Supplementary Material for *Semiparametric Estimation in Secondary Analysis of Case-Control Studies*

Yanyuan Ma and Raymond J. Carroll

Department of Statistics, 3143 TAMU, Texas A&M University, College Station, TX 77843-3143

ma@stat.tamu.edu and carroll@stat.tamu.edu

S.1 Derivation of the Efficient Score

Having obtained both the score function and the two spaces Λ and $\Lambda^\perp$, we only need to project the score function onto $\Lambda^\perp$ to obtain the efficient score $\mathbf{S}_{\rm eff}$. To do this, we write $\mathbf{S}_\theta = \mathbf{S} - E(\mathbf{S} \mid D) = \mathbf{g}(\epsilon, \mathbf{X}) - E(\mathbf{g} \mid D) + \mathbf{S}_{\rm eff}$, where $E_{\rm true}(\mathbf{g}) = E_{\rm true}(\epsilon\mathbf{g} \mid \mathbf{X}) = \mathbf{0}$. We alternatively write $\mathbf{S}_{\rm eff} = \mathbf{S} - \mathbf{g}(\epsilon, \mathbf{X}) - E(\mathbf{S} - \mathbf{g} \mid D)$ and $\mathbf{S}_{\rm eff}$ satisfies $E\{\mathbf{S}_{\rm eff} - E(\mathbf{S}_{\rm eff} \mid D) \mid \epsilon, \mathbf{X}\} \sum_d (N_d/N) H(d, \mathbf{x}, Y)/p_D^{\rm true}(d) = \epsilon\mathbf{a}(\mathbf{X})$ and $E(\mathbf{S}_{\rm eff}) = 0$. However, $E(\mathbf{S}_{\rm eff} \mid d) = 0$ automatically, hence we can ignore the second requirement and the first requirement simplifies to $E(\mathbf{S}_{\rm eff} \mid \epsilon, \mathbf{X}) \sum_d (N_d/N)\{H(d, \mathbf{X}, Y)/p_D^{\rm true}(d)\} = \epsilon\mathbf{a}(\mathbf{X})$. This gives

$$\epsilon\mathbf{a}(\mathbf{X}) \left\{ \sum_d \frac{N_d}{N} \frac{H(d, \mathbf{X}, Y)}{p_D^{\rm true}(d)} \right\}^{-1} = E(\mathbf{S} - \mathbf{g} \mid \epsilon, \mathbf{X}) - E\left\{ E(\mathbf{S} - \mathbf{g} \mid D) \mid \epsilon, \mathbf{X} \right\}.$$

It follows that

$$f_{D|X,Y}(d, \mathbf{x}, y) = \frac{N_d}{N} \frac{H(d, \mathbf{x}, y)}{p_D^{\rm true}(d)} \left\{ \sum_d \frac{N_d}{N} \frac{H(d, \mathbf{x}, y)}{p_D^{\rm true}(d)} \right\}^{-1}.$$

To simplify notation, in the following calculation we denote

$$\begin{aligned}
\pi_0 &= p_D^{\text{true}}(0) = \int \eta_1(\mathbf{x})\eta_2(\epsilon,\mathbf{x})H(0,\mathbf{x},y)d\mu(\mathbf{x})d\mu(y);\\
\pi_1 &= p_D^{\text{true}}(1) = \int \eta_1(\mathbf{x})\eta_2(\epsilon,\mathbf{x})H(1,\mathbf{x},y)d\mu(\mathbf{x})d\mu(y);\\
b_0 &= E\{f_{D|\mathbf{X},Y}(1,\mathbf{X},Y) \mid D=0\};\\
b_1 &= E\{f_{D|\mathbf{X},Y}(0,\mathbf{X},Y) \mid D=1\};\\
\mathbf{c}_0 &= E(\mathbf{S} \mid D=0) - E\{E(\mathbf{S} \mid \epsilon,\mathbf{X}) \mid D=0\};\\
\mathbf{c}_1 &= E(\mathbf{S} \mid D=1) - E\{E(\mathbf{S} \mid \epsilon,\mathbf{X}) \mid D=1\};\\
\kappa(\mathbf{x},y) &= \left[\textstyle\sum_{d=0}^{1}\{N_d H(d,\mathbf{x},y)\}/(N\pi_d)\right]^{-1};\\
\mathbf{u}_0 &= E\left\{\epsilon\mathbf{a}(\mathbf{X})\kappa(\mathbf{X},Y) \mid D=0\right\};\\
\mathbf{u}_1 &= E\left\{\epsilon\mathbf{a}(\mathbf{X})\kappa(\mathbf{X},Y) \mid D=1\right\};\\
\mathbf{v}_0 &= E(\mathbf{S}-\mathbf{g} \mid D=0);\\
\mathbf{v}_1 &= E(\mathbf{S}-\mathbf{g} \mid D=1).
\end{aligned}$$

Note that $\pi_0+\pi_1=1$, $b_0N_0=b_1N_1$, $\mathbf{c}_0N_0+\mathbf{c}_1N_1=\mathbf{0}$ and $\mathbf{v}_0\pi_0+\mathbf{v}_1\pi_1=\mathbf{0}$.

Under a true model, $\pi_0,\pi_1,b_0,b_1,\mathbf{c}_0,\mathbf{c}_1$ are known quantities, while $\mathbf{u}_1,\mathbf{u}_1,\mathbf{v}_0,\mathbf{v}_1$ are not known because $\mathbf{g}=\mathbf{g}(\epsilon,\mathbf{x})$ and $\mathbf{a}=\mathbf{a}(\mathbf{x})$ are not specified. To further obtain $\mathbf{u}_1,\mathbf{u}_1,\mathbf{v}_0,\mathbf{v}_1$,

$$\begin{aligned}
\epsilon\mathbf{a}(\mathbf{x})\kappa(\mathbf{x},y) &= E(\mathbf{S}-\mathbf{g} \mid \epsilon,\mathbf{X}=\mathbf{x}) - \mathbf{v}_0 f_{D|X,Y}(0,\mathbf{x},y) - \mathbf{v}_1 f_{D|X,Y}(1,\mathbf{x},y)\\
&= E(\mathbf{S} \mid \epsilon,\mathbf{X}=\mathbf{x}) - \mathbf{g} - \mathbf{v}_0 f_{D|X,Y}(0,\mathbf{x},y) - \mathbf{v}_1 f_{D|X,Y}(1,\mathbf{x},y).
\end{aligned}$$

Alternatively, we also have

$$\mathbf{g}(\epsilon,\mathbf{x}) = E(\mathbf{S} \mid \epsilon,\mathbf{X}=\mathbf{x}) - \epsilon\mathbf{a}(\mathbf{x})\kappa(\mathbf{x},y) - \mathbf{v}_0 f_{D|X,Y}(0,\mathbf{x},y) - \mathbf{v}_1 f_{D|X,Y}(1,\mathbf{x},y). \quad \text{(S.1)}$$

Since $\mathbf{v}_0 = E(\mathbf{S}-\mathbf{g} \mid D=0)$, we obtain

$$\begin{aligned}
\mathbf{v}_0 &= E(\mathbf{S} \mid D=0) - E\big\{E(\mathbf{S} \mid \epsilon,\mathbf{X}) - \epsilon\mathbf{a}(\mathbf{X})\kappa(\mathbf{X},Y)\\
&\qquad -\mathbf{v}_0 f_{D|X,Y}(0,\mathbf{X},Y) - \mathbf{v}_1 f_{D|X,Y}(1,\mathbf{X},Y) \mid D=0\big\}\\
&= \mathbf{c}_0 + \mathbf{u}_0 + \mathbf{v}_0(1-b_0) + \mathbf{v}_1 b_0.
\end{aligned}$$

Thus, we have $b_0\mathbf{v}_0 - b_0\mathbf{v}_1 - \mathbf{u}_0 = \mathbf{c}_0$. Similarly, from $\mathbf{v}_1 = E(\mathbf{S}-\mathbf{g} \mid D=1)$, we obtain

$$\begin{aligned}
\mathbf{v}_1 &= E(\mathbf{S} \mid D=1) - E\big\{E(\mathbf{S} \mid \epsilon,\mathbf{X}) - \epsilon\mathbf{a}(\mathbf{X})\kappa(\mathbf{X},Y)\\
&\qquad -\mathbf{v}_0 f_{D|X,Y}(0,\mathbf{X},Y) - \mathbf{v}_1 f_{D|X,Y}(1,\mathbf{X},Y) \mid D=1\big\}\\
&= \mathbf{c}_1 + \mathbf{u}_1 + \mathbf{v}_0 b_1 + \mathbf{v}_1(1-b_1).
\end{aligned}$$

Thus, we have $-b_1\mathbf{v}_0 + b_1\mathbf{v}_1 - \mathbf{u}_1 = \mathbf{c}_1$. Since

$$E\{\epsilon\mathbf{a}(\mathbf{X})\kappa(\mathbf{X},Y)\} = \mathbf{0},$$

we have $\mathbf{u}_0 N_0 + \mathbf{u}_1 N_1 = \mathbf{0}$. Since $E_{\text{true}}(\mathbf{S}-\mathbf{g}) = \mathbf{0}$, we have

$$\begin{aligned}
\mathbf{0} &= \sum_d \int (\mathbf{S}-\mathbf{g}) f_{\mathbf{X},Y|D}(\mathbf{x},y,d)\{f^{\text{true}}_{X,Y,D}(\mathbf{x},y,d)/f_{\mathbf{X},Y|D}(\mathbf{x},y,d)\}d\mu(\mathbf{x})d\mu(y) \\
&= \sum_d \int (\mathbf{S}-\mathbf{g}) f_{\mathbf{X},Y|D}(\mathbf{x},y,d)\{f^{\text{true}}_{X,Y,D}(\mathbf{x},y,d)/f^{\text{true}}_{\mathbf{X},Y|D}(\mathbf{x},y,d)\}d\mu(\mathbf{x})d\mu(y) \\
&= \sum_d \int (\mathbf{S}-\mathbf{g}) f_{\mathbf{X},Y|D}(\mathbf{x},y,d) p^{\text{true}}_D(d) d\mu(\mathbf{x})d\mu(y) \\
&= \pi_0\mathbf{v}_0 + \pi_1\mathbf{v}_1.
\end{aligned}$$

Combining the above relations, we have obtained $N_0\mathbf{u}_0 + N_1\mathbf{u}_1 = \mathbf{0}$, $\pi_0\mathbf{v}_0 + \pi_1\mathbf{v}_1 = \mathbf{0}$, $b_0\mathbf{v}_0 - b_0\mathbf{v}_1 - \mathbf{u}_0 = \mathbf{c}_0$ and $-b_1\mathbf{v}_0 + b_1\mathbf{v}_1 - \mathbf{u}_1 = \mathbf{c}_1$. The last two equations are equivalent so one is redundant. Using these relations, we can rewrite $\mathbf{u}_1, \mathbf{v}_0, \mathbf{v}_1$ as a function of $\mathbf{u}_0$:

$$\mathbf{u}_1 = -(N_0/N_1)\mathbf{u}_0, \quad \mathbf{v}_0 = (\pi_1/b_0)(\mathbf{u}_0 + \mathbf{c}_0), \quad \mathbf{v}_1 = -(\pi_0/b_0)(\mathbf{u}_0 + \mathbf{c}_0). \tag{S.2}$$

We cannot obtain a more explicit expression for $\mathbf{u}_0$ at this stage, but we can further obtain $\mathbf{a}(\mathbf{x})$ as a function of $\mathbf{u}_0$. Using (S.1) and since $E_{\text{true}}(\epsilon\mathbf{g} \mid \mathbf{X}) = \mathbf{0}$, we have

$$\begin{aligned}
&E_{\text{true}}\{\epsilon E(\mathbf{S} \mid \epsilon, \mathbf{X}) \mid \mathbf{X}\} - E_{\text{true}}\left\{\epsilon^2\kappa(\mathbf{X},Y) \mid \mathbf{X}\right\}\mathbf{a}(\mathbf{X}) \\
&-\mathbf{v}_0 E_{\text{true}}\left\{\epsilon f_{D|X,Y}(0,\mathbf{X},Y) \mid \mathbf{X}\right\} - \mathbf{v}_1 E_{\text{true}}\left\{\epsilon f_{D|X,Y}(1,\mathbf{X},Y) \mid \mathbf{X}\right\} = \mathbf{0}.
\end{aligned}$$

Hence

$$\begin{aligned}
\mathbf{a}(\mathbf{X}) &= \left[E_{\text{true}}\left\{\epsilon^2\kappa(\mathbf{X},Y) \mid \mathbf{X}\right\}\right]^{-1} \\
&\quad \left[E_{\text{true}}\{\epsilon E(\mathbf{S} \mid \epsilon, \mathbf{X}) \mid \mathbf{X}\} - \mathbf{v}_0 E_{\text{true}}\left\{\epsilon f_{D|X,Y}(0,\mathbf{X},Y) \mid \mathbf{X}\right\}\right. \\
&\qquad \left. -\mathbf{v}_1 E_{\text{true}}\left\{\epsilon f_{D|X,Y}(1,\mathbf{X},Y) \mid \mathbf{X}\right\}\right] \\
&= \left[E_{\text{true}}\left\{\epsilon^2\kappa(\mathbf{X},Y) \mid \mathbf{X}\right\}\right]^{-1}\left[E_{\text{true}}\{\epsilon E(\mathbf{S} \mid \epsilon, \mathbf{X}) \mid \mathbf{X}\}\right. \\
&\quad -(\pi_1/b_0)(\mathbf{u}_0 + \mathbf{c}_0)E_{\text{true}}\left\{\epsilon f_{D|X,Y}(0,\mathbf{X},Y) \mid \mathbf{X}\right\} \\
&\qquad \left. +(\pi_0/b_0)(\mathbf{u}_0 + \mathbf{c}_0)E_{\text{true}}\left\{\epsilon f_{D|X,Y}(1,\mathbf{X},Y) \mid \mathbf{X}\right\}\right].
\end{aligned}$$

To further simplify notation, denote

$$\begin{aligned}
t_1(\mathbf{X}) &= \left[E_{\text{true}}\left\{\epsilon^2\kappa(\mathbf{X},Y) \mid \mathbf{X}\right\}\right]^{-1}; \qquad \text{(S.3)} \\
\mathbf{t}_2(\mathbf{X}) &= E_{\text{true}}\{\epsilon E(\mathbf{S} \mid \epsilon, \mathbf{X}) \mid \mathbf{X}\} - (\pi_1/b_0)\mathbf{c}_0 E_{\text{true}}\left\{\epsilon f_{D|X,Y}(0,\mathbf{X},Y) \mid \mathbf{X}\right\} \\
&\quad +(\pi_0/b_0)\mathbf{c}_0 E_{\text{true}}\left\{\epsilon f_{D|X,Y}(1,\mathbf{X},Y) \mid \mathbf{X}\right\} \\
&= E_{\text{true}}\{\epsilon E(\mathbf{S} \mid \epsilon, \mathbf{X}) \mid \mathbf{X}\} - (\mathbf{c}_0/b_0)E_{\text{true}}\left\{\epsilon f_{D|X,Y}(0,\mathbf{X},Y) \mid \mathbf{X}\right\}; \\
t_3(\mathbf{X}) &= -(\pi_1/b_0)E_{\text{true}}\left\{\epsilon f_{D|X,Y}(0,\mathbf{X},Y) \mid \mathbf{X}\right\} + (\pi_0/b_0)E_{\text{true}}\left\{\epsilon f_{D|X,Y}(1,\mathbf{X},Y) \mid \mathbf{X}\right\} \\
&= -b_0^{-1}E_{\text{true}}\left\{\epsilon f_{D|X,Y}(0,\mathbf{X},Y) \mid \mathbf{X}\right\}.
\end{aligned}$$

Then

$$\mathbf{a}(\mathbf{x}) = t_1(\mathbf{x})\{\mathbf{t}_2(\mathbf{x}) + t_3(\mathbf{x})\mathbf{u}_0\}, \tag{S.4}$$

hence the definition of $\mathbf{u}_0$ yields

$$\begin{aligned}\mathbf{u}_0 &= E\left(\epsilon\left[t_1(\mathbf{X})\{\mathbf{t}_2(\mathbf{X}) + t_3(\mathbf{X})\mathbf{u}_0\}\right]\kappa(\mathbf{X}, Y) \mid D = 0\right)\\ &= E\left[\epsilon t_1(\mathbf{X})\mathbf{t}_2(\mathbf{X})\kappa(\mathbf{X}, Y) \mid D = 0\right] + E\left[\epsilon t_1(\mathbf{X})t_3(\mathbf{X})\kappa(\mathbf{X}, Y) \mid D = 0\right]\mathbf{u}_0.\end{aligned}$$

This yields

$$\mathbf{u}_0 = \left(1 - E\left[\epsilon t_1(\mathbf{X})t_3(\mathbf{X})\kappa(\mathbf{X}, Y) \mid D = 0\right]\right)^{-1} E\left[\epsilon t_1(\mathbf{X})\mathbf{t}_2(\mathbf{X})\kappa(\mathbf{X}, Y) \mid D = 0\right]. \tag{S.5}$$

Combining the above results, we have obtained the analytic form of $\mathbf{S}_{\text{eff}} = \mathbf{S} - \mathbf{g} - E(\mathbf{S} - \mathbf{g} \mid D = d)$, where $\mathbf{g}$ is given in (S.1), $\mathbf{a}(\mathbf{x})$ is given in (S.4), $\mathbf{v}_0, \mathbf{v}_1$ are given in (S.2) $\mathbf{u}_0$ is given in (S.5) and the functions $t_1, \mathbf{t}_2, t_3$ are given in (S.3).

In forming the estimating equation $\sum_{i=1}^N \mathbf{S}_{\text{eff}} = 0$, we will have $\sum_{i=1}^N [\mathbf{S}(\mathbf{X}_i, Y_i, D_i) - \mathbf{g}\{Y_i - m(\mathbf{X}_i, \boldsymbol{\beta}), \mathbf{X}_i\}] - N_0 E(\mathbf{S} - \mathbf{g} \mid D = 0) - N_1 E(\mathbf{S} - \mathbf{g} \mid D = 1) = \mathbf{0}$. Using (S.1), we obtain

$$\begin{aligned}E(\mathbf{S} - \mathbf{g} \mid D = 0) &= E(\mathbf{S} \mid D = 0) - E\{E(\mathbf{S} \mid \epsilon, \mathbf{X}) \mid D = 0\} + E\{\epsilon\mathbf{a}(\mathbf{X})\kappa(\mathbf{X}, Y) \mid D = 0\}\\ &\quad + \mathbf{v}_0 E\{f_{D|X,Y}(0, \mathbf{X}, Y) \mid D = 0\} + \mathbf{v}_1 E\{f_{D|X,Y}(1, \mathbf{X}, Y) \mid D = 0\}\\ &= \mathbf{c}_0 + \mathbf{u}_0 + \mathbf{v}_0(1 - b_0) + \mathbf{v}_1 b_0\end{aligned}$$

and

$$\begin{aligned}E(\mathbf{S} - \mathbf{g} \mid D = 1) &= E(\mathbf{S} \mid D = 1) - E\{E(\mathbf{S} \mid \epsilon, \mathbf{X}) \mid D = 1\} + E\{\epsilon\mathbf{a}(\mathbf{X})\kappa(\mathbf{X}, Y) \mid D = 1\}\\ &\quad + \mathbf{v}_0 E\{f_{D|X,Y}(0, \mathbf{X}, Y) \mid D = 1\} + \mathbf{v}_1 E\{f_{D|X,Y}(1, \mathbf{X}, Y) \mid D = 1\}\\ &= \mathbf{c}_1 + \mathbf{u}_1 + \mathbf{v}_0 b_1 + \mathbf{v}_1(1 - b_1),\end{aligned}$$

hence

$$\begin{aligned}&N_0 E(\mathbf{S} - \mathbf{g} \mid D = 0) + N_1 E(\mathbf{S} - \mathbf{g} \mid D = 1)\\ &\quad N_0\{\mathbf{c}_0 + \mathbf{u}_0 + \mathbf{v}_0(1 - b_0) + \mathbf{v}_1 b_0\} + N_1\{\mathbf{c}_1 + \mathbf{u}_1 + \mathbf{v}_0 b_1 + \mathbf{v}_1(1 - b_1)\}\\ &\quad (N_0\mathbf{c}_0 + N_1\mathbf{c}_1) + (N_0\mathbf{u}_0 + N_1\mathbf{u}_1) + (N_0\mathbf{v}_0 + N_1\mathbf{v}_1) + (\mathbf{v}_1 - \mathbf{v}_0)(N_0\mathbf{b}_0 - N_1\mathbf{b}_1).\end{aligned}$$

Thus, the estimating equation simplifies to $\sum_{i=1}^N [\mathbf{S}(\mathbf{X}_i, Y_i, D_i) - \mathbf{g}\{Y_i - m(\mathbf{X}_i, \boldsymbol{\beta}), \mathbf{X}_i\}] - N_0\mathbf{v}_0 - N_1\mathbf{v}_1 = \mathbf{0}$.

S.2 Verification of Assumptions 1-2 in the Linear Model as in Remark 2

Let $\mathcal{D}^c = [-K, K]$ for a sufficiently large K. We have assumed that the conditional distribution of ϵ satisfies the property that $\mathrm{pr}(|\epsilon| > K|\mathbf{X} = \mathbf{x}) \to 0$ as $K \to \infty$ uniformly in $\mathbf{x}$, and hence $\mathrm{pr}(\epsilon \in \mathcal{D})$ can be made arbitrarily small uniformly in $\mathbf{x}$.

Without loss of generality, assume the first component of $\boldsymbol{\alpha}_1 + \boldsymbol{\beta}_1\alpha_2$ is not zero. We consider three situations, 1. $\widetilde{\beta}_c - \beta_c \neq 0$, 2. $\widetilde{\beta}_c - \beta_c = 0$ and the first component of $\widetilde{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}_1$ is not zero and 3) $\widetilde{\beta}_c - \beta_c = 0$ and the first component of $\widetilde{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}_1$ is zero. Note that ϵ has conditional mean zero. Thus for most common density functions with bounded variance function, this requirement is satisfied.

Case 1 and Case 2: If the first component of $\boldsymbol{\alpha}_1 + \boldsymbol{\beta}_1\alpha_2$ is positive, let $\mathbf{c}_1 = (\infty, \mathbf{0}^{\mathrm{T}})^{\mathrm{T}}$, otherwise, let $\mathbf{c}_1 = (-\infty, \mathbf{0}^{\mathrm{T}})^{\mathrm{T}}$. Let $c_2 = 0$, $\mathcal{D}^c = [-K, K]$ for a sufficiently large K. Then

$$\begin{aligned}
&\sup_{\epsilon\in\mathcal{D}^c}\lim_{\mathbf{x}\to\mathbf{c}_1} |(1+\exp[\alpha_c + u\{\mathbf{x}, m(\mathbf{x},\boldsymbol{\beta}) + \epsilon, \boldsymbol{\alpha}_1, \alpha_2\}])^{-1} - c_2| \\
= &\sup_{-K\le\epsilon\le K}\lim_{\mathbf{x}\to\mathbf{c}_1} |[1+\exp\{\alpha_c + \mathbf{x}^{\mathrm{T}}(\boldsymbol{\alpha}_1 + \boldsymbol{\beta}_1\alpha_2) + \beta_c\alpha_2 + \epsilon\alpha_2\}]^{-1}| \\
= &\ 0.
\end{aligned}$$

Finally, since $\mathcal{D} = (-\infty, -K) \cup (K, \infty)$, we have that every element $u \in \mathcal{D}$ satisfies $u > 1$ as long as $K > 1$. We have thus verified Assumption 1.

We have $c(\boldsymbol{\beta}, \widetilde{\boldsymbol{\beta}}) = \mathbf{c}_1^{\mathrm{T}}(\widetilde{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}_1) + (\widetilde{\boldsymbol{\beta}}_c - \boldsymbol{\beta}_c)$, which is not zero if $\widetilde{\boldsymbol{\beta}}_c - \boldsymbol{\beta}_c \neq 0$ (case 1), and is also not zero if the first component of $\widetilde{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}_1$ is not zero (case 2). Thus, Assumption 2 also holds.

Case 3: Since $\widetilde{\boldsymbol{\beta}} \neq \boldsymbol{\beta}$, without loss of generality, we assume the second component of $\widetilde{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}_1$ is not zero in this case. We select $\mathbf{c}_1$ as follows. If the first component of $\boldsymbol{\alpha}_1 + \boldsymbol{\beta}_1\alpha_2$ is positive and the second component of $\boldsymbol{\alpha}_1 + \boldsymbol{\beta}_1\alpha_2$ is non-negative, let $\mathbf{c}_1 = (\infty, \infty, \mathbf{0}^{\mathrm{T}})^{\mathrm{T}}$. If the first component of $\boldsymbol{\alpha}_1 + \boldsymbol{\beta}_1\alpha_2$ is positive and the second component of $\boldsymbol{\alpha}_1 + \boldsymbol{\beta}_1\alpha_2$ is negative, let $\mathbf{c}_1 = (\infty, -\infty, \mathbf{0}^{\mathrm{T}})^{\mathrm{T}}$. If the first component of $\boldsymbol{\alpha}_1 + \boldsymbol{\beta}_1\alpha_2$ is negative and the second component of $\boldsymbol{\alpha}_1 + \boldsymbol{\beta}_1\alpha_2$ is non-negative, let $\mathbf{c}_1 = (-\infty, \infty, \mathbf{0}^{\mathrm{T}})^{\mathrm{T}}$. If the first component of $\boldsymbol{\alpha}_1 + \boldsymbol{\beta}_1\alpha_2$ is negative and the second component of $\boldsymbol{\alpha}_1 + \boldsymbol{\beta}_1\alpha_2$ is negative, let $\mathbf{c}_1 = (-\infty, -\infty, \mathbf{0}^{\mathrm{T}})^{\mathrm{T}}$. The selection of $c_2, \mathcal{D}, K$ remains the same as in Cases 1 and 2. We an see the same arguments lead to the verification of Assumption 1. In addition, $c(\boldsymbol{\beta}, \widetilde{\boldsymbol{\beta}}) = \mathbf{c}_1^{\mathrm{T}}(\widetilde{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}_1)$, which is either ∞ or $-\infty$, and is thus not zero. Thus, Assumption 2 also holds.

S.3 Verification of Nonidentifiability in the Special Case in Remark 2

Here, we provide the details of the proof of the nonidentifiability result in Remark 2, which happens when when $m(\mathbf{x}, \boldsymbol{\beta}) = \mathbf{x}^{\mathrm{T}}\boldsymbol{\beta}_1 + \beta_c$ and $u(\mathbf{x}, y, \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_c) = \mathbf{x}^{\mathrm{T}}\boldsymbol{\alpha}_1 + y\alpha_2 + \alpha_c$, and $\boldsymbol{\alpha}_1 + \boldsymbol{\beta}_1\alpha_2 = \mathbf{0}$. We first exclude a special case when $\alpha_2 = 0$. This special case implies $\boldsymbol{\alpha}_1 = \mathbf{0}$ and $\alpha_2 = 0$, hence the case-control sampling is in fact random sampling. Thus, in the following, we assume $\alpha_2 \neq 0$, We point out that

$$m_1 \equiv \int \epsilon \frac{1+\exp(\widetilde{\alpha}_c + \beta_c\alpha_2 + \epsilon\alpha_2)}{1+\exp(\alpha_c + \beta_c\alpha_2 + \epsilon\alpha_2)} \eta_2(\epsilon) d\mu(\epsilon) \neq 0.$$

This is because we can use the mean value theorem for integration to obtain

$$\begin{aligned}
&\int \epsilon \frac{1+\exp(\widetilde{\alpha}_c + \beta_c\alpha_2 + \epsilon\alpha_2)}{1+\exp(\alpha_c + \beta_c\alpha_2 + \epsilon\alpha_2)} \eta_2(\epsilon) d\mu(\epsilon) \\
=\;& \int_{-\infty}^{0} \epsilon \frac{1+\exp(\widetilde{\alpha}_c + \beta_c\alpha_2 + \epsilon\alpha_2)}{1+\exp(\alpha_c + \beta_c\alpha_2 + \epsilon\alpha_2)} \eta_2(\epsilon) d\mu(\epsilon) + \int_{0}^{\infty} \epsilon \frac{1+\exp(\widetilde{\alpha}_c + \beta_c\alpha_2 + \epsilon\alpha_2)}{1+\exp(\alpha_c + \beta_c\alpha_2 + \epsilon\alpha_2)} \eta_2(\epsilon) d\mu(\epsilon) \\
=\;& \frac{1+\exp(\widetilde{\alpha}_c + \beta_c\alpha_2 - k_1\alpha_2)}{1+\exp(\alpha_c + \beta_c\alpha_2 - k_1\alpha_2)} \int_{-\infty}^{0} \epsilon\eta_2(\epsilon) d\mu(\epsilon) + \frac{1+\exp(\widetilde{\alpha}_c + \beta_c\alpha_2 + k_2\alpha_2)}{1+\exp(\alpha_c + \beta_c\alpha_2 + k_2\alpha_2)} \int_{0}^{\infty} \epsilon\eta_2(\epsilon) d\mu(\epsilon) \\
=\;& \left\{ \frac{1+\exp(\widetilde{\alpha}_c + \beta_c\alpha_2 + k_2\alpha_2)}{1+\exp(\alpha_c + \beta_c\alpha_2 + k_2\alpha_2)} - \frac{1+\exp(\widetilde{\alpha}_c + \beta_c\alpha_2 - k_1\alpha_2)}{1+\exp(\alpha_c + \beta_c\alpha_2 - k_1\alpha_2)} \right\} \int_{0}^{\infty} \epsilon\eta_2(\epsilon) d\mu(\epsilon) \\
=\;& \frac{\{\exp(\widetilde{\alpha}_c) - \exp(\alpha_c)\}\exp(\beta_c\alpha_2)\{\exp(k_2\alpha_2) - \exp(-k_1\alpha_2)\}}{\{1+\exp(\alpha_c + \beta_c\alpha_2 + k_2\alpha_2)\}\{1+\exp(\alpha_c + \beta_c\alpha_2 - k_1\alpha_2)\}} \int_{0}^{\infty} \epsilon\eta_2(\epsilon) d\mu(\epsilon) \neq 0,
\end{aligned}$$

where k_1, k_2 are positive constants.

Following the notation in the proof of Proposition 1, we define $\eta_2(\epsilon, \mathbf{x}) = \eta_2(\epsilon)$,

$$\begin{aligned}
\widetilde{\eta}_1(\mathbf{x}) &= \frac{\widetilde{\pi}_0}{\pi_0}\eta_1(x) \int \frac{1+\exp(\widetilde{\alpha}_c + \beta_c\alpha_2 + \epsilon\alpha_2)}{1+\exp(\alpha_c + \beta_c\alpha_2 + \epsilon\alpha_2)} \eta_2(\epsilon) d\mu(\epsilon); \\
\widetilde{\beta}_c &= \beta_c + \left\{ \int \frac{1+\exp(\widetilde{\alpha}_c + \beta_c\alpha_2 + \epsilon\alpha_2)}{1+\exp(\alpha_c + \beta_c\alpha_2 + \epsilon\alpha_2)} \eta_2(\epsilon) d\mu(\epsilon) \right\}^{-1} m_1; \\
\widetilde{\eta}_2(\epsilon) &= c_0 \frac{1+\exp(\widetilde{\alpha}_c + \widetilde{\beta}_c\alpha_2 + \epsilon\alpha_2)}{1+\exp(\alpha_c + \widetilde{\beta}_c\alpha_2 + \epsilon\alpha_2)} \eta_2(\epsilon + \widetilde{\beta}_c - \beta_c),
\end{aligned}$$

where

$$c_0^{-1} = \int \frac{1+\exp(\widetilde{\alpha}_c + \widetilde{\beta}_c\alpha_2 + \epsilon\alpha_2)}{1+\exp(\alpha_c + \widetilde{\beta}_c\alpha_2 + \epsilon\alpha_2)} \eta_2(\epsilon + \widetilde{\beta}_c - \beta_c) d\mu(\epsilon).$$

Obviously $\int \widetilde{\eta}_2(\epsilon)d\mu(\epsilon) = 1$. We can easily verify that

$$\begin{aligned}\int \widetilde{\eta}_1(\mathbf{x})d\mu(\mathbf{x}) &= \frac{\widetilde{\pi}_0}{\pi_0}\int \eta_1(x)\frac{1+\exp(\widetilde{\alpha}_c+\beta_c\alpha_2+\epsilon\alpha_2)}{1+\exp(\alpha_c+\beta_c\alpha_2+\epsilon\alpha_2)}\eta_2(\epsilon)d\mu(\epsilon)d\mu(\mathbf{x}) \\ &= \frac{\widetilde{\pi}_0}{\pi_0}\{\pi_0+\exp(\widetilde{\alpha}_c-\alpha_c)\pi_1\} \\ &= \widetilde{\pi}_0+\frac{\widetilde{\pi}_0\exp(\widetilde{\alpha}_c)\pi_1}{\pi_0\exp(\alpha_c)} = \widetilde{\pi}_0+\widetilde{\pi}_1 = 1,\end{aligned}$$

using the intermediate results in the proof of Proposition 1. We can also obtain

$$\begin{aligned}\int \epsilon\widetilde{\eta}_2(\epsilon)d\mu(\epsilon) &= c_0\int \epsilon\frac{1+\exp(\widetilde{\alpha}_c+\widetilde{\beta}_c\alpha_2+\epsilon\alpha_2)}{1+\exp(\alpha_c+\widetilde{\beta}_c\alpha_2+\epsilon\alpha_2)}\eta_2(\epsilon+\widetilde{\beta}_c-\beta_c)d\mu(\epsilon) \\ &= c_0\int (t+\beta_c-\widetilde{\beta}_c)\frac{1+\exp(\widetilde{\alpha}_c+\beta_c\alpha_2+t\alpha_2)}{1+\exp(\alpha_c+\beta_c\alpha_2+t\alpha_2)}\eta_2(t)d\mu(\epsilon) = 0.\end{aligned}$$

Now we can easily verify that the two parameter sets $\{\alpha_c, \boldsymbol{\alpha}_1, \alpha_2, \boldsymbol{\beta}_1, \beta_c, \eta_1(\mathbf{x}), \eta_2(\epsilon)\}$ and $\{\widetilde{\alpha}_c, \boldsymbol{\alpha}_1, \alpha_2, \boldsymbol{\beta}_1, \widetilde{\beta}_c, \widetilde{\eta}_1(\mathbf{x}), \widetilde{\eta}_2(\epsilon)\}$ satisfy (A.1), hence the problem is not identifiable.

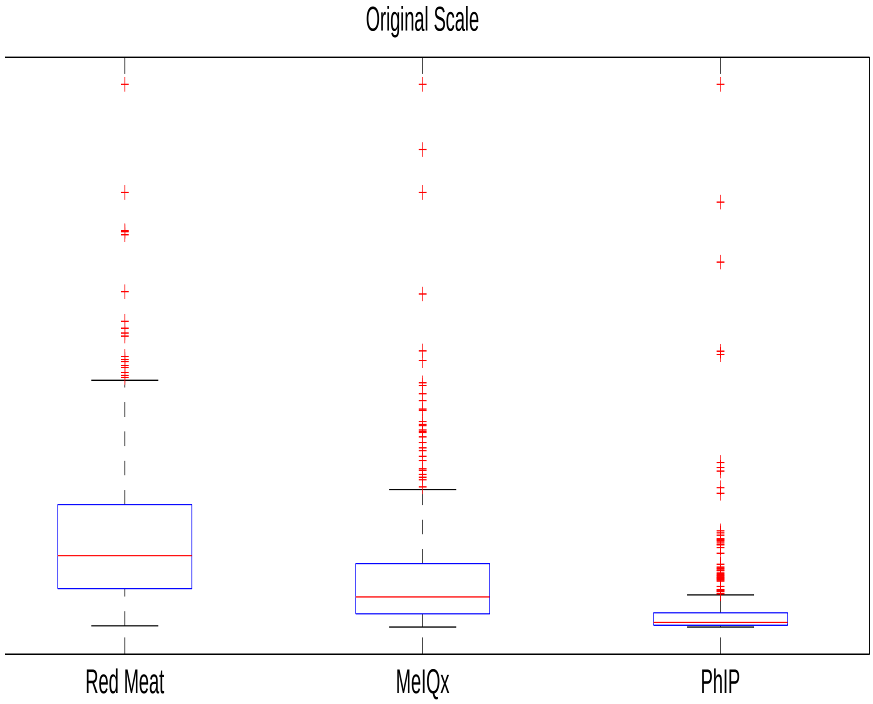

Figure S.1: Boxplots of the variables in the original data scale among the controls. Each variable has been normalized to have maximum value 1.0. This and Figure S.2 indicate a need for data transformation.

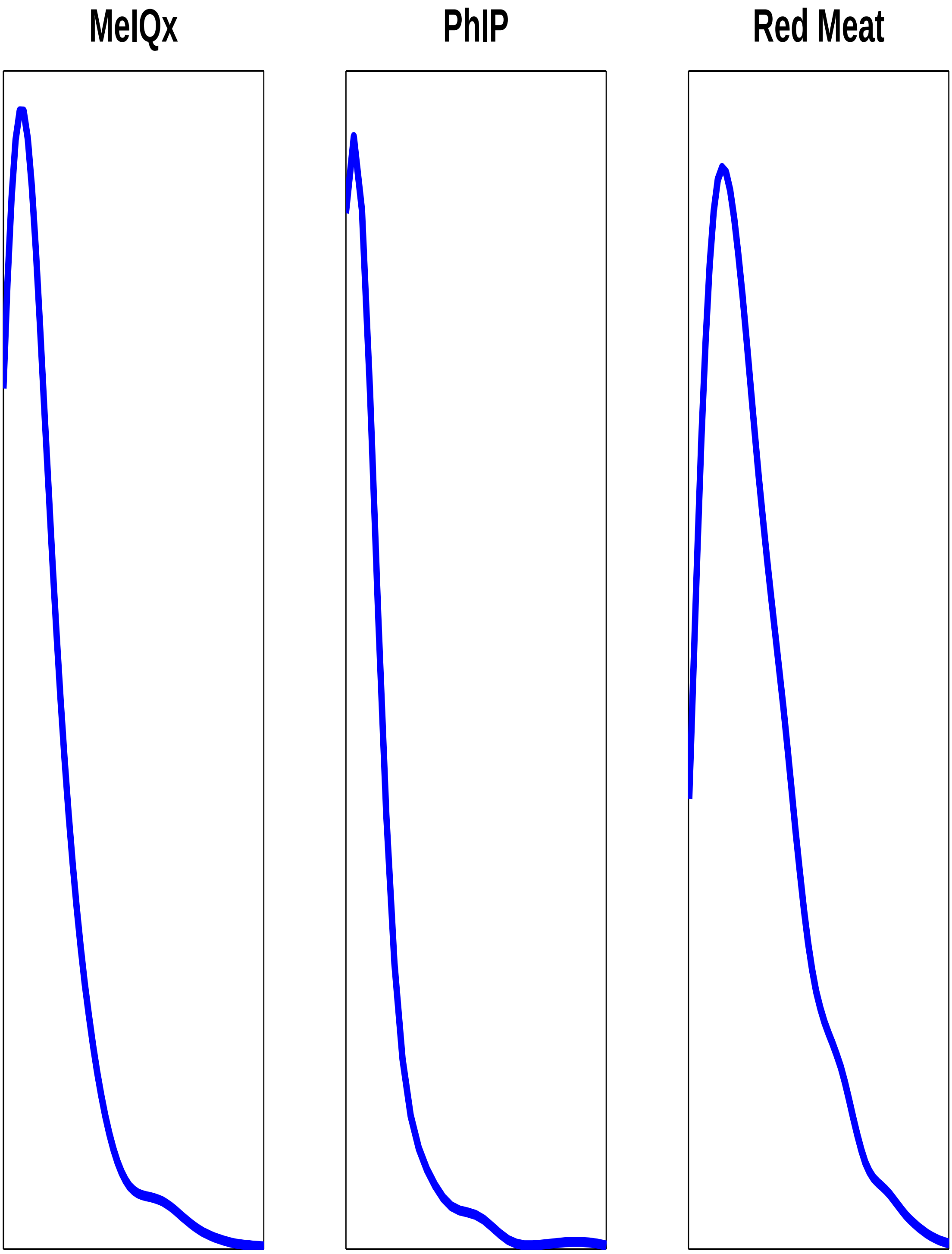

Figure S.2: Kernel density estimates in the original data scale, among the controls. This and Figure S.1 indicate a need for data transformation.

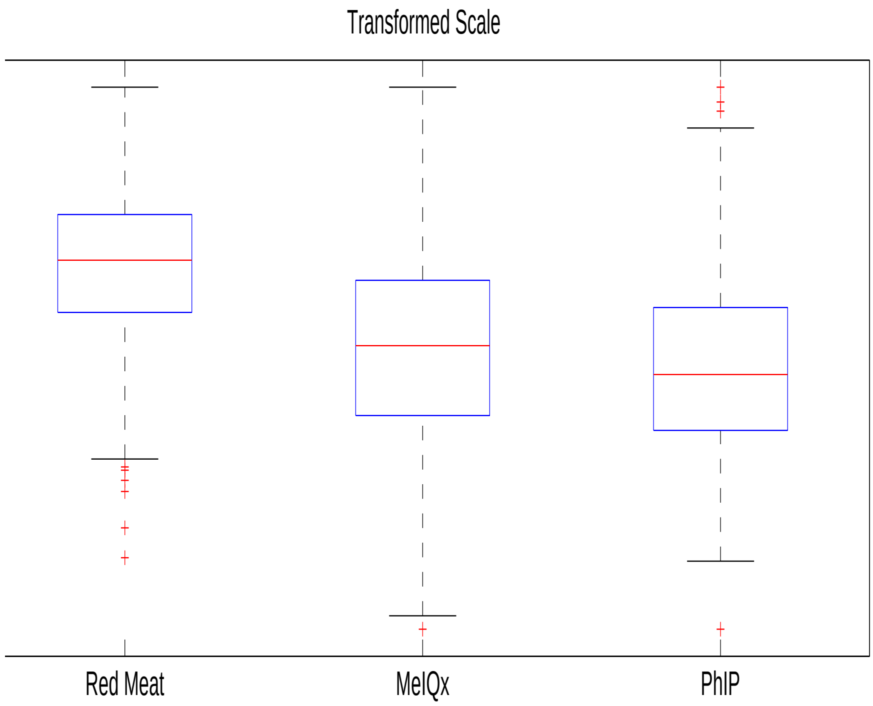

Figure S.3: Boxplots of the variables in the transformed data scale, among the controls. Each variable has been normalized to have maximum value 1.0. Contrast with the boxplots in the original data scale in Figure S.1.

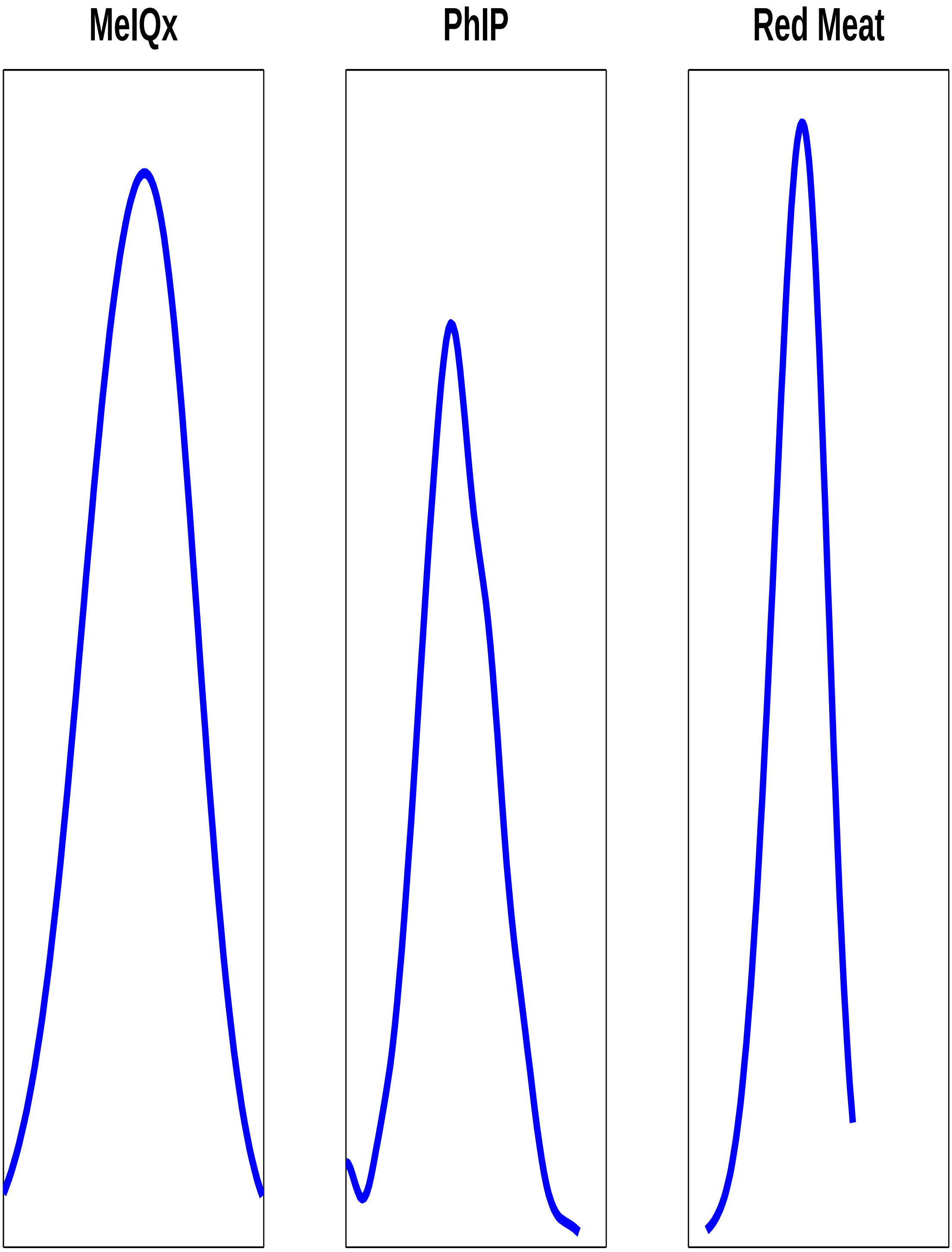

Figure S.4: Kernel density estimates in the transformed data scale, among the controls. Contrast with the boxplots in the original data scale in Figure S.2

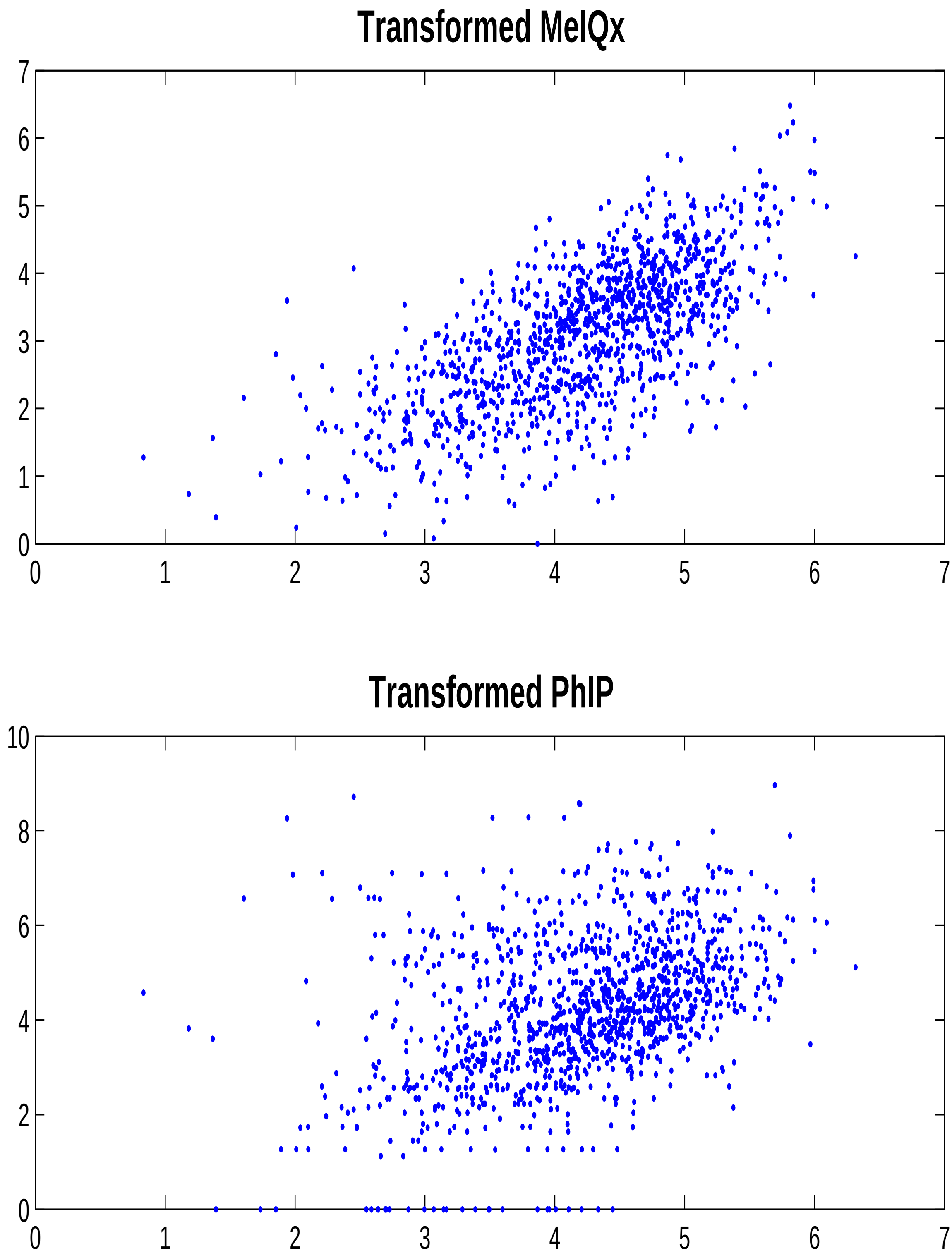

Figure S.5: Scatterplots of transformed MeIQX and PhIP against transformed Red Meat, among the controls.