to 0\end{aligned}$$

by Assumption 2(iii). Combining the preceding three displays, we see that

$$\left\| \mathcal{J}_{1,n} - \mathbb{E}^*\left[\dot{f}_{\mu_0}\left[Z^* + Z_{n,\mu_0}\right] \middle| Z_{n,\mu_0} \right] \right\| \to 0 \quad \text{on } A_n.$$

It follows that $\|\mathcal{J}_{1,n} - \mathbb{E}^*[\dot{f}_{\mu_0}[Z^* + Z_{n,\mu_0}]|Z_{n,\mu_0}]\| \overset{P_{n,\mu_0}}{\to} 0$ because $P_{n,\mu_0}(A_n^c) \to 0$.

It remains to show that $\|\mathcal{J}_{2,n}\| \overset{P_{n,\mu_0}}{\to} 0$. First note that by Assumptions 2(i)-(iii) we may invoke Lemma 10.3 of van der Vaart (1998), using Assumption 2(v) and Proposition 6.2 of Clarke and Barron (1990) to deduce existence of uniformly consistent tests as required in Theorem 10.1 of van der Vaart (1998). Hence, there exists a sequence of tests $\{\phi_n\}$ and a constant $c > 0$ such that for every sufficiently large n and every $\mu \in \mathcal{M}$ with $\|\mu - \mu_0\| \ge B_n/(2\sqrt{n})$, we have $\mathbb{E}_{n,\mu_0}[\phi_n(X^n)] \to 0$ and $\mathbb{E}_{n,\mu}[1-\phi_n(X^n)] \le e^{-cn(\|\mu-\mu_0\|^2 \wedge 1)}$. Without loss of generality we may take the tests to be nonrandomized (van der Vaart, 1998, p. 148). Then

$$\mathcal{J}_{2,n} = \phi_n(X^n)\mathcal{J}_{2,n} + (1-\phi_n(X^n))\mathcal{J}_{2,n},$$

where $P_{n,\mu_0}(\|\phi_n(X^n)\mathcal{J}_{2,n}\| > 0) \le P_{n,\mu_0}(\phi_n(X^n) = 1) \to 0$. The remaining term is

$$(1-\phi_n(X^n)) \frac{\int_{\mu:B_n<\|\sqrt{n}(\mu-\mu_0)-Z_{n,\mu_0}\|} \sqrt{n}\left(f(\mu)-f(\mu_0)\right) \prod_{i=1}^n p_\mu(X_i)/p_{\mu_0}(X_i) d\pi(\mu)}{\mathcal{J}_n},$$

where $\mathcal{J}_n = \int \prod_{i=1}^n p_\mu(X_i)/p_{\mu_0}(X_i) d\pi(\mu)$. By Lemma 6.26 of Ghosal and van der Vaart (2017) and Lemma 10, there exists $C > 0$ such that for any $C_n > 0$, the inequality

$$P_{n,\mu_0}\left(\mathcal{J}_n < Cn^{-K/2}e^{-2C_n}\right) \leq \frac{1}{C_n}$$

holds for all n sufficiently large. Below we shall specify a sequence $\{C_n\} \subset \mathbb{R}_+$ with $C_n \uparrow \infty$. Letting A_n' denote the event upon which $\|Z_{n,\mu_0}\| \leq \frac{1}{2}B_n$ and $\mathcal{J}_n \geq Cn^{-K/2}e^{-2C_n}$ both hold, we have by the above display and Assumption 2(iii) that $P_{n,\mu_0}(A_n'^c) \to 0$. It follows by the union bound, Markov's inequality, and Fubini's theorem that for any $\epsilon > 0$,

$$\begin{aligned} &P_{n,\mu_0}\left(\|(1-\phi_n(X^n))\mathcal{J}_{2,n}\| > \epsilon\right) \\ &\qquad \leq \frac{\int_{\mu:\|\sqrt{n}(\mu-\mu_0)\|>\frac{1}{2}B_n} \|\sqrt{n}\left(f(\mu)-f(\mu_0)\right)\| e^{-cn(\|\mu-\mu_0\|^2\wedge 1)} d\pi(\mu)}{\epsilon C n^{-K/2}e^{-2C_n}} + P_{n,\mu_0}(A_n'^c). \end{aligned}$$

Split the numerator into integrals over $N_{1,n} = \{\mu : \frac{1}{2}B_n < \|\sqrt{n}(\mu-\mu_0)\| \leq \epsilon'\sqrt{n}\}$ and $N_{2,n} = \{\mu : \|\mu-\mu_0\| > \epsilon'\}$ for $\epsilon' \in (0,1)$ sufficiently small that $\pi(\mu) \leq C_2$ on $\{\mu : \|\mu-\mu_0\| \leq \epsilon'\}$. Finally, Proposition 3.3 of Shapiro (1990) implies $\sup_{h:\|h\|\leq 1} \|f(\mu_0+th)-f(\mu_0)-\dot{f}_{\mu_0}[th]\| = o(t)$ as $t \downarrow 0$. Hence, we may choose $\bar{t} > 0$ such that $\|f(\mu_0+th)-f(\mu_0)-\dot{f}_{\mu_0}[th]\| \leq t$ holds for all $t \leq \bar{t}$ and all $\|h\| \leq 1$. Let $\epsilon' = \min\{\epsilon', \bar{t}\}$. It then follows for any $\mu \in N_{1,n}$ that, by setting $t = \|\mu-\mu_0\|$ and $h = \frac{\mu-\mu_0}{\|\mu-\mu_0\|}$, we have $\|f(\mu)-f(\mu_0)\| \leq \|\dot{f}_{\mu_0}[\mu-\mu_0]\| + \|\mu-\mu_0\|$. Hence, by (33), we have that $\|f(\mu)-f(\mu_0)\| \leq (1+C_f)\|\mu-\mu_0\|$ holds on $N_{1,n}$. Then

$$\begin{aligned} &\int_{N_{1,n}} \|\sqrt{n}\left(f(\mu)-f(\mu_0)\right)\| e^{-cn(\|\mu-\mu_0\|^2\wedge 1)} d\pi(\mu) \\ &\qquad \leq C_2(1+C_f)\int_{N_{1,n}} \|\sqrt{n}(\mu-\mu_0)\| e^{-c\|\sqrt{n}(\mu-\mu_0)\|^2} d\mu \\ &\qquad \leq C_2(1+C_f)n^{-K/2}\int_{h:\|h\|>\frac{1}{2}B_n} \|h\| e^{-c\|h\|^2} dh, \end{aligned}$$

where the final line is by a change of variables from μ to $h = \sqrt{n}(\mu-\mu_0)$. The final integral vanishes as $n \to \infty$ by dominated convergence. Further, by Assumption 4(i), we have

$$\int_{N_{2,n}} \|\sqrt{n}\left(f(\mu)-f(\mu_0)\right)\| e^{-cn(\|\mu-\mu_0\|^2\wedge 1)} d\pi(\mu) \leq \sqrt{n}e^{-nc(\epsilon')^2} \int \|f(\mu)-f(\mu_0)\| d\pi(\mu).$$

Choosing $\{C_n\} \subset \mathbb{R}_+$ with $C_n \uparrow \infty$ sufficiently slowly that

$$e^{2C_n}\int_{h:\|h\|>\frac{1}{2}B_n} \|h\| e^{-c\|h\|^2} dh \to 0, \qquad \text{and} \qquad n^{(K+1)/2}e^{2C_n - nc(\epsilon')^2} \to 0,$$

it follows from the preceding four displays that $P_{n,\mu_0}\left(\|(1-\phi_n(X^n))\mathcal{J}_{2,n}\| > \epsilon\right) \to 0$. ■

For the next result, recall that $Z \sim N(h, I_{\mu_0}^{-1})$, $\mathbb{E}^*$ denotes expectation with respect to $Z^* \sim N(0, I_{\mu_0}^{-1})$ independent of Z, and $\overset{P_{n,h}}{\rightsquigarrow}$ denotes convergence in distribution along $\{P_{n,h}\}$.

Proposition 2 *Suppose that Assumptions 2 and 4 hold and $\pi \in \Pi$. Then for every $h \in \mathbb{R}^K$,*

$$\bar{f}_n \overset{P_{n,h}}{\rightsquigarrow} \mathbb{E}^*\left[\dot{f}_{\mu_0}\left[Z^* + Z\right]\middle| Z\right].$$

Proof of Proposition 2. Assumption 2(i)-(iii) imply $\mathcal{P} = \{P_\mu : \mu \in \mathcal{M}\}$ is locally asymptotically normal (van der Vaart, 1998, Example 7.15) and so $\{P_{n,\mu_0}\}$ and $\{P_{n,h}\}$ are mutually contiguous (van der Vaart, 1998, Example 6.5). Hence, with $\overset{P_{n,h}}{\to}$ denoting convergence in probability under $\{P_{n,h}\}$, we have by Lemma 4 that

$$\left\|\bar{f}_n - \mathbb{E}^*\left[\dot{f}_{\mu_0}\left[Z^* + Z_{n,\mu_0}\right]\middle| Z_{n,\mu_0}\right]\right\| \overset{P_{n,h}}{\to} 0.$$

Moreover, under Assumption 2(i)-(iii) we have by Le Cam's third lemma (van der Vaart, 1998, Example 6.7) that

$$Z_{n,\mu_0} \overset{P_{n,h}}{\rightsquigarrow} Z \sim N(h, I_{\mu_0}^{-1}).$$

The result follows by the continuous mapping theorem, noting $z \mapsto \mathbb{E}^*[\dot{f}_{\mu_0}[Z^* + z]]$ is everywhere continuous by Assumption 4(ii) and Proposition 3.1 of Shapiro (1990). ■

C.2 Semiparametric Models

We now extend Proposition 2 to semiparametric models. In this case,

$$\bar{f}_n = \int \sqrt{n}(f(\mu) - f(\mu_0))d\pi_n(\mu)$$

with π_n denoting the quasi-posterior in (25) based on a Gaussian quasi-likelihood. Before giving the main result we first state and prove several lemmas. The first is a Bernstein–von Mises result for the quasi-posterior. Let $Z_{n,(\mu_0,\eta_0)} = \sqrt{n}(\hat{\mu} - \mu_0)$. Let $\tilde{\pi}_n(h) \propto e^{-\frac{1}{2}h^T\hat{I}h}\pi(\mu_0 + n^{-1/2}(h + Z_{n,(\mu_0,\eta_0)})$ denote the quasi-posterior for h under a change of variables from μ to $h = \sqrt{n}(\mu - \mu_0) - Z_{n,(\mu_0,\eta_0)}$. Let P^* denote the distribution of $Z^* \sim N(0, I_{(\mu_0,\eta_0)}^{-1})$.

Lemma 5 *Suppose that Assumption 3(i),(iii),(v) holds and $\pi \in \Pi$. Then*

$$\|\tilde{\pi}_n - P^*\|_{TV} \overset{P_{n,(\mu_0,\eta_0)}}{\to} 0.$$

Proof of Lemma 5. Fix any Borel set $B \subset \mathbb{R}^K$. We have

$$\tilde{\pi}_n(B) = \frac{\int_{h:\mu_0+n^{-1/2}(h+Z_{n,(\mu_0,\eta_0)})\in\mathcal{M}} \mathbb{I}[h \in B] e^{-\frac{1}{2}h^T\hat{I}h}\pi(\mu_0 + n^{-1/2}(h + Z_{n,(\mu_0,\eta_0)})dh}{\int_{h:\mu_0+n^{-1/2}(h+Z_{n,(\mu_0,\eta_0)})\in\mathcal{M}} e^{-\frac{1}{2}h^T\hat{I}h}\pi(\mu_0 + n^{-1/2}(h + Z_{n,(\mu_0,\eta_0)})dh}.$$

Let $\{M_n\} \subset \mathbb{R}_+$ be a sequence with $M_n \uparrow +\infty$, $M_n = o(n^{1/2})$. Let A_n denote the event upon which $\|Z_{n,(\mu_0,\eta_0)}\| \leq M_n$ and $c \leq \hat{\lambda}_{\min}, \hat{\lambda}_{\max} \leq c^{-1}$ hold, where $\hat{\lambda}_{\min}$ and $\hat{\lambda}_{\max}$ are the smallest and largest eigenvalues of $\hat{I}$, and $c \in (0,1)$ is chosen such that the minimum and maximum eigenvalues of $I_{(\mu_0,\eta_0)}$ are contained in (c, c^{-1}) (such a c exists by Assumption 3(iii)). We have $P_{n,(\mu_0,\eta_0)}(A_n) \to 1$ by Assumptions 3(iii) and 3(v). Moreover, by Assumption 3(i),

$$\{\mu_0 + n^{-1/2}(h + Z_{n,(\mu_0,\eta_0)}) : \|h\| \leq M_n\} \subset \{\mu_0 + n^{-1/2}h : \|h\| \leq 2M_n\} \subset \mathcal{M}$$

holds on A_n for all n sufficiently large. It follows that for n sufficiently large, we have

$$\sup_{\|h\|\leq M_n} |\pi(\mu_0 + n^{-1/2}(h + Z_{n,(\mu_0,\eta_0)})) - \pi(\mu_0)| \leq \sup_{\|h\|\leq 2M_n} |\pi(\mu_0 + n^{-1/2}h) - \pi(\mu_0)| =: \varepsilon_n \pi(\mu_0)$$

on A_n, where $\varepsilon_n \downarrow 0$ because $\pi \in \Pi$. The density $\pi(\mu)$ is uniformly bounded above by some $\overline{\pi} < +\infty$ because $\pi \in \Pi$. Hence, on A_n, for all n sufficiently large (independent of B),

$$\begin{aligned}
\tilde{\pi}_n(B) &\leq \frac{\int_{h:\mu_0+n^{-1/2}(h+Z_{n,(\mu_0,\eta_0)})\in\mathcal{M}} \mathbb{I}[h \in B] e^{-\frac{1}{2}h^T\hat{I}h}\pi(\mu_0 + n^{-1/2}(h + Z_{n,(\mu_0,\eta_0)})dh}{(1-\varepsilon_n)\pi(\mu_0)\int_{h:\|h\|\leq M_n} e^{-\frac{1}{2}h^T\hat{I}h}dh} \\
&\leq \frac{(1+\varepsilon_n)\int_{h:\|h\|\leq M_n} \mathbb{I}[h \in B] e^{-\frac{1}{2}h^T\hat{I}h}\, dh}{(1-\varepsilon_n)\int_{h:\|h\|\leq M_n} e^{-\frac{1}{2}h^T\hat{I}h}dh} + \frac{\overline{\pi}\int_{h:\|h\|>M_n} e^{-\frac{1}{2}h^T\hat{I}h}dh}{(1-\varepsilon_n)\pi(\mu_0)\int_{h:\|h\|\leq M_n} e^{-\frac{1}{2}h^T\hat{I}h}dh} \\
&\leq \frac{(1+\varepsilon_n)\int \mathbb{I}[h \in B] e^{-\frac{1}{2}h^T\hat{I}h}\, dh}{(1-\varepsilon_n)\int_{h:\|h\|\leq M_n} e^{-\frac{1}{2}h^T\hat{I}h}dh} + \frac{\overline{\pi}\int_{h:\|h\|>M_n} e^{-\frac{c}{2}\|h\|^2}dh}{(1-\varepsilon_n)\pi(\mu_0)\int_{h:\|h\|\leq M_n} e^{-\frac{1}{2c}\|h\|^2}dh},
\end{aligned}$$

where the second term on the right-hand side, say $t_{1,n}$, converges to zero by dominated convergence. Moreover,

$$\delta_{1,n} := \frac{\int e^{-\frac{1}{2}h^T\hat{I}h}dh}{\int_{h:\|h\|\leq M_n} e^{-\frac{1}{2}h^T\hat{I}h}dh} - 1 = \frac{\int_{h:\|h\|>M_n} e^{-\frac{1}{2}h^T\hat{I}h}dh}{\int_{h:\|h\|\leq M_n} e^{-\frac{1}{2}h^T\hat{I}h}dh} \leq \frac{\int_{h:\|h\|>M_n} e^{-\frac{c}{2}\|h\|^2}dh}{\int_{h:\|h\|\leq M_n} e^{-\frac{1}{2c}\|h\|^2}dh} \to 0,$$

on A_n, again by dominated convergence. Hence,

$$\tilde{\pi}_n(B) \leq \frac{(1+\varepsilon_n)}{(1-\varepsilon_n)} \frac{\int e^{-\frac{1}{2}h^T\hat{I}h}dh}{\int_{h:\|h\|\leq M_n} e^{-\frac{1}{2}h^T\hat{I}h}dh} \frac{\int \mathbb{I}[h\in B]e^{-\frac{1}{2}h^T\hat{I}h}\,dh}{\int e^{-\frac{1}{2}h^T\hat{I}h}dh} + t_{1,n}$$
$$\leq \frac{(1+\varepsilon_n)(1+\delta_{1,n})}{1-\varepsilon_n}\hat{P}(B) + t_{1,n}, \quad (34)$$

holds on A_n for n sufficiently large (independent of B), where $\hat{P}$ denotes the distribution of $Z^* \sim N(0, \hat{I}^{-1})$, and ε_n, $\delta_{1,n}$ and $t_{1,n}$ are all independent of B and converge (either deterministically or in probability) to zero.

Similarly, for all n sufficiently large (independently of B), on A_n, we have

$$\begin{aligned}
\tilde{\pi}_n(B) &\geq \frac{\int_{h:\|h\|\leq M_n} \mathbb{I}[h\in B]e^{-\frac{1}{2}h^T\hat{I}h}\pi(\mu_0+n^{-1/2}(h+Z_{n,(\mu_0,\eta_0)}))dh}{\int_{h:\mu_0+n^{-1/2}(h+Z_{n,(\mu_0,\eta_0)})\in\mathcal{M}} e^{-\frac{1}{2}h^T\hat{I}h}\pi(\mu_0+n^{-1/2}(h+Z_{n,(\mu_0,\eta_0)}))dh} \\
&\geq \frac{(1-\varepsilon_n)\pi(\mu_0)\int_{h:\|h\|\leq M_n} \mathbb{I}[h\in B]e^{-\frac{1}{2}h^T\hat{I}h}dh}{\int_{h:\mu_0+n^{-1/2}(h+Z_{n,(\mu_0,\eta_0)})\in\mathcal{M}} e^{-\frac{1}{2}h^T\hat{I}h}\pi(\mu_0+n^{-1/2}(h+Z_{n,(\mu_0,\eta_0)}))dh} \\
&\geq \frac{(1-\varepsilon_n)\pi(\mu_0)\int_{h:\|h\|\leq M_n} \mathbb{I}[h\in B]e^{-\frac{1}{2}h^T\hat{I}h}dh}{(1+\delta_{2,n})\int_{h:\|h\|\leq M_n} e^{-\frac{1}{2}h^T\hat{I}h}\pi(\mu_0+n^{-1/2}(h+Z_{n,(\mu_0,\eta_0)}))dh},
\end{aligned}$$

where

$$\begin{aligned}
\delta_{2,n} &:= \left| \frac{\int_{h:\mu_0+n^{-1/2}(h+Z_{n,(\mu_0,\eta_0)})\in\mathcal{M}} e^{-\frac{1}{2}h^T\hat{I}h}\pi(\mu_0+n^{-1/2}(h+Z_{n,(\mu_0,\eta_0)}))dh}{\int_{h:\|h\|\leq M_n} e^{-\frac{1}{2}h^T\hat{I}h}\pi(\mu_0+n^{-1/2}(h+Z_{n,(\mu_0,\eta_0)}))dh} - 1 \right| \\
&\leq \frac{\overline{\pi}\int_{h:\|h\|>M_n} e^{-\frac{1}{2}h^T\hat{I}h}dh}{(1-\varepsilon_n)\pi(\mu_0)\int_{h:\|h\|\leq M_n} e^{-\frac{1}{2}h^T\hat{I}h}dh} \leq \frac{\overline{\pi}\int_{h:\|h\|>M_n} e^{-\frac{c}{2}\|h\|^2}dh}{(1-\varepsilon_n)\pi(\mu_0)\int_{h:\|h\|\leq M_n} e^{-\frac{1}{2c}\|h\|^2}dh} \to 0.
\end{aligned}$$

Hence, on A_n, for all n sufficiently large (independent of B),

$$\begin{aligned}
\tilde{\pi}_n(B) &\geq \frac{(1-\varepsilon_n)}{(1+\delta_{2,n})(1+\varepsilon_n)} \frac{\int e^{-\frac{1}{2}h^T\hat{I}h}dh}{\int_{h:\|h\|\leq M_n} e^{-\frac{1}{2}h^T\hat{I}h}dh} \frac{\int_{h:\|h\|\leq M_n} \mathbb{I}[h\in B]e^{-\frac{1}{2}h^T\hat{I}h}dh}{\int e^{-\frac{1}{2}h^T\hat{I}h}dh} \\
&\geq \frac{(1-\varepsilon_n)(1-\delta_{1,n})}{(1+\delta_{2,n})(1+\varepsilon_n)} \frac{\int_{h:\|h\|\leq M_n} \mathbb{I}[h\in B]e^{-\frac{1}{2}h^T\hat{I}h}dh}{\int e^{-\frac{1}{2}h^T\hat{I}h}dh} \\
&\geq \frac{(1-\varepsilon_n)(1-\delta_{1,n})}{(1+\delta_{2,n})(1+\varepsilon_n)}\hat{P}(B) - \frac{(1-\varepsilon_n)(1-\delta_{1,n})}{(1+\delta_{2,n})(1+\varepsilon_n)} \frac{\int_{h:\|h\|>M_n} e^{-\frac{c}{2}\|h\|^2}dh}{\int e^{-\frac{1}{2c}\|h\|^2}dh},
\end{aligned}$$

where the second term on the right-hand side, say $t_{2,n}$, converges to zero by dominated convergence. Hence,

$$\tilde{\pi}_n(B) \geq \frac{(1-\varepsilon_n)(1-\delta_{1,n})}{(1+\delta_{2,n})(1+\varepsilon_n)}\hat{P}(B) - t_{2,n} \tag{35}$$

holds on A_n for n sufficiently large (independent of B), where ε_n, $\delta_{1,n}$, $\delta_{2,n}$, and $t_{2,n}$ are all independent of B and converge (either deterministically or in probability) to zero.

Combining (34) and (35) and using the fact that $P_{n,(\mu_0,\eta_0)}(A_n) \to 1$, we may deduce

$$\|\tilde{\pi}_n - \hat{P}\|_{TV} \equiv \sup_B \left|\tilde{\pi}_n(B) - \hat{P}(B)\right| \overset{P_{n,(\mu_0,\eta_0)}}{\to} 0,$$

where the supremum is over all Borel sets. Finally, it is easy to deduce (e.g., by using Pinsker's inequality) that the total variation distance between the $N(0,\Sigma_1)$ and $N(0,\Sigma_2)$ distributions converges to zero as $\Sigma_1 \to \Sigma_2$ for Σ_2 positive definite. It therefore follows by Assumptions 3(iii) and 3(v)b that $\|\hat{P} - P^*\|_{TV} \overset{P_{n,(\mu_0,\eta_0)}}{\to} 0$. ■

Lemma 6 *Suppose that Assumptions 3(i),(iii),(v) and 4(ii) hold and $\pi \in \Pi$. Then there exists a sequence $\{B_n\} \subset \mathbb{R}_+$ with $B_n \uparrow +\infty$, $B_n = o(\sqrt{n})$, such that for all n sufficiently large,*

$$\sup_{h:\|h\|\leq 2B_n} \left\| \sqrt{n}\left(f(\mu_0 + h/\sqrt{n}) - f(\mu_0)\right) - \dot{f}_{\mu_0}[h] \right\| \leq \frac{1}{B_n},$$

and

$$P_{n,(\mu_0,\eta_0)}\left(\|\tilde{\pi}_n - P^*\|_{TV} > \frac{1}{B_n^2}\right) \leq \frac{1}{B_n}.$$

Proof of Lemma 6. The result follows by similar arguments to the proof of Lemma 3, using Lemma 5 in place of the (parametric) Bernstein–von Mises theorem. ■

For the next result we invoke Assumption 3(v)a, but we only require $\|\sqrt{n}(\hat{\mu} - \mu_0)\|$ to be bounded in probability under $\{P_{n,(\mu_0,\eta_0)}\}$. Let $\mathbb{E}^*$ denote expectation with respect to $Z^* \sim N(0, I^{-1}_{(\mu_0,\eta_0)})$ independent of $Z_{n,(\mu_0,\eta_0)}$.

Lemma 7 *Suppose Assumptions 3(i),(iii),(v) and 4 hold and $\pi \in \Pi$. Then*

$$\left\| \bar{f}_n - \mathbb{E}^*\left[\dot{f}_{\mu_0}\left[Z^* + Z_{n,(\mu_0,\eta_0)}\right] \middle| Z_{n,(\mu_0,\eta_0)}\right] \right\| \overset{P_{n,(\mu_0,\eta_0)}}{\to} 0$$

as $n \to \infty$.

Proof of Lemma 7. Let $\{B_n\}$ be as in Lemma 6. By a change of variables, we have

$$\bar{f}_n = \int_{h:\|h\|\le B_n} \sqrt{n}\left(f(\mu_0 + n^{-1/2}(Z_{n,(\mu_0,\eta_0)} + h)) - f(\mu_0)\right) d\tilde{\pi}_n(h) + \int_{\mu:\|\sqrt{n}(\mu-\hat{\mu})\|>B_n} \sqrt{n}\left(f(\mu) - f(\mu_0)\right) d\pi_n(\mu) =: \mathcal{J}_{1,n} + \mathcal{J}_{2,n}.$$

Let A_n denote the event $\|Z_{n,(\mu_0,\eta_0)}\| \le B_n$, $\|\tilde{\pi}_n - P^*\|_{TV} \le B_n^{-2}$, and $c \le \hat{\lambda}_{\min}, \hat{\lambda}_{\max} \le c^{-1}$ for some $c \in (0,1)$. Assumptions 3(iii) and 3(v) and Lemma 6 imply $\lim_{n\to\infty} P_{n,(\mu_0,\eta_0)}(A_n^c) = 0$. Again by Lemma 6, on A_n we have

$$\left\| \mathcal{J}_{1,n} - \int_{h:\|h\|\le B_n} \dot{f}_{\mu_0}\left[Z_{n,(\mu_0,\eta_0)} + h\right] d\tilde{\pi}_n(h) \right\| \le \frac{1}{B_n} \to 0,$$

where, as in the proof of Lemma 4, we may deduce that

$$\left\| \int_{h:\|h\|\le B_n} \dot{f}_{\mu_0}\left[Z_{n,(\mu_0,\eta_0)} + h\right] d\tilde{\pi}_n(h) - \int_{h:\|h\|\le B_n} \dot{f}_{\mu_0}\left[Z_{n,(\mu_0,\eta_0)} + h\right] dP^*(h) \right\| \le \frac{2C_f}{B_n} \to 0,$$

on A_n, and

$$\left\| \int_{h:\|h\|> B_n} \dot{f}_{\mu_0}\left[Z_{n,(\mu_0,\eta_0)} + h\right] dP^*(h) \right\| \le \frac{2C_f \mathrm{tr}(I^{-1}_{(\mu_0,\eta_0)})}{B_n} \to 0,$$

on A_n, where the latter convergence holds in view of Assumption 3(iii). It follows that

$$\|\mathcal{J}_{1,n} - \mathbb{E}^*[\dot{f}_{\mu_0}[Z^* + Z_{n,(\mu_0,\eta_0)}]|Z_{n,(\mu_0,\eta_0)}]\| \overset{P_{n,(\mu_0,\eta_0)}}{\to} 0.$$

It remains to show that $\|\mathcal{J}_{2,n}\| \overset{P_{n,(\mu_0,\eta_0)}}{\to} 0$. To this end, write

$$\mathcal{J}_{2,n} = \int_{\mu:B_n<\|\sqrt{n}(\mu-\hat{\mu})\|\le\epsilon\sqrt{n}} \sqrt{n}\left(f(\mu) - f(\mu_0)\right) d\pi_n(\mu) + \int_{\mu:\|\sqrt{n}(\mu-\hat{\mu})\|>\epsilon\sqrt{n}} \sqrt{n}\left(f(\mu) - f(\mu_0)\right) d\pi_n(\mu) =: \mathcal{J}_{2,n,a} + \mathcal{J}_{2,n,b},$$

where, as in the proof of Lemma 4, we choose $\epsilon > 0$ sufficiently small that $\|f(\mu) - f(\mu_0)\| \le (1 + C_f)\|\mu - \mu_0\|$ holds for μ with $\|\mu - \mu_0\| \le 2\epsilon$. Then on A_n, we have

$$\|\mu - \mu_0\| \le \|\mu - \hat{\mu}\| + \|\hat{\mu} - \mu_0\| \le \|\mu - \hat{\mu}\| + \frac{B_n}{\sqrt{n}},$$

so $\{\mu : \|\mu - \hat{\mu}\| \leq \epsilon\} \subseteq \{\mu : \|\mu - \mu_0\| \leq 2\epsilon\}$ for all n sufficiently large, in which case

$$\|\mathcal{I}_{2,n,a}\| \leq (1+C_f)\left(\int_{\mu: B_n < \|\sqrt{n}(\mu-\hat{\mu})\| \leq \epsilon\sqrt{n}} \|\sqrt{n}(\mu-\hat{\mu})\| d\pi_n(\mu) + B_n \pi_n(\{\mu : B_n < \sqrt{n}(\mu-\hat{\mu})\}) \right).$$

Taking ϵ smaller if necessary, we have $\{\mu : \|\mu - \mu_0\| \leq 2\epsilon\} \subset \mathcal{M}$ in view of Assumption 3(i). Then by dominated convergence and the fact that $c \leq \hat{\lambda}_{\min}$, $\hat{\lambda}_{\max} \leq c^{-1}$ holds on A_n, we have

$$\int_{\mu: B_n < \|\sqrt{n}(\mu-\hat{\mu})\| \leq \epsilon\sqrt{n}} \|\sqrt{n}(\mu-\hat{\mu})\| d\pi_n(\mu) \leq \frac{\int_{h:\|h\|>B_n} \|h\| e^{-\frac{c}{2}\|h\|^2} \pi(\mu_0 + n^{-1/2}(h + Z_{n,(\mu_0,\eta_0)}))dh}{\int e^{-\frac{1}{2c}\|h\|^2} \pi(\mu_0 + n^{-1/2}(h + Z_{n,(\mu_0,\eta_0)}))dh}$$

and

$$B_n \pi_n(\{\mu : B_n < \sqrt{n}(\mu - \hat{\mu})\}) \leq B_n \frac{\int_{h:\|h\|>B_n} e^{-\frac{c}{2}\|h\|^2} \pi(\mu_0 + n^{-1/2}(h + Z_{n,(\mu_0,\eta_0)}))dh}{\int e^{-\frac{1}{2c}\|h\|^2} \pi(\mu_0 + n^{-1/2}(h + Z_{n,(\mu_0,\eta_0)}))dh}.$$

It then follows by similar arguments to the proof of Lemma 5 that $\|\mathcal{J}_{2,n,a}\| \overset{P_{n,(\mu_0,\eta_0)}}{\to} 0$.

Finally, for $\mathcal{J}_{2,n,b}$, on A_n we have

$$\|\mathcal{J}_{2,n,b}\| \leq \frac{\sqrt{n} e^{-\frac{cn\epsilon^2}{2}} \int f(\mu) - f(\mu_0)\pi(\mu)d\mu}{n^{-K/2} \int e^{-\frac{1}{2c}\|h\|^2} \pi(\mu_0 + n^{-1/2}(h + Z_{n,(\mu_0,\eta_0)}))dh}.$$

The integral in the numerator is finite by Assumption 4(i) and the integral in the denominator may be shown to be bounded below by a positive constant on A_n by similar arguments to the proof of Lemma 5, from which it follows that $\|\mathcal{J}_{2,n,b}\| \overset{P_{n,(\mu_0,\eta_0)}}{\to} 0$. ■

For the next result, let $Z \sim N(h, I_{\mu_0}^{-1})$ and let $\mathbb{E}^*$ denote expectation with respect to $Z^* \sim N(0, I_{\mu_0}^{-1})$ independent of Z.

Proposition 3 *Suppose that Assumptions 3 and 4 hold and $\pi \in \Pi$. Then for every $h \in \mathbb{R}^K$,*

$$\bar{f}_n \overset{P_{n,h}}{\rightsquigarrow} \mathbb{E}^* \left[\dot{f}_{\mu_0} [Z^* + Z] \middle| Z \right].$$

Proof of Proposition 3. Assumption 3(i)-(iii) imply $\mathcal{P} = \{P_{\beta(t)} : t \in \mathcal{M}_{(\mu_0,\eta_0)}\}$ is locally asymptotically normal (van der Vaart, 1998, Example 7.15) and so $\{P_{n,(\mu_0,\eta_0)}\}$ and $\{P_{n,h}\}$ are mutually contiguous (van der Vaart, 1998, Example 6.5). Hence, with $\overset{P_{n,h}}{\to}$ denoting convergence in probability under $\{P_{n,h}\}$, we have by Lemma 7 that

$$\left\| \bar{f}_n - \mathbb{E}^* \left[\dot{f}_{\mu_0} \left[Z^* + Z_{n,(\mu_0,\eta_0)} \right] \middle| Z_{n,(\mu_0,\eta_0)} \right] \right\| \overset{P_{n,h}}{\to} 0.$$

Moreover, under Assumption 3(v)a we have

$$Z_{n,(\mu_0,\eta_0)} \overset{P_{n,h}}{\rightsquigarrow} Z \sim N(h, I^{-1}_{(\mu_0,\eta_0)}).$$

The result follows by the continuous mapping theorem, noting $z \mapsto \mathbb{E}^*[\dot{f}_{\mu_0}[Z^* + z]]$ is everywhere continuous by Assumption 4(ii) and Proposition 3.1 of Shapiro (1990). ■

D Bounds on Mistake Probabilities

Definition 1(ii) requires mistake probabilities to vanish faster than $n^{-1/2}$. The main results of this section are Lemma 11, which shows that mistake probabilities for Bayes decisions vanish at rate n^{-1}, and Lemma 12, which presents a result for quasi-Bayes decisions.

D.1 Parametric Models

We first present some preliminary lemmas before proving Lemma 11.

Lemma 8 *Suppose that Assumption 1(i) holds. Then for any $\mu_0 \in \mathcal{M}$ there exists $\epsilon, \epsilon' > 0$ such that*

$$\pi_n(\{\mu : \|\mu - \mu_0\| \geq \epsilon\}) \leq 2\epsilon' \tag{36}$$

implies $\delta_n^*(X^n; \pi) \in \underline{\mathcal{D}}_{\mu_0}$.

Proof of Lemma 8. If $\underline{\mathcal{D}}_{\mu_0} = \mathcal{D}$ then the result is trivial. Now suppose $\underline{\mathcal{D}}_{\mu_0} \subsetneq \mathcal{D}$. By Assumption 1(i), we may choose $\epsilon, \varepsilon > 0$ such that

$$\min_{i \notin \underline{\mathcal{D}}_{\mu_0}} R(i,\mu) - \max_{i \in \underline{\mathcal{D}}_{\mu_0}} R(i,\mu) \geq \varepsilon, \qquad \mu \in N := \{\mu : \|\mu - \mu_0\| < \epsilon\}.$$

We have

$$\begin{aligned}\max_{i \in \underline{\mathcal{D}}_{\mu_0}} \bar{R}_n(i) &\leq \max_{i \in \underline{\mathcal{D}}_{\mu_0}} \int_N R(i,\mu)\, d\pi_n(\mu) + \max_{i \in \underline{\mathcal{D}}_{\mu_0}} \int_{N^c} R(i,\mu)\, d\pi_n(\mu) \\ &\leq \int_N \max_{i \in \underline{\mathcal{D}}_{\mu_0}} R(i,\mu)\, d\pi_n(\mu) + \pi_n(N^c)\|R\|_\infty,\end{aligned}$$

where $\|R\|_\infty := \sup_{(d,\mu) \in \mathcal{D} \times \mathcal{M}} |R(d,\mu)| < \infty$ by Assumption 1(i), and

$$\begin{aligned}\min_{i \notin \underline{\mathcal{D}}_{\mu_0}} \bar{R}_n(i) &\geq \min_{i \notin \underline{\mathcal{D}}_{\mu_0}} \int_N R(i,\mu)\, d\pi_n(\mu) + \min_{i \notin \underline{\mathcal{D}}_{\mu_0}} \int_{N^c} R(i,\mu)\, d\pi_n(\mu) \\ &\geq \int_N \min_{i \notin \underline{\mathcal{D}}_{\mu_0}} R(i,\mu)\, d\pi_n(\mu) - \pi_n(N^c)\|R\|_\infty.\end{aligned}$$

Hence,

$$\begin{aligned}\min_{i\notin\underline{\mathcal{D}}_{\mu_0}}\bar{R}_n(i)-\max_{i\in\underline{\mathcal{D}}_{\mu_0}}\bar{R}_n(i) &\geq \int_N\left(\min_{i\notin\underline{\mathcal{D}}_{\mu_0}}R(i,\mu)-\max_{i\in\underline{\mathcal{D}}_{\mu_0}}R(i,\mu)\right)d\pi_n(\mu)-2\pi_n(N^c)\|R\|_\infty\\ &\geq \pi_n(N)\varepsilon-2\pi_n(N^c)\|R\|_\infty\\ &= \varepsilon-\pi_n(N^c)(\varepsilon+2\|R\|_\infty).\end{aligned}$$

If $\pi_n(N^c)\leq\frac{\varepsilon}{2(\varepsilon+2\|R\|_\infty)}=:2\epsilon'$ holds, then

$$\min_{i\notin\underline{\mathcal{D}}_{\mu_0}}\bar{R}_n(i)-\max_{i\in\underline{\mathcal{D}}_{\mu_0}}\bar{R}_n(i)\geq\varepsilon-\frac{\varepsilon}{2}=\frac{\varepsilon}{2},$$

which implies

$$\max_{i\in\underline{\mathcal{D}}_{\mu_0}}\bar{R}_n(i)\leq\min_{i\notin\underline{\mathcal{D}}_{\mu_0}}\bar{R}_n(i)-\frac{\varepsilon}{2}<\min_{i\notin\underline{\mathcal{D}}_{\mu_0}}\bar{R}_n(i),$$

and hence that $\delta_n^*(X^n;\pi)\in\underline{\mathcal{D}}_{\mu_0}$. ■

Lemma 9 *Suppose that Assumption 1(i) holds. Take any $\mu_0\in\mathcal{M}$ and $h\in\mathbb{R}^K$, and let ϵ,ϵ' be as in (36) of Lemma 8. Suppose that there exist tests $\{\phi_n\}$ with the property that $\mathbb{E}_{n,h}[\phi_n(X^n)]\leq C'e^{-nC}$ and $\sup_{\mu:\|\mu-\mu_0\|\geq\epsilon}\mathbb{E}_{n,\mu}[1-\phi_n(X^n)]\leq C'e^{-nC}$ for $C,C'>0$. Then*

$$P_{n,h}\left(\delta_n^*(X^n;\pi)\notin\underline{\mathcal{D}}_{\mu_0}\right)\leq\frac{C'e^{-nC}}{\epsilon'}+\frac{C'e^{-nC/2}}{\epsilon'\pi(\{\mu:D_{KL}(p_{\mu_0+h/\sqrt{n}}\|p_\mu)<n^{-1}\})}+\frac{4}{nC}.$$

Proof of Lemma 9. In view of Lemma 8, we have

$$P_{n,h}\left(\delta_n^*(X^n;\pi)\notin\underline{\mathcal{D}}_{\mu_0}\right)\leq P_{n,h}\left(\pi_n(\{\mu:\|\mu-\mu_0\|\geq\epsilon\})>2\epsilon'\right).$$

We proceed by similar arguments to the consistency theorem of Schwartz (1965) but centering along $P_{n,h}$. Write

$$\begin{aligned}P_{n,h}\left(\pi_n(\{\mu:\|\mu-\mu_0\|\geq\epsilon\})>2\epsilon'\right)&\leq P_{n,h}\left(\phi_n(X^n)>\epsilon'\right)\\ &\quad+P_{n,h}\left(\frac{(1-\phi_n(X^n))\int_{N^c}\prod_{i=1}^n p_\mu(X_i)/p_{\mu_0+h/\sqrt{n}}(X_i)\,d\pi(\mu)}{\mathcal{J}_n}>\epsilon'\right),\end{aligned}$$

where $\mathcal{J}_n=\int\prod_{i=1}^n p_\mu(X_i)/p_{\mu_0+h/\sqrt{n}}(X_i)\,d\pi(\mu)$ and $N^c=\{\mu:\|\mu-\mu_0\|>\epsilon\}$. The first probability on the right-hand side is bounded by $C'e^{-nC}/\epsilon'$ by Markov's inequality. For the second, Lemma 6.26 of Ghosal and van der Vaart (2017) implies that for any $B>0$ and

$\varepsilon \geq n^{-1}$,[17]

$$P_{n,h}\left(\mathcal{J}_n < \pi(\{\mu : D_{KL}(p_{\mu_0+h/\sqrt{n}}\|p_\mu) < \varepsilon\})e^{-2nB\varepsilon}\right) \leq \frac{1}{B}.$$

Hence we may deduce

$$P_{n,h}\left(\frac{(1-\phi_n'(X^n))\int_{N^c}\prod_{i=1}^n p_\mu(X_i)/p_{\mu_0+h/\sqrt{n}}(X_i)\,d\pi(\mu)}{\mathcal{J}_n} > \epsilon'\right)$$
$$\leq \frac{\mathbb{E}_{n,h}\left[(1-\phi_n'(X^n))\int_{N^c}\prod_{i=1}^n p_\mu(X_i)/p_{\mu_0+h/\sqrt{n}}(X_i)\,d\pi(\mu)\right]}{\epsilon'\pi(\{\mu : D_{KL}(p_{\mu_0+h/\sqrt{n}}\|p_\mu) < \varepsilon\})e^{-2nB\varepsilon}} + \frac{1}{B}.$$

The numerator in the first term on the right-hand side may be bounded by $C'e^{-Cn}$ using Fubini's theorem (see, e.g., the proof of Lemma 6.17 of Ghosal and van der Vaart (2017)). Setting $B = Cn/4$ and $\varepsilon = n^{-1}$, the right-hand side of the above display is bounded by

$$\frac{C'e^{-nC/2}}{\epsilon'\pi(\{\mu : D_{KL}(p_{\mu_0+h/\sqrt{n}}\|p_\mu) < n^{-1}\})} + \frac{4}{nC},$$

as required. ■

Lemma 10 *Suppose that Assumption 2(i),(iii)-(iv) hold and $\pi \in \Pi$. Then for $\mu_0 \in \mathcal{M}$ and $h \in \mathbb{R}^K$,*

$$\liminf_{n\to\infty} n^{K/2}\pi(\{\mu : D_{KL}(p_{\mu_0+h/\sqrt{n}}\|p_\mu) < n^{-1}\}) > 0.$$

Proof of Lemma 10. Assumption 2(iv) implies $D_{KL}(p_{\mu_0+h/\sqrt{n}}\|p_\mu) \leq B\|\mu-\mu_0-h/\sqrt{n}\|^2$ on a neighborhood of μ_0 for all sufficiently large n, with $B = 2\lambda_{\max}(I_{\mu_0}) < \infty$ by Assumption 2(iii). Hence, for all n sufficiently large, $\|\mu-\mu_0-h/\sqrt{n}\|^2 < (Bn)^{-1}$ implies $D_{KL}(p_{\mu_0+h/\sqrt{n}}\|p_\mu) < n^{-1}$. Assumption 2(i) and the positivity and continuity of π imply that $\pi(\{\mu : \|\mu-\mu_0-h/\sqrt{n}\|^2 < (Bn)^{-1}\})$ is bounded below by a multiple of $n^{-K/2}$ for all n sufficiently large, and hence that $\pi(\{\mu : D_{KL}(p_{\mu_0+h/\sqrt{n}}\|p_\mu) < n^{-1}\})$, is bounded below by a multiple of $n^{-K/2}$ for all n sufficiently large. ■

Lemma 11 *Suppose that Assumptions 1(i) and 2 hold and $\pi \in \Pi$. Then there exists $C > 0$ such that*

$$P_{n,h}\left(\delta_n^*(X^n;\pi) \notin \underline{\mathcal{D}}_{\mu_0}\right) \leq \frac{C}{n}.$$

Proof of Lemma 11. We will prove the result by applying Lemma 9. We first establish existence of tests as required by Lemma 9. We proceed as in Propositions 6.1 and 6.2 of Clarke and Barron (1990). Let d_G denote the metric they construct to metrize weak convergence.

[17] The result is stated requiring $B > 1$, but inspection of the proof shows any $B > 0$ is valid: restricting to $B > 1$ is simply so the probability bound of $1/B$ is not vacuous.

Let $\hat{P}_n \equiv \hat{P}_n(X^n)$ denote the empirical measure. As shown in the proof of Proposition 6.2 of Clarke and Barron (1990), for any $\varepsilon > 0$ there exists an integer k depending only on ε such that

$$P_{n,\mu}\left(d_G(\hat{P}_n, P_\mu) \geq \varepsilon\right) \leq 2ke^{-n\varepsilon^2}, \qquad \mu \in \mathcal{M}. \tag{37}$$

By Assumption 2(v), there exists $\epsilon' > 0$ such that $\|\mu - \mu_0\| \geq \epsilon$ implies $d_G(P_\mu, P_{\mu_0}) \geq \epsilon'$. Consider the test $\phi_n(X^n) = \mathbb{I}[d_G(\hat{P}_n, P_{\mu_0}) > \epsilon'/2]$. Again using Assumption 2(v), take n large enough that $d_G(P_{\mu_0+h/\sqrt{n}}, P_{\mu_0}) < \epsilon'/4$. Then by (37), we have

$$\mathbb{E}_{n,h}[\phi_n(X^n)] \leq P_{n,h}\left(d_G(\hat{P}_n, P_{\mu_0+h/\sqrt{n}}) \geq \epsilon'/4\right) \leq C'e^{-nC}$$

for suitable C, C'. We may similarly deduce as in the proof of Proposition 6.1 of Clarke and Barron (1990) that for any μ with $\|\mu - \mu_0\| \geq \epsilon$ we have $\mathbb{E}_{n,\mu}[1 - \phi_n(X^n)] \leq C'e^{-nC}$ for suitable C, C' not depending on μ. Thus, we have verified existence of a sequence of tests with the desired properties.

By Lemma 9, it suffices to show $e^{nc}\pi(\{\mu : D_{KL}(p_{\mu_0+h/\sqrt{n}} \| p_\mu) < n^{-1}\}) \to +\infty$ for any $c > 0$. This follows immediately from Lemma 10. ■

D.2 Semiparametric Models

Lemma 12 *Suppose that Assumptions 1(i) and 3 hold and $\pi \in \Pi$. Then for any $(\mu_0, \eta_0) \in \mathcal{M} \times \mathcal{H}$,*

$$\sqrt{n}\, P_{n,h}\left(\delta_n^*(X^n; \pi) \notin \underline{\mathcal{D}}_{\mu_0}\right) \to 0$$

for all $h \in \mathbb{R}^K$.

Proof of Lemma 12. In view of Lemma 8, we have

$$P_{n,h}\left(\delta_n^*(X^n; \pi) \notin \underline{\mathcal{D}}_{\mu_0}\right) \leq P_{n,h}\left(\pi_n(\{\mu : \|\mu - \mu_0\| \geq \epsilon\}) > 2\epsilon'\right)$$

for some $\epsilon, \epsilon' > 0$. By Assumption 3(i), choose $C > 0$ such that $\{\mu : \|\mu - \mu_0\| \leq C\} \subset \mathcal{M}$. Taking ϵ smaller if needed, we may assume $\epsilon \leq C$. Recall that $\hat{\lambda}_{\min}$ and $\hat{\lambda}_{\max}$ denote the smallest and largest eigenvalues of $\hat{I}$. Let A_n denote the event upon which $\|\hat{\mu} - \mu_0\| \leq \epsilon/2$ and $c < \hat{\lambda}_{\min} \leq \hat{\lambda}_{\max} < c^{-1}$ both hold, where the constant $c \in (0, 1)$ is from Assumption 3(iv)b. Note π has a positive, continuous, and bounded Lebesgue density $\pi(\mu)$ on $\mathcal{M}$, hence $\overline{\pi} :=$

$\sup_{\mu\in\mathcal{M}} \pi(\mu) < \infty$ and $\underline{\pi} := \inf_{\mu:\|\mu-\mu_0\|\leq C} \pi(\mu) > 0$. It follows that on A_n, we have

$$\begin{aligned}\pi_n(\{\mu : \|\mu-\mu_0\| \geq \epsilon\}) &\leq \pi_n(\{\mu : \|\mu-\hat{\mu}\| \geq \epsilon/2\}) \\ &\leq \frac{\int_{\mu:\|\mu-\hat{\mu}\|\geq\epsilon/2} e^{-\frac{1}{2}(\mu-\hat{\mu})^T(n\hat{I})(\mu-\hat{\mu})}\, d\pi(\mu)}{\int_{\mu:\|\mu-\hat{\mu}\|\leq C-\epsilon/2} e^{-\frac{1}{2}(\mu-\hat{\mu})^T(n\hat{I})(\mu-\hat{\mu})}\, d\pi(\mu)} \\ &\leq \frac{\overline{\pi}\int_{h:\|h\|\geq\sqrt{n}\epsilon/2} e^{-\frac{c}{2}\|h\|^2}\, dh}{\underline{\pi}\int_{h:\|h\|\leq\sqrt{n}(C-\epsilon/2)} e^{-\frac{1}{2c}\|h\|^2}\, dh} \to 0,\end{aligned}$$

where the first inequality is by the reverse triangle inequality, the final inequality is by a change of variables from μ to $h = \sqrt{n}(\mu-\hat{\mu})$, and convergence to zero is by the dominated convergence theorem. We therefore have that $\pi_n(\{\mu : \|\mu-\mu_0\| \geq \epsilon\}) < 2\epsilon'$ holds on A_n for all n sufficiently large, in which case

$$P_{n,h}\left(\pi_n(\{\mu : \|\mu-\mu_0\| \geq \epsilon\}) > 2\epsilon'\right) \leq P_{n,h}(A_n^c).$$

Finally, $\sqrt{n}P_{n,h}(A_n^c) \to 0$ by Assumption 3(iv). ■

References

Clarke, B., and A. Barron (1990): "Information-theoretic asymptotics of Bayes methods," *IEEE Transactions on Information Theory*, 36(3), 453–471.

Ghosal, S., and A. van der Vaart (2017): *Fundamentals of Nonparametric Bayesian Inference*. Cambridge University Press.

Kitagawa, T., J. L. Montiel Olea, J. Payne, and A. Velez (2020): "Posterior distribution of nondifferentiable functions," *Journal of Econometrics*, 217(1), 161–175.

Schwartz, L. (1965): "On Bayes procedures," *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 4(1), 10–26.

Shapiro, A. (1990): "On concepts of directional differentiability," *Journal of Optimization Theory and Applications*, 66(3), 477–487.

van der Vaart, A. (1991): "An Asymptotic Representation Theorem," *International Statistical Review*, 59(1), 97–121.

——— (1998): *Asymptotic Statistics*. Cambridge University Press, Cambridge.