Semiparametric Modelling of Two-component Mixtures with Stochastic Dominance

Jingjing Wu[1]*, Tasnima Abedin[2] and Qiang Zhao[3]

[1]*Department of Mathematics and Statistics, University of Calgary*
2500 University Drive NW, Calgary, AB Canada T2N 1N4
[2]*Clinical Research Unit & Translational Laboratories, Alberta Health Services*
1331 29 Street NW, Calgary, AB Canada T2N 4N2
[3]*School of Mathematics and Statistics, Shandong Normal University*
No.1 University Road, Science Park, Changqing District, Jinan, Shandong China 250358

Supplementary material: derivations and proofs

S.1. Derivations of the MELE.

To find the MELE, we use the Lagrange multipliers and maximize

$$\sum_{i=1}^{m+n} \log p_i + \sum_{j=1}^{n} \log\left[(1-\lambda) + \lambda e^{\alpha+\beta Y_j}\right] - t_1 \left[\sum_{i=1}^{m+n} p_i - 1\right] - t_2 \left[\sum_{i=1}^{m+n} p_i e^{\alpha+\beta T_i} - 1\right].$$

Taking partial derivatives gives the estimating equation system

$$\frac{1}{p_i} - t_1 - t_2 e^{\alpha+\beta T_i} \;=\; 0, \qquad i = 1, ..., m+n, \tag{S.1}$$

$$\sum_{j=1}^{n} \frac{e^{\alpha+\beta Y_j} - 1}{(1-\lambda) + \lambda e^{\alpha+\beta Y_j}} \;=\; 0, \tag{S.2}$$

$$\sum_{j=1}^{n} \frac{\lambda e^{\alpha+\beta Y_j}}{(1-\lambda) + \lambda e^{\alpha+\beta Y_j}} - t_2 \sum_{i=1}^{m+n} p_i e^{\alpha+\beta T_i} \;=\; 0, \tag{S.3}$$

$$\sum_{j=1}^{n} \frac{Y_j \lambda e^{\alpha+\beta Y_j}}{(1-\lambda) + \lambda e^{\alpha+\beta Y_j}} - t_2 \sum_{i=1}^{m+n} p_i T_i e^{\alpha+\beta T_i} \;=\; 0,$$

*Corresponding author. Email: jinwu@ucalgary.ca. Phone: 1-403-2206303. Fax: 1-403-2825150.

$$\sum_{i=1}^{m+n} p_i \;=\; 1, \qquad \sum_{i=1}^{m+n} p_i e^{\alpha+\beta T_i} \;=\; 1. \tag{S.4}$$

From (S.2) and $\displaystyle\sum_{j=1}^{n} \frac{(1-\lambda)+\lambda e^{\alpha+\beta Y_j}}{(1-\lambda)+\lambda e^{\alpha+\beta Y_j}} = n$, we have $\sum_{j=1}^{n} \frac{e^{\alpha+\beta Y_j}}{(1-\lambda)+\lambda e^{\alpha+\beta Y_j}} = \sum_{j=1}^{n} \frac{1}{(1-\lambda)+\lambda e^{\alpha+\beta Y_j}} = n$ and plugging it into (S.3) gives

$$n\lambda - t_2 \sum_{i=1}^{m+n} p_i e^{\alpha+\beta T_i} \;=\; 0. \tag{S.5}$$

From (S.1) we get $(m+n) - t_1 \sum_{i=1}^{m+n} p_i - t_2 \sum_{i=1}^{m+n} p_i e^{\alpha+\beta T_i} = 0$. This together with (S.4) and (S.5) gives $t_2 = n\lambda$ and $t_1 = m+n-n\lambda$. Then by (S.1) again we have

$$p_i \;=\; \frac{1}{(m+n)\left[1+\rho_N \lambda \left(e^{\alpha+\beta T_i}-1\right)\right]},$$

where $\rho_N = n/(m+n)$ with $N = m+n$.

S.2. Proof of Theorem 1.

Since we have a sample from f, as a result f is identifiable. Given this, when $h_{\theta_1} = h_{\theta_2}$, i.e.,

$$\left\{1-\lambda_1+\lambda_1 \exp\left[\alpha_1+\beta_1^\top r(x)\right]\right\} f(x) = \left\{1-\lambda_2+\lambda_2 \exp\left[\alpha_2+\beta_2^\top r(x)\right]\right\} f(x) \ \text{ for all } x,$$

we must have $\lambda_1 = \lambda_2$, $\alpha_1 = \alpha_2$ and $\beta_1 = \beta_2$ by the assumption that $(1, r(y))^\top$ is linearly independent on the support of f. Thus h_θ is identifiable.

Let $s(x) = f(x) - g(x) = f(x)\left[1-\exp(\alpha+\beta x)\right]$ and $S(x) = \int_{-\infty}^{x} s(t)dt = F(x) - G(x)$. Let x_0 denote the solution to $1-\exp(\alpha+\beta x) = 0$. Then $s(x) > 0$ when $x < x_0$ and $s(x) \le 0$ when $x \ge x_0$, and hence $S(x)$ increases for $x < x_0$ and decreases for $x \ge x_0$. If $F(x') < G(x')$ for some x', i.e., $S(x') < 0$, then $x' \ge x_0$ since $S(x) > 0$ for all $x < x_0$. Since $S(x)$ decreases when $x \ge x_0$, we have $S(x) \le S(x') < 0$ for all $x > x'$ and thus $S(\infty) < 0$. However $S(\infty) = F(\infty) - G(\infty) = 1-1 = 0$, a contradiction. Therefore $F \ge G$. □

S.3. Proof of Lemma 1.

Define $w_{2N}(y) = 1 - \rho_N\lambda + \rho_N\lambda e^{\alpha+\beta y}$. The second-order partial derivatives of the empirical log-likelihood function l in (7) are

$$\frac{\partial^2 l}{\partial \lambda^2} \;=\; -\sum_{j=1}^{n} \frac{(e^{\alpha+\beta Y_j}-1)^2}{w_1^2(Y_j)} + \sum_{i=1}^{m+n} \frac{\rho_N^2 (e^{\alpha+\beta T_i}-1)^2}{w_{2N}^2(T_i)},$$

$$\frac{\partial^2 l}{\partial \alpha^2} = \sum_{j=1}^{n} \frac{\lambda(1-\lambda)e^{\alpha+\beta Y_j}}{w_1^2(Y_j)} - \sum_{i=1}^{m+n} \frac{\rho_N \lambda(1-\rho_N\lambda)e^{\alpha+\beta T_i}}{w_{2N}^2(T_i)},$$

$$\frac{\partial^2 l}{\partial \beta^2} = \sum_{j=1}^{n} \frac{\lambda(1-\lambda)Y_j^2 e^{\alpha+\beta Y_j}}{w_1^2(Y_j)} - \sum_{i=1}^{m+n} \frac{\rho_N \lambda(1-\rho_N\lambda)T_i^2 e^{\alpha+\beta T_i}}{w_{2N}^2(T_i)},$$

$$\frac{\partial^2 l}{\partial \lambda \partial \alpha} = \sum_{j=1}^{n} \frac{e^{\alpha+\beta Y_j}}{w_1^2(Y_j)} - \sum_{i=1}^{m+n} \frac{\rho_N e^{\alpha+\beta T_i}}{w_{2N}^2(T_i)},$$

$$\frac{\partial^2 l}{\partial \lambda \partial \beta} = \sum_{j=1}^{n} \frac{Y_j e^{\alpha+\beta Y_j}}{w_1^2(Y_j)} - \sum_{i=1}^{m+n} \frac{\rho_N T_i e^{\alpha+\beta T_i}}{w_{2N}^2(T_i)},$$

$$\frac{\partial^2 l}{\partial \alpha \partial \beta} = \sum_{j=1}^{n} \frac{\lambda(1-\lambda)Y_j e^{\alpha+\beta Y_j}}{w_1^2(Y_j)} - \sum_{i=1}^{m+n} \frac{\rho_N \lambda(1-\rho_N\lambda)T_i e^{\alpha+\beta T_i}}{w_{2N}^2(T_i)}.$$

Straight calculation gives

$$E\left[-\frac{1}{N}\cdot\frac{\partial^2 l}{\partial \lambda^2}\right] = \rho_N(1-\rho_N)\int (e^{\alpha+\beta y}-1)^2 \frac{f}{w_1 w_{2N}}(y)\, dy \longrightarrow \rho(1-\rho)S_{11}.$$

By WLLN, $-\frac{1}{N}\cdot\frac{\partial^2 l}{\partial \lambda^2} \xrightarrow{P} \rho(1-\rho)S_{11}$. Similarly we have the convergence of other components of the matrix S_N. □

S.4. Proof of Theorem 2.

Let $Q_N = \frac{1}{N}\left(\frac{\partial l}{\partial \lambda}, \frac{\partial l}{\partial \alpha}, \frac{\partial l}{\partial \beta}\right)^\top$, then $E[Q_N]=0$. Note that as $N\to\infty$,

$$\begin{aligned}
\frac{1}{N}Var\left[\frac{\partial l}{\partial \lambda}\right] &= \frac{1}{N}Var\left[\sum_{j=1}^{n}\left(\frac{e^{\alpha+\beta Y_j}-1}{w_1(Y_j)} - \frac{\rho_N(e^{\alpha+\beta Y_j}-1)}{w_{2N}(Y_j)}\right) - \sum_{i=1}^{m}\frac{\rho_N(e^{\alpha+\beta X_i}-1)}{w_{2N}(X_i)}\right] \\
&= \rho_N Var\left[\frac{(1-\rho_N)(e^{\alpha+\beta Y_1}-1)}{w_1(Y_1)w_{2N}(Y_1)}\right] + (1-\rho_N)Var\left[\frac{\rho_N(e^{\alpha+\beta X_1}-1)}{w_{2N}(X_1)}\right] \\
&= \rho_N(1-\rho_N)^2\left\{\int \frac{(e^{\alpha+\beta y}-1)^2}{w_1^2(y)w_{2N}^2(y)}w_1(y)f(y)dy - \left[\int \frac{e^{\alpha+\beta y}-1}{w_1(y)w_{2N}(y)}w_1(y)f(y)dy\right]^2\right\} \\
&\quad + \rho_N^2(1-\rho_N)\left\{\int \frac{(e^{\alpha+\beta y}-1)^2}{w_{2N}^2(y)}f(y)dy - \left[\int \frac{e^{\alpha+\beta y}-1}{w_{2N}(y)}f(y)dy\right]^2\right\} \\
&= \rho_N(1-\rho_N)\left[\int (e^{\alpha+\beta y}-1)^2\frac{f}{w_1 w_{2N}}(y)dy - \left[\int (e^{\alpha+\beta y}-1)\frac{f}{w_{2N}}(y)dy\right]^2\right] \\
&\to \rho(1-\rho)V_{11}.
\end{aligned}$$

Similarly we have $\frac{1}{N}Var\left[\frac{\partial l}{\partial \alpha}\right] \to \rho(1-\rho)V_{22}$ and $\frac{1}{N}Var\left[\frac{\partial l}{\partial \beta}\right] \to \rho(1-\rho)V_{33}$ as $N\to\infty$.

Note that

$$\begin{aligned}
\frac{1}{N}Cov\left[\frac{\partial l}{\partial\lambda},\frac{\partial l}{\partial\alpha}\right] &= \frac{1}{N}E\left[\frac{\partial l}{\partial\lambda}\cdot\frac{\partial l}{\partial\alpha}\right] \\
&= \frac{1}{N}E\left[\left\{\sum_{j=1}^{n}\left(\frac{e^{\alpha+\beta Y_j}-1}{w_1(Y_j)}-\frac{\rho_N(e^{\alpha+\beta Y_j}-1)}{w_{2N}(Y_j)}\right)-\sum_{i=1}^{m}\frac{\rho_N(e^{\alpha+\beta X_i}-1)}{w_{2N}(X_i)}\right\}\right. \\
&\qquad\left.\cdot\left\{\sum_{j=1}^{n}\left(\frac{\lambda e^{\alpha+\beta Y_j}}{w_1(Y_j)}-\frac{\rho_N\lambda e^{\alpha+\beta Y_j}}{w_{2N}(Y_j)}\right)-\sum_{i=1}^{m}\frac{\rho_N\lambda e^{\alpha+\beta X_i}}{w_{2N}(X_i)}\right\}\right] \\
&= \frac{1}{N}E\left[\left\{(1-\rho_N)\sum_{j=1}^{n}\frac{e^{\alpha+\beta Y_j}-1}{w_1(Y_j)w_{2N}(Y_j)}-\rho_N\sum_{i=1}^{m}\frac{e^{\alpha+\beta X_i}-1}{w_{2N}(X_i)}\right\}\right. \\
&\qquad\left.\cdot\left\{(1-\rho_N)\lambda\sum_{j=1}^{n}\frac{e^{\alpha+\beta Y_j}}{w_1(Y_j)w_{2N}(Y_j)}-\rho_N\lambda\sum_{i=1}^{m}\frac{e^{\alpha+\beta X_i}}{w_{2N}(X_i)}\right\}\right] \\
&= \frac{1}{N}E\left[(A-B)(C-D)\right],\ \text{say} \\
&= \frac{1}{N}\left\{E[AC]+E[BD]-E[A]E[D]-E[B]E[C]\right\},
\end{aligned}$$

where

$$\begin{aligned}
E\left[AC\right] &= (1-\rho_N)^2\lambda\left\{nE\left[\frac{e^{\alpha+\beta Y_1}(e^{\alpha+\beta Y_1}-1)}{w_1^2(Y_1)w_{2N}^2(Y_1)}\right]\right. \\
&\qquad\left.+\ n(n-1)E\left[\frac{e^{\alpha+\beta Y_1}-1}{w_1(Y_1)w_{2N}(Y_1)}\right]E\left[\frac{e^{\alpha+\beta Y_1}}{w_1(Y_1)w_{2N}(Y_1)}\right]\right\} \\
&= n(1-\rho_N)^2\lambda\left\{\int e^{\alpha+\beta y}(e^{\alpha+\beta y}-1)\frac{f}{w_1w_{2N}^2}(y)\,dy\right. \\
&\qquad\left.+\ (n-1)\int(e^{\alpha+\beta y}-1)\frac{f}{w_{2N}}(y)\,dy\int e^{\alpha+\beta y}\frac{f}{w_{2N}}(y)\,dy\right\},
\end{aligned}$$

$$\begin{aligned}
E\left[BD\right] &= m\rho_N^2\lambda\left\{\int e^{\alpha+\beta y}(e^{\alpha+\beta y}-1)\frac{f}{w_{2N}^2}(y)\,dy\right. \\
&\qquad\left.+\ (m-1)\int(e^{\alpha+\beta y}-1)\frac{f}{w_{2N}}(y)\,dy\int e^{\alpha+\beta y}\frac{f}{w_{2N}}(y)\,dy\right\},
\end{aligned}$$

$$E[A]E[D] = E[B]E[C] = mn\rho_N(1-\rho_N)\lambda\int(e^{\alpha+\beta y}-1)\frac{f}{w_{2N}}(y)\,dy\int e^{\alpha+\beta y}\frac{f}{w_{2N}}(y)\,dy,$$

and thus as $N \to \infty$,

$$\begin{aligned}
\frac{1}{N} Cov\left[\frac{\partial l}{\partial \lambda}, \frac{\partial l}{\partial \alpha}\right] &= \rho_N(1-\rho_N)^2 \lambda \int e^{\alpha+\beta y}(e^{\alpha+\beta y}-1)\frac{f}{w_1 w_{2N}^2}(y)\, dy \\
&\quad + (n-1)\rho_N(1-\rho_N)^2 \lambda \int (e^{\alpha+\beta y}-1)\frac{f}{w_{2N}}(y)\, dy \int e^{\alpha+\beta y}\frac{f}{w_{2N}}(y)\, dy \\
&\quad + \rho_N^2(1-\rho_N)\lambda \int e^{\alpha+\beta y}(e^{\alpha+\beta y}-1)\frac{f}{w_{2N}^2}(y)\, dy \\
&\quad + (m-1)\rho_N^2(1-\rho_N)\lambda \int (e^{\alpha+\beta y}-1)\frac{f}{w_{2N}}(y)\, dy \int e^{\alpha+\beta y}\frac{f}{w_{2N}}(y)\, dy \\
&\quad - 2m\rho_N^2(1-\rho_N)\lambda \int (e^{\alpha+\beta y}-1)\frac{f}{w_{2N}}(y)\, dy \int e^{\alpha+\beta y}\frac{f}{w_{2N}}(y)\, dy \\
&= \rho_N(1-\rho_N)\lambda \Bigg[\int e^{\alpha+\beta y}(e^{\alpha+\beta y}-1)\frac{f}{w_1 w_{2N}^2}(y)\, dy \\
&\qquad - \int (e^{\alpha+\beta y}-1)\frac{f}{w_{2N}}(y)\, dy \int e^{\alpha+\beta y}\frac{f}{w_{2N}}(y)\, dy\Bigg] \\
&\to \rho(1-\rho)V_{12}.
\end{aligned}$$

Similarly we have $\frac{1}{N} Cov\left[\frac{\partial l}{\partial \lambda}, \frac{\partial l}{\partial \beta}\right] \to \rho(1-\rho)V_{13}$ and $\frac{1}{N} Cov\left[\frac{\partial l}{\partial \alpha}, \frac{\partial l}{\partial \beta}\right] \to \rho(1-\rho)V_{23}$ as $N \to \infty$. Thus by CLT, $\sqrt{N}Q_N \xrightarrow{L} N(0, \rho(1-\rho)V)$. From Lemma 1 along with Slutsky's theorem, we have $\sqrt{N}(\hat{\theta}_{MELE} - \theta) \xrightarrow{L} N(0, \Sigma)$.

From above calculation we see that $\frac{1}{N\rho(1-\rho)} Cov[\frac{\partial l}{\partial \theta}] \to V$. Though α is uniquely determined by β from relationship (3), α depends on β non-linearly. As a result the vector $Q_N = \frac{1}{N}\left(\frac{\partial l}{\partial \lambda}, \frac{\partial l}{\partial \alpha}, \frac{\partial l}{\partial \beta}\right)^\top$ is linearly independent. Thus $Cov[\frac{\partial l}{\partial \theta}]$ is positive definite, and further so do V and Σ. □

The proofs of Theorems 3-6 are similar to Wu, Karunamuni and Zhang (2010) but for a different model. For completeness we present the proofs below.

S.5. Proof of Theorem 3.

In order to prove Theorem 3, we need the following lemma.

Lemma 2. *For any $\varphi \in \mathcal{H}$, $d(t) = \|h_t^{1/2} - \varphi^{1/2}\|$ is continuous at point $t = \theta$.*

Proof. Suppose $\theta_k \to \theta$ as $k \to \infty$. From Minkowski's inequality,

$$|d(\theta_k) - d(\theta)| \;\le\; \left\|h_{\theta_k}^{1/2} - h_\theta^{1/2}\right\| \;\le\; \left[\int |h_{\theta_k}(x) - h_\theta(x)|\, dx\right]^{1/2}. \tag{S.6}$$

For any $x \in \mathbb{R}$, as $k \to \infty$, $\theta_k \to \theta$ implies

$$
\begin{aligned}
|h_{\theta_k}(x) - h_\theta(x)| &= \left|-(\lambda_k - \lambda)f(x) + \lambda_k f(x)(e^{\alpha_k+\beta_k x} - e^{\alpha+\beta x}) + (\lambda_k - \lambda)e^{\alpha+\beta x} f(x)\right| \\
&\to 0.
\end{aligned}
$$

Thus by Scheffé's theorem we have $\int |h_{\theta_k}(x) - h_\theta(x)|dx \to 0$ as $k \to \infty$, i.e., $d(\theta_k) \to d(\theta)$ as $k \to \infty$ and $d(t)$ is continuous at point $t = \theta$. □

Proof of Theorem 3. (i) Let $d_m(t) = \|\hat{h}_t^{1/2} - \varphi^{1/2}\|$. Suppose sequence $\{t_k\} \subset \Theta$ such that $t_k \to t$ as $k \to \infty$. Since Θ is compact, $t \in \Theta$. Similar to (S.6), we have

$$
|d_m(t_k) - d_m(t)| \le \left[\int \left|\lambda_k - \lambda - \lambda_k e^{\alpha_k+\beta_k x} + \lambda e^{\alpha+\beta x}\right| f_m(x)dx\right]^{1/2}.
$$

Since f_m is compactly supported, we have by (D1) and the Dominated Convergence Theorem (DCT) that $d_m(t_k) \to d_m(t)$ as $k \to \infty$, i.e., $d_m(t)$ is continuous and achieves a minimum over $t \in \Theta$. Let $d(t) = \|h_t^{1/2} - \varphi^{1/2}\|$. By Lemma 2, $d(t)$ is continuous in t and therefore achieves a minimum over $t \in \Theta$.

(ii) Suppose $\|\varphi_n^{1/2} - \varphi^{1/2}\| \to 0$ and $\sup_{t\in\Theta} \|\hat{h}_t^{1/2} - h_t^{1/2}\| \to 0$ as $N \to \infty$. Let $d_N(t) = \|\hat{h}_t^{1/2} - \varphi_n^{1/2}\|$ and $d(t) = \|h_t^{1/2} - \varphi^{1/2}(x)\|$. By Minkowski's inequality,

$$
\begin{aligned}
|d_N(t) - d(t)| &\le \left\{\int \left[\hat{h}_t^{1/2}(x) - \varphi_n^{1/2}(x) - h_t^{1/2}(x) + \varphi^{1/2}(x)\right]^2 dx\right\}^{1/2} \\
&\le \left\{2\int \left[\hat{h}_t^{1/2}(x) - h_t^{1/2}\right]^2 dx + 2\int \left[\varphi_n^{1/2} - \varphi^{1/2}(x)\right]^2 dx\right\}^{1/2}
\end{aligned}
$$

and consequently $\sup_{t\in\Theta} |d_N(t) - d(t)| \to 0$ as $N \to \infty$. Thus as $N \to \infty$, $d_N(\theta) \to d(\theta)$ and $d_N(\theta_N) - d(\theta_N) \to 0$. If $\theta_N \nrightarrow \theta$, then there exists a subsequence $\{\theta_{N_i}\} \subseteq \{\theta_N\}$ such that, $\theta_{N_i} \to \theta' \ne \theta$. Since Θ is compact, $\theta' \in \Theta$. Lemma 2 yields that $d(\theta_{N_i}) \to d(\theta')$. From the above results we have $d_{N_i}(\theta_{N_i}) - d_{N_i}(\theta) \to d(\theta') - d(\theta)$. By the definition of θ_{N_i}, $d_{N_i}(\theta_{N_i}) - d_{N_i}(\theta) \le 0$. Hence, $d(\theta') - d(\theta) \le 0$. But by the definition and uniqueness of θ, $d(\theta') > d(\theta)$. This is a contradiction. Therefore $\theta_N \to \theta$.

(iii) Since by Theorem 1 $\{h_t\}_{t\in\Theta}$ is identifiable, we have $T(f, h_\theta) = \theta$ uniquely for any $\theta \in \Theta$. □

S.6. Proof of Theorem 4.

In order to prove Theorem 4, we need the following lemma.

Lemma 3. *Suppose (D3) holds. Then as $m \to \infty$,*

$$\sup_{\theta\in\Theta} \int w_1(x) \left[f_m^{1/2}(x) - f^{1/2}(x)\right]^2 dx \ \xrightarrow{P}\ 0.$$

Proof. By the continuity of the function w_1 in θ and the compactness of Θ, there exists a $\theta_m \in \Theta$ which maximizes $\int w_1(x)[f_m^{1/2}(x) - f^{1/2}(x)]^2 dx$. By (S.6), (S.7) and a Taylor expansion, one has

$$\begin{aligned}
& E\left|\int I_{\{|x|>\alpha_m\}} w_1(x) f_m(x) dx\right| \\
= & \int\int I_{\{|x|>\alpha_m\}} w_1(x) \frac{1}{b_m} K_0\left(\frac{y-x}{b_m}\right) f(y) dy\, dx \\
= & \int I_{\{|x|>\alpha_m\}} w_1(x) \int K_0(t) f(x+tb_m) dt\, dx \\
= & \int I_{\{|x|>\alpha_m\}} w_1(x) \int K_0(t) \left[f(x) + f^{(1)}(x) t b_m + \frac{1}{2} f^{(2)}(\xi) t^2 b_m^2\right] dt\, dx \\
\le & \int I_{\{|x|>\alpha_m\}} h_\theta(x) dx + \frac{1}{2} b_m^2 \int I_{\{|x|>\alpha_m\}} h_\theta(x) \sup_{|t|\le a_0} \frac{f^{(2)}(x+tb_m)}{f(x)} dx \int t^2 K_0(t) dt \\
\le & \sup_{\theta\in\Theta} \int I_{\{|x|>\alpha_m\}} h_\theta(x) dx + \frac{1}{2} b_m^2 \int t^2 K_0(t) dt \sup_{\theta\in\Theta} \int I_{\{|x|>\alpha_m\}} h_\theta(x) \sup_{|t|\le a_0} \frac{f^{(2)}(x+tb_m)}{f(x)} dx \\
\to & \ 0.
\end{aligned}$$

Thus as $m \to \infty$, $\int I_{\{|x|>\alpha_m\}} w_1(x) f_m(x) dx \ \xrightarrow{P}\ 0$ and

$$\begin{aligned}
& \int I_{\{|x|>\alpha_m\}} w_1(x) \left[f_m^{1/2}(x) - f^{1/2}(x)\right]^2 dx \\
\le & \ 2\int I_{\{|x|>\alpha_m\}} w_1(x) \left[f_m(x) + f(x)\right] dx \\
\le & \ 2\int I_{\{|x|>\alpha_m\}} w_1(x) f_m(x) dx + 2\int I_{\{|x|>\alpha_m\}} h_\theta(x) dx \\
\xrightarrow{P} & \ 0.
\end{aligned} \tag{S.7}$$

On the other hand,

$$\begin{aligned}
\left|\int I_{\{|x|\le\alpha_m\}} w_1(x) \left[f_m^{1/2}(x) - f^{1/2}(x)\right]^2 dx\right| = & \int I_{\{|x|\le\alpha_m\}} w_1(x) \frac{[f_m(x) - f(x)]^2}{\left[f_m^{1/2}(x) + f^{1/2}(x)\right]^2} dx \\
\le & \int I_{\{|x|\le\alpha_m\}} w_1(x) f^{-1}(x) [f_m(x) - f(x)]^2 dx \\
\le & \ 2\int I_{\{|x|\le\alpha_m\}} w_1(x) f^{-1}(x) [f_m(x) - E[f_m(x)]]^2 dx \\
& +2\int I_{\{|x|\le\alpha_m\}} w_1(x) f^{-1}(x) \left[E[f_m(x)] - f(x)\right]^2 dx \\
= & \ 2(A_{1m} + A_{2m}), \quad \text{say.}
\end{aligned}$$

Now by (A.3) as $m \to \infty$

$$
\begin{aligned}
E[A_{1m}] &= \int I_{\{|x|\le\alpha_m\}} w_1(x) f^{-1}(x) E[f_m(x) - E[f_m(x)]]^2 dx \\
&\le \int I_{\{|x|\le\alpha_m\}} w_1(x) f^{-1}(x) \frac{1}{m b_m^2} \int K_0^2\left(\frac{y-x}{b_m}\right) f(y)dydx \\
&= m^{-1} b_m^{-1} \int I_{\{|x|\le\alpha_m\}} w_1(x) \int_{-a_0}^{a_0} K_0^2(t) f(x+tb_m) f^{-1}(x) dtdx \\
&\le m^{-1} b_m^{-1} \int_{-a_0}^{a_0} K_0^2(t)dt \sup_{\theta\in\Theta} \int I_{\{|x|\le\alpha_m\}} h_\theta(x) \sup_{|t|\le a_0} \frac{f(x+tb_m)}{f^2(x)} dx \\
&\to 0,
\end{aligned}
$$

i.e, $A_{1m} \xrightarrow{P} 0$ as $m \to \infty$. By a Taylor expansion and (A.4),

$$
\begin{aligned}
|A_{2m}| &= \int I_{\{|x|\le\alpha_m\}} w_1(x) f^{-1}(x) \left[\int_{-a_0}^{a_0} K_0(t)(f(x+tb_m) - f(x))dt\right]^2 dx \\
&\le \frac{1}{4} b_m^4 \int I_{\{|x|\le\alpha_m\}} w_1(x) f^{-1}(x) \left[\sup_{|t|\le a_0} |f^{(2)}(x+tb_m)| \int_{-a_0}^{a_0} t^2 K_0(t)\right]^2 dx \\
&\le \frac{1}{4} b_m^4 \left[\int_{-a_0}^{a_0} K_0(t) t^2 dt\right]^2 \sup_{\theta\in\Theta} \int I_{\{|x|\le\alpha_m\}} h_\theta(x) \sup_{|t|\le a_0} \left[\frac{f^{(2)}(x+tb_m)}{f(x)}\right]^2 dx \\
&\to 0
\end{aligned}
$$

Therefore, $\int I_{\{|x|\le\alpha_m\}} w_1(x)[f_m^{1/2}(x) - f^{1/2}(x)]^2 dx \xrightarrow{P} 0$ as $m \to \infty$. This combined with (S.7) gives $\int w_1(x)[f_m^{1/2}(x) - f^{1/2}(x)]^2 dx \xrightarrow{P} 0$ for any $\theta \in \Theta$. By the continuity of the function in θ and the compactness of Θ, hence the result. □

Proof of Theorem 4. If we can prove that $\|h_n^{1/2} - h_\theta^{1/2}\| \xrightarrow{P} 0$ and $\sup_{t\in\Theta} \|\hat{h}_t^{1/2} - h_t^{1/2}\| \xrightarrow{P} 0$ as $N \to \infty$, then by Theorem 3 (iii) and then (ii) we have $\hat{\theta}_{MHDE} \xrightarrow{P} \theta$ as $N \to \infty$.

It is known that $f_m \xrightarrow{P} f$ and $h_n \xrightarrow{P} h$ as $N \to \infty$ (see Rao, 1983). Since $\int h_\theta(x)dx = \int h_n(x)dx = 1$, $\int[h_\theta(x) - h_n(x)]^+ dx = \int[h_\theta(x) - h_n(x)]^- dx$ and $\|h_n^{1/2} - h_\theta^{1/2}\|^2 \le \int |h_\theta(x) - h_n(x)|dx = 2\int[h_\theta(x) - h_n(x)]^+ dx$. Since, $[h_\theta(x) - h_n(x)]^+ < h_\theta(x)$, by the DCT it follows that $\|h_n^{1/2} - h_\theta^{1/2}\| \xrightarrow{P} 0$ as $n \to \infty$. Similarly $\|f_m^{1/2} - f^{1/2}\| \xrightarrow{P} 0$ as $m \to \infty$.

Note that $\int[\hat{h}_\theta^{1/2}(x) - h_\theta^{1/2}(x)]^2 dx = \int w_1(x)[f_m^{1/2}(x) - f^{1/2}(x)]^2 dx \le \int w_1(x)|f_m(x) - f(x)|dx$. If (D2) holds then $f_m - f$ will have a compact support on which $w_1(x)$ is bounded. Therefore, $\int[\hat{h}_\theta^{1/2}(x) - h_\theta^{1/2}(x)]^2 dx \le C_1 \int |f_m(x) - f(x)|dx = 2C_1 \int[f(x) - f_m(x)]^+ dx$ for some positive number C_1. Since $f_m \xrightarrow{P} f$, by the DCT we have $\sup_{\theta\in\Theta} \|\hat{h}_\theta^{1/2} - h_\theta^{1/2}\| \xrightarrow{P} 0$. If (D3) holds then Lemma 3 gives $\sup_{\theta\in\Theta} \|\hat{h}_\theta^{1/2} - h_\theta^{1/2}\| \xrightarrow{P} 0$. □

S.7. Proof of Theorem 5.

From Theorem 4 we have $\hat{\theta}_{MHDE} \xrightarrow{P} \theta$ as $N \to \infty$. Since $t = \hat{\theta}_{MHDE} \in \Theta$ minimizes the Hellinger distance between $\hat{h}_t$ and h_n, $\hat{\theta}_{MHDE}$ maximizes $2\int \hat{h}_t^{1/2}(x)h_n^{1/2}(x)dx - \int \hat{h}_t(x)dx$. Also since K_0 has compact support, we have

$$\int \frac{\partial}{\partial t}\left[2\hat{h}_t^{1/2}(x)h_n^{1/2}(x)dx - \hat{h}_t(x)\right]\Bigg|_{t=\hat{\theta}_{MHDE}} dx \;=\; 0.$$

For notation simplicity we use $\hat{\theta}$ to denote $\hat{\theta}_{MHDE}$ and use $\hat{w}_1$ to denote w_1 in (10) with θ replaced by $\hat{\theta}_{MHDE}$. Let $M_\theta(x) \;=\; 2\hat{h}_\theta^{1/2}(x)h_n^{1/2}(x)dx - \hat{h}_\theta(x)$, then by a Taylor expansion of $\hat{\theta}$ at θ it follows that

$$\int \frac{\partial M_\theta(x)}{\partial \theta}dx + \left[\int \frac{\partial^2 M_\theta(x)}{\partial\theta\partial\theta^\top}dx + R_N\right]\cdot\left(\hat{\lambda}-\lambda, \hat{\alpha}-\alpha, \hat{\beta}-\beta\right)^\top \;=\; 0, \tag{S.8}$$

where R_N is a 3×3 matrix with elements tending to zero in probability as $N \to \infty$. Direct calculation gives

$$\frac{\partial M_\theta(x)}{\partial\lambda} \;=\; (e^{\alpha+\beta x}-1)\left[\frac{f_m^{1/2}h_n^{1/2}}{w_1^{1/2}}(x) - f_m(x)\right],$$

$$\frac{\partial M_\theta(x)}{\partial\alpha} \;=\; \lambda e^{\alpha+\beta x}\left[\frac{f_m^{1/2}h_n^{1/2}}{w_1^{1/2}}(x) - f_m(x)\right],$$

$$\frac{\partial M_\theta(x)}{\partial\beta} \;=\; \lambda x e^{\alpha+\beta x}\left[\frac{f_m^{1/2}h_n^{1/2}}{w_1^{1/2}}(x) - f_m(x)\right],$$

$$\frac{\partial^2 M_\theta(x)}{\partial\lambda^2} \;=\; -\frac{(e^{\alpha+\beta x}-1)^2}{2w_1^{3/2}(x)}f_m^{1/2}(x)h_n^{1/2}(x), \tag{S.9}$$

$$\frac{\partial^2 M_\theta(x)}{\partial\lambda\partial\alpha} \;=\; \frac{e^{\alpha+\beta x}(w_1(x)+1)}{2w_1^{3/2}}f_m^{1/2}(x)h_n^{1/2}(x) - e^{\alpha+\beta x}f_m(x), \tag{S.10}$$

$$\frac{\partial^2 M_\theta(x)}{\partial\lambda\partial\beta} \;=\; \frac{xe^{\alpha+\beta x}(w_1(x)+1)}{2w_1^{3/2}(x)}f_m^{1/2}(x)h_n^{1/2}(x) - xe^{\alpha+\beta x}f_m(x), \tag{S.11}$$

$$\frac{\partial^2 M_\theta(x)}{\partial\alpha^2} \;=\; \frac{\lambda e^{\alpha+\beta x}(w_1(x)+1-\lambda)}{2w_1^{3/2}(x)}f_m^{1/2}(x)h_n^{1/2}(x) - \lambda e^{\alpha+\beta x}f_m(x), \tag{S.12}$$

$$\frac{\partial^2 M_\theta(x)}{\partial\alpha\partial\beta} \;=\; \frac{\lambda x e^{\alpha+\beta x}(w_1(x)+1-\lambda)}{2w_1^{3/2}(x)}f_m^{1/2}(x)h_n^{1/2}(x) - \lambda x e^{\alpha+\beta x}f_m(x), \tag{S.13}$$

$$\frac{\partial^2 M_\theta(x)}{\partial\beta^2} \;=\; \frac{\lambda x^2 e^{\alpha+\beta x}(w_1(x)+1-\lambda)}{2w_1^{3/2}(x)}f_m^{1/2}(x)h_n^{1/2}(x) - \lambda x^2 e^{\alpha+\beta x}f_m(x). \tag{S.14}$$

For (S.9) we have by Theorem 4 that

$$\begin{aligned}
& \left| \int \frac{(e^{\alpha+\beta x}-1)^2}{2w_1^{3/2}(x)} \left[f_m^{1/2}(x) h_n^{1/2}(x) - f^{1/2}(x) h_\theta^{1/2}(x) \right] dx \right| \\
\leq \; & C \left[\int f_m^{1/2}(x) \left| h_n^{1/2}(x) - h_\theta^{1/2}(x) \right| dx + \int h_\theta^{1/2}(x) \left| f_m^{1/2}(x) - f^{1/2}(x) \right| dx \right] \\
\leq \; & C \left[\left\| h_n^{1/2} - h_\theta^{1/2} \right\| + \left\| f_m^{1/2} - f^{1/2} \right\| \right] \\
\xrightarrow{P} \; & 0.
\end{aligned}$$

Thus for (S.9),

$$\begin{aligned}
-\int \frac{(e^{\alpha+\beta x}-1)^2}{2w_1^{3/2}(x)} f_m^{1/2}(x) h_n^{1/2}(x) \quad & \xrightarrow{P} \quad -\int \frac{(e^{\alpha+\beta x}-1)^2}{2w_1^{3/2}(x)} f^{1/2}(x) h_\theta^{1/2}(x) dx \\
& = \quad -\frac{1}{2} \int (e^{\alpha+\beta x}-1)^2 \frac{f}{w_1}(x) dx \\
& = \quad -\frac{1}{2} \Delta_{11}(\theta).
\end{aligned} \tag{S.15}$$

For (S.10), similarly we have

$$\begin{aligned}
\int \frac{e^{\alpha+\beta x}(w_1(x)+1)}{2w_1^{3/2}(x)} f_m^{1/2}(x) h_n^{1/2}(x) dx \quad & \xrightarrow{P} \quad \int \frac{e^{\alpha+\beta x}(w_1(x)+1)}{2w_1^{3/2}(x)} f^{1/2}(x) h_\theta^{1/2}(x) dx \\
& = \quad \int \frac{g(x)(w_1(x)+1)}{2w_1(x)} dx \\
& = \quad \frac{1}{2} + \frac{1}{2} \int \frac{e^{\alpha+\beta x}}{w_1(x)} f(x) dx
\end{aligned}$$

and

$$\begin{aligned}
\left| \int e^{\alpha+\beta x} \left[f_m(x) - f(x) \right] dx \right| \quad & \leq \quad C \int \left| \left[f_m^{1/2}(x) - f^{1/2}(x) \right] \left[f_m^{1/2}(x) + f^{1/2}(x) \right] \right| dx \\
& \leq \quad C \left\| f_m^{1/2} - f^{1/2} \right\| \cdot \left\| f_m^{1/2} + f^{1/2} \right\| \\
& \leq \quad 2C \left\| f_m^{1/2} - f^{1/2} \right\| \\
& \xrightarrow{P} \quad 0,
\end{aligned}$$

i.e. $\int e^{\alpha+\beta x} f_m(x) dx \xrightarrow{P} \int e^{\alpha+\beta x} f(x) dx$. Thus for (S.10),

$$\begin{aligned}
& \int \frac{e^{\alpha+\beta x}(w_1(x)+1)}{2w_1^{3/2}(x)} f_m^{1/2}(x) h_n^{1/2}(x) - e^{\alpha+\beta x} f_m(x) dx \\
& \qquad \xrightarrow{P} \; -\frac{1}{2} \lambda \int e^{\alpha+\beta x}(e^{\alpha+\beta x}-1) \frac{f}{w_1}(x) dx \;=\; -\frac{1}{2} \Delta_{12}(\theta).
\end{aligned} \tag{S.16}$$

Similarly for (S.11)-(S.14),

$$\begin{aligned}
& \int \frac{x e^{\alpha+\beta x}(w_1(x)+1)}{2w_1^{3/2}(x)} f_m^{1/2}(x) h_n^{1/2}(x) - x e^{\alpha+\beta x} f_m(x) dx \\
& \qquad \xrightarrow{P} \; -\frac{1}{2} \lambda \int x e^{\alpha+\beta x}(e^{\alpha+\beta x}-1) \frac{f}{w_1}(x) dx \;=\; -\frac{1}{2} \Delta_{13}(\theta),
\end{aligned} \tag{S.17}$$

$$\int \frac{\lambda e^{\alpha+\beta x}(w_1(x)+1-\lambda)}{2w_1^{3/2}(x)} f_m^{1/2}(x)h_n^{1/2}(x) - \lambda e^{\alpha+\beta x} f_m(x)dx$$
$$\xrightarrow{P} -\frac{1}{2}\lambda^2 \int e^{2\alpha+2\beta x}\frac{f}{w_1}(x)dx \ = \ -\frac{1}{2}\Delta_{22}(\theta), \tag{S.18}$$

$$\int \frac{\lambda x e^{\alpha+\beta x}(w_1(x)+1-\lambda)}{2w_1^{3/2}(x)} f_m^{1/2}(x)h_n^{1/2}(x) - \lambda x e^{\alpha+\beta x} f_m(x)dx$$
$$\xrightarrow{P} -\frac{1}{2}\lambda^2 \int x e^{2\alpha+2\beta x}\frac{f}{w_1}(x)dx \ = \ -\frac{1}{2}\Delta_{23}(\theta), \tag{S.19}$$

$$\int \frac{\lambda x^2 e^{\alpha+\beta x}(w_1(x)+1-\lambda)}{2w_1^{3/2}(x)} f_m^{1/2}(x)h_n^{1/2}(x) - \lambda x^2 e^{\alpha+\beta x} f_m(x)dx$$
$$\xrightarrow{P} -\frac{1}{2}\lambda^2 \int x^2 e^{2\alpha+2\beta x}\frac{f}{w_1}(x)dx \ = \ -\frac{1}{2}\Delta_{33}(\theta). \tag{S.20}$$

Now together with (S.9)-(S.20), (S.8) is reduced to

$$A_N(\theta) + \left[-\frac{1}{2}\Delta(\theta) + R_N\right]\left(\hat{\theta} - \theta\right) = 0,$$

where, $\Delta(\theta)$ and $A_N(\theta)$ are given in (19) and (20) respectively. Hence the result. □

S.8. Proof of Theorem 6.

We give here the sketch of the proof and readers are referred to Abedin (2018) for details.

In order to find the asymptotic distribution of $\hat{\theta}_{MHDE} - \theta$, by (21) we only need to find the asymptotic distribution of $\sqrt{N}A_N(\theta)$. Note that by (20),

$$\begin{aligned}
A_N(\theta) &= \int \frac{\partial w_1}{\partial \theta}(x)\left[\frac{f_m^{1/2}h_n^{1/2}}{w_1^{1/2}}(x) - f_m(x)\right] dx \\
&= \int \frac{\partial w_1}{\partial \theta}(x)\frac{f_m^{1/2}}{w_1^{1/2}}(x)\left[h_n^{1/2}(x) - h_\theta^{1/2}(x)\right] dx - \int \frac{\partial w_1}{\partial \theta}(x) f_m^{1/2}(x)\left[f_m^{1/2}(x) - f^{1/2}(x)\right] dx \\
&= \int \frac{\partial w_1}{\partial \theta}(x)\frac{f^{1/2}}{w_1^{1/2}}(x)\left[h_n^{1/2}(x) - h_\theta^{1/2}(x)\right] dx \\
&\qquad + \int \frac{\partial w_1}{\partial \theta}(x)\frac{1}{w_1^{1/2}(x)}\left[f_m^{1/2}(x) - f^{1/2}(x)\right]\left[h_n^{1/2}(x) - h_\theta^{1/2}(x)\right] dx \\
&\quad - \int \frac{\partial w_1}{\partial \theta}(x) f^{1/2}(x)\left[f_m^{1/2}(x) - f^{1/2}(x)\right] dx \\
&\qquad - \int \frac{\partial w_1}{\partial \theta}(x)\left[f_m^{1/2}(x) - f^{1/2}(x)\right]\left[f_m^{1/2}(x) - f^{1/2}(x)\right] dx.
\end{aligned}$$

We can prove that as $N \to \infty$,

$$\sqrt{N}\int \frac{\partial w_1}{\partial \theta}(x)\frac{1}{w_1^{1/2}(x)}\left[f_m^{1/2}(x) - f^{1/2}(x)\right]\left[h_n^{1/2}(x) - h_\theta^{1/2}(x)\right] dx \ \xrightarrow{P} \ 0,$$

$$\sqrt{N}\int \frac{\partial w_1}{\partial \theta}(x)\left[f_m^{1/2}(x)-f^{1/2}(x)\right]\left[f_m^{1/2}(x)-f^{1/2}(x)\right]dx \xrightarrow{P} 0.$$

Thus we only need to give the asymptotic distribution of

$$\int \frac{\partial w_1}{\partial \theta}(x)\frac{f^{1/2}}{w_1^{1/2}}(x)\left[h_n^{1/2}(x)-h_\theta^{1/2}(x)\right]dx$$

and

$$\int \frac{\partial w_1}{\partial \theta}(x) f^{1/2}(x)\left[f_m^{1/2}(x)-f^{1/2}(x)\right]dx$$

separately as they are independent. □