Supplement to "Sufficient Dimension Reduction for Feasible and Robust Estimation of Average Causal Effect"

Trinetri Ghosh, Yanyuan Ma and Xavier de Luna

Pennsylvania State University, Pennsylvania State University and Umeå University

S1 Proof of Properties of IPW

$$
\begin{aligned}
& \widehat{E}(Y_1) \qquad\qquad (S1.1)\\
= \;& n^{-1}\sum_{i=1}^{n}\frac{t_i y_i[1+\exp\{\widehat{\eta}(\widehat{\boldsymbol{\alpha}}^{\mathrm{T}}\mathbf{x}_i)\}]}{\exp\{\widehat{\eta}(\widehat{\boldsymbol{\alpha}}^{\mathrm{T}}\mathbf{x}_i)\}}\\
= \;& n^{-1}\sum_{i=1}^{n} t_i y_{1i}\left[\exp\{-\widehat{\eta}(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_i)\}+1\right]\\
& +\left\{\frac{\partial}{\partial \mathrm{vecl}(\boldsymbol{\alpha})^{\mathrm{T}}}\left(n^{-1}\sum_{i=1}^{n} t_i y_{1i}\left[\exp\{-\eta(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_i)\}+1\right]\right)+o_p(1)\right\}\mathrm{vecl}(\widehat{\boldsymbol{\alpha}}-\boldsymbol{\alpha})\\
= \;& n^{-1}\sum_{i=1}^{n} t_i y_{1i}\left[\exp\{-\widehat{\eta}(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_i)\}+1\right]\\
& +\left\{E\left(T_i Y_{1i}\mathrm{vec}\left[-\exp\{-\eta(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{X}_i)\}\mathbf{X}_{Li}\boldsymbol{\eta}'(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{X}_i)^{\mathrm{T}}\right]\right)\right\}^{\mathrm{T}}\mathrm{vecl}(\widehat{\boldsymbol{\alpha}}-\boldsymbol{\alpha})\\
& +o_p(n^{-1/2}).
\end{aligned}
$$

Inserting (2.13) from the paper, we have that

$$
\begin{aligned}
& \left\{E\left(T_i Y_{1i} \mathrm{vec}\left[-\exp\{-\eta(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{X}_i)\}\mathbf{X}_{Li}\boldsymbol{\eta}'(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{X}_i)^{\mathrm{T}}\right]\right)\right\}^{\mathrm{T}} \mathrm{vecl}(\widehat{\boldsymbol{\alpha}}-\boldsymbol{\alpha}) \\
= \quad & n^{-1}\sum_{i=1}^{n}\left(E\left[\frac{m_{1i}(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{X}_i)}{1+\exp\{\eta(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{X}_i)\}}\mathrm{vec}\{\mathbf{X}_{Li}\boldsymbol{\eta}'(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{X}_i)^{\mathrm{T}}\}\right]\right)^{\mathrm{T}}\mathbf{B} \\
& \times(t_i-p_i)\mathrm{vec}[\{\mathbf{x}_{Li}-E(\mathbf{X}_{Li}\mid\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_i)\}\boldsymbol{\eta}'(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_i)^{\mathrm{T}}]+o_p(n^{-1/2}).
\end{aligned}
$$

In addition, using (S1.1) and Condition C2 and C3,

$$
\begin{aligned}
& n^{-1}\sum_{i=1}^{n} t_i y_{1i}\left[\exp\{-\widehat{\eta}(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_i)\}+1\right] \\
= \quad & n^{-1}\sum_{j=1}^{n} t_j y_{1j}\frac{n^{-1}\sum_{i=1}^{n}K_h(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_i-\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_j)}{n^{-1}\sum_{i=1}^{n}t_iK_h(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_i-\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_j)} \\
= \quad & n^{-1}\sum_{j=1}^{n} t_j y_{1j}\left[\frac{f(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_j)}{p_j f(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_j)}+\frac{n^{-1}\sum_{i=1}^{n}K_h(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_i-\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_j)-f(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_j)}{p_j f(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_j)}\right. \\
& \left.-\frac{f(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_j)\{n^{-1}\sum_{i=1}^{n}t_iK_h(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_i-\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_j)-p_jf(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_j)\}}{\{p_jf(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_j)\}^2}\right] \\
& +O_p\{h^{2m}+(nh^d)^{-1}\} \\
= \quad & n^{-1}\sum_{j=1}^{n}\frac{t_j y_{1j}}{p_j}\left\{1+\frac{n^{-1}\sum_{i=1}^{n}K_h(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_i-\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_j)}{f(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_j)}\right. \\
& \left.-\frac{n^{-1}\sum_{i=1}^{n}t_iK_h(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_i-\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_j)}{p_jf(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_j)}\right\}+o_p(n^{-1/2}) \\
= \quad & n^{-1}\sum_{j=1}^{n}\frac{t_j y_{1j}}{p_j}+n^{-2}\sum_{j=1}^{n}\sum_{i=1}^{n}\frac{t_jy_{1j}(p_j-t_i)K_h(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_i-\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_j)}{p_j^2f(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_j)}+o_p(n^{-1/2})
\end{aligned}
$$

$$
\begin{aligned}
&= n^{-1}\sum_{j=1}^{n}\frac{t_j y_{1j}}{p_j} + n^{-1}\sum_{j=1}^{n} E\left\{\frac{t_j y_{1j}(p_j - T_i)K_h(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{X}_i - \boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_j)}{p_j^2 f(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_j)}\right\} \\
&\quad + n^{-1}\sum_{i=1}^{n} E\left\{\frac{T_j Y_{1j}(P_j - t_i)K_h(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_i - \boldsymbol{\alpha}^{\mathrm{T}}\mathbf{X}_j)}{P_j^2 f(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{X}_j)}\right\} \\
&\quad - n^{-1}\sum_{i=1}^{n} E\left\{\frac{T_j Y_{1j}(P_j - T_i)K_h(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{X}_i - \boldsymbol{\alpha}^{\mathrm{T}}\mathbf{X}_j)}{P_j^2 f(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{X}_j)}\right\} + o_p(n^{-1/2}) \\
&= n^{-1}\sum_{i=1}^{n}\frac{t_i y_{1i}}{p_i} + n^{-1}\sum_{i=1}^{n} E\left\{m_1(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{X}_i) \mid \boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_i\right\}\left(1 - \frac{t_i}{p_i}\right) + o_p(n^{-1/2}).
\end{aligned}
$$

We thus obtain

$$
\begin{aligned}
\sqrt{n}\widehat{E}(Y_1) &= n^{-1/2}\sum_{i=1}^{n}\frac{t_i y_{1i}}{p_i} + n^{-1/2}\sum_{i=1}^{n} E\left\{m_1(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{X}_i) \mid \boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_i\right\}\left(1 - \frac{t_i}{p_i}\right) \\
&\quad + n^{-1/2}\sum_{i=1}^{n}\left(E\left[\frac{m_{1i}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_i)}{1+\exp\{\eta(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{X}_i)\}}\mathrm{vec}\{\mathbf{X}_{Li}\boldsymbol{\eta}'(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{X}_i)^{\mathrm{T}}\}\right]\right)^{\mathrm{T}}\mathbf{B} \\
&\quad \times (t_i - p_i)\mathrm{vec}[\{\mathbf{x}_{Li} - E(\mathbf{X}_{Li} \mid \boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_i)\}\boldsymbol{\eta}'(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_i)^{\mathrm{T}}] \\
&\quad + o_p(1). \qquad \text{(S1.2)}
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
\widehat{E}(Y_0) &= n^{-1}\sum_{i=1}^{n}(1-t_i)y_i[1+\exp\{\widehat{\eta}(\widehat{\boldsymbol{\alpha}}^{\mathrm{T}}\mathbf{x}_i)\}] \\
&= n^{-1}\sum_{i=1}^{n}(1-t_i)y_{0i}\left[\exp\{\widehat{\eta}(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_i)\}+1\right] \\
&\quad +\left\{\frac{\partial}{\partial \mathrm{vecl}(\boldsymbol{\alpha})^{\mathrm{T}}}\left(n^{-1}\sum_{i=1}^{n}(1-t_i)y_{0i}\left[\exp\{\eta(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_i)\}+1\right]\right)\right. \\
&\quad \left.+o_p(1)\right\}\mathrm{vecl}(\widehat{\boldsymbol{\alpha}}-\boldsymbol{\alpha}) \\
&= n^{-1}\sum_{i=1}^{n}(1-t_i)y_{0i}\left[\exp\{\widehat{\eta}(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_i)\}+1\right] \\
&\quad +\left\{E\left((1-T_i)Y_{0i}\mathrm{vec}\left[\exp\{\eta(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{X}_i)\}\mathbf{X}_{Li}\boldsymbol{\eta}'(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{X}_i)^{\mathrm{T}}\right]\right)\right\}^{\mathrm{T}}\mathrm{vecl}(\widehat{\boldsymbol{\alpha}}-\boldsymbol{\alpha}) \\
&\quad +o_p(n^{-1/2}).
\end{aligned}
$$

We further have that

$$
\begin{aligned}
&\left\{E\left((1-T_i)Y_{0i}\mathrm{vec}\left[\exp\{\eta(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{X}_i)\}\mathbf{X}_{Li}\boldsymbol{\eta}'(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{X}_i)^{\mathrm{T}}\right]\right)\right\}^{\mathrm{T}}\mathrm{vecl}(\widehat{\boldsymbol{\alpha}}-\boldsymbol{\alpha}) \\
=\ & -n^{-1}\sum_{i=1}^{n}\left(E\left[\frac{m_{0i}(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i)\exp\{\eta(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{X}_i)\}}{1+\exp\{\eta(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{X}_i)\}}\mathrm{vec}\{\mathbf{X}_{Li}\boldsymbol{\eta}'(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{X}_i)^{\mathrm{T}}\}\right]\right)^{\mathrm{T}}\mathbf{B} \\
&\times(t_i-p_i)\mathrm{vec}[\{\mathbf{x}_{Li}-E(\mathbf{X}_{Li}\mid\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_i)\}\boldsymbol{\eta}'(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_i)^{\mathrm{T}}]+o_p(n^{-1/2}).
\end{aligned}
$$

In addition, using (S1.1) and Conditon C2 and C3,

$$
\begin{aligned}
& n^{-1}\sum_{i=1}^{n}(1-t_i)y_{0i}\left[\exp\{\widehat{\eta}(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_i)\}+1\right] \\
= \; & n^{-1}\sum_{j=1}^{n}(1-t_j)y_{0j}\frac{n^{-1}\sum_{i=1}^{n}K_h(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_i-\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_j)}{n^{-1}\sum_{i=1}^{n}(1-t_i)K_h(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_i-\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_j)} \\
= \; & n^{-1}\sum_{j=1}^{n}(1-t_j)y_{0j}\left[\frac{f(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_j)}{(1-p_j)f(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_j)}+\frac{n^{-1}\sum_{i=1}^{n}K_h(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_i-\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_j)-f(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_j)}{(1-p_j)f(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_j)}\right. \\
& \left.-\frac{f(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_j)\{n^{-1}\sum_{i=1}^{n}(1-t_i)K_h(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_i-\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_j)-(1-p_j)f(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_j)\}}{\{(1-p_j)f(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_j)\}^2}\right] \\
& +O_p\{h^{2m}+(nh)^{-1}\} \\
= \; & n^{-1}\sum_{j=1}^{n}\frac{(1-t_j)y_{0j}}{1-p_j}\left\{1+\frac{n^{-1}\sum_{i=1}^{n}K_h(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_i-\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_j)}{f(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_j)}\right. \\
& \left.-\frac{n^{-1}\sum_{i=1}^{n}(1-t_i)K_h(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_i-\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_j)}{(1-p_j)f(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_j)}\right\}+o_p(n^{-1/2}) \\
= \; & n^{-1}\sum_{j=1}^{n}\frac{(1-t_j)y_{0j}}{1-p_j}-n^{-2}\sum_{j=1}^{n}\sum_{i=1}^{n}\frac{(1-t_j)y_{0j}(p_j-t_i)K_h(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_i-\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_j)}{(1-p_j)^2f(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_j)}+o_p(n^{-1/2}) \\
= \; & n^{-1}\sum_{j=1}^{n}\frac{(1-t_j)y_{0j}}{1-p_j}-n^{-1}\sum_{j=1}^{n}E\left\{\frac{(1-t_j)y_{0j}(p_j-T_i)K_h(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{X}_i-\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_j)}{(1-p_j)^2f(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_j)}\right\} \\
& -n^{-1}\sum_{i=1}^{n}E\left\{\frac{(1-T_j)Y_{0j}(P_j-t_i)K_h(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_i-\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{X}_j)}{(1-P_j)^2f(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{X}_j)}\right\} \\
& +E\left\{\frac{(1-T_j)Y_{0j}(P_j-T_i)K_h(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{X}_i-\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{X}_j)}{(1-P_j)^2f(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{X}_j)}\right\}+o_p(n^{-1/2}) \\
= \; & n^{-1}\sum_{i=1}^{n}\frac{(1-t_i)y_{0i}}{1-p_i}+n^{-1}\sum_{i=1}^{n}E\left\{m_0(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i)\mid\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_i\right\}\left(\frac{t_i-p_i}{1-p_i}\right)+o_p(n^{-1/2}).
\end{aligned}
$$

We thus obtain

$$\begin{aligned}
\sqrt{n}\widehat{E}(Y_0) &= n^{-1/2}\sum_{i=1}^{n}\frac{(1-t_i)y_{0i}}{1-p_i} + n^{-1/2}\sum_{i=1}^{n} E\left\{m_0(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i) \mid \boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_i\right\}\left(\frac{t_i-p_i}{1-p_i}\right) \\
&\quad -n^{-1/2}\sum_{i=1}^{n}\left(E\left[\frac{m_{0i}(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i)\exp\{\eta(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{X}_i)\}}{1+\exp\{\eta(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{X}_i)\}}\mathrm{vec}\{\mathbf{X}_{Li}\boldsymbol{\eta}'(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{X}_i)^{\mathrm{T}}\}\right]\right)^{\mathrm{T}}\mathbf{B} \\
&\quad \times(t_i-p_i)\mathrm{vec}[\{\mathbf{x}_{Li}-E(\mathbf{X}_{Li}\mid\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_i)\}\boldsymbol{\eta}'(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_i)^{\mathrm{T}}] + o_p(1). \qquad \text{(S1.3)}
\end{aligned}$$

Combining the results of $\widehat{E}(Y_1)$ and $\widehat{E}(Y_0)$, we get

$$\begin{aligned}
&\sqrt{n}[\{\widehat{E}(Y_1)-\widehat{E}(Y_0)\}-\{E(Y_1)-E(Y_0)\}] \\
= &\; n^{-1/2}\sum_{i=1}^{n}\left\{\frac{t_iy_{1i}}{p_i}-E(Y_1)-\frac{(1-t_i)y_{0i}}{1-p_i}+E(Y_0)\right\} \\
&+n^{-1/2}\sum_{i=1}^{n}\left(1-\frac{t_i}{p_i}\right)E\left\{m_1(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{X}_i)\mid\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_i\right\} \\
&-n^{-1/2}\sum_{i=1}^{n}\left(\frac{t_i-p_i}{1-p_i}\right)E\left\{m_0(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i)\mid\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_i\right\} \\
&+n^{-1/2}\sum_{i=1}^{n}\left(E\left[\frac{m_{1i}(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{X}_i)+\exp\{\eta(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{X}_i)\}m_{0i}(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i)}{1+\exp\{\eta(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{X}_i)\}}\mathrm{vec}\{\mathbf{X}_{Li}\boldsymbol{\eta}'(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{X}_i)^{\mathrm{T}}\}\right]\right)^{\mathrm{T}}\mathbf{B} \\
&\times(t_i-p_i)\mathrm{vec}[\{\mathbf{x}_{Li}-E(\mathbf{X}_{Li}\mid\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_i)\}\boldsymbol{\eta}'(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_i)^{\mathrm{T}}]+o_p(1).
\end{aligned}$$

$\square$

S2 Proof of Properties of AIPW

$$
\begin{aligned}
& \sqrt{n}\{\widehat{E}(Y_1) - E(Y_1)\} \\
= \; & n^{-1/2} \sum_{i=1}^{n} \Big\{ t_i y_i [1 + \exp\{-\widehat{\eta}(\widehat{\boldsymbol{\alpha}}^{\mathrm{T}} \mathbf{x}_i)\}] - E(Y_1) \\
& + \Big(1 - t_i[1 + \exp\{-\widehat{\eta}(\widehat{\boldsymbol{\alpha}}^{\mathrm{T}} \mathbf{x}_i)\}]\Big) \widehat{m}_1(\widehat{\boldsymbol{\beta}}_1^{\mathrm{T}} \mathbf{x}_i) \Big\} \\
= \; & n^{-1/2} \sum_{i=1}^{n} \big\{ t_i y_{1i} [1 + \exp\{-\widehat{\eta}(\boldsymbol{\alpha}^{\mathrm{T}} \mathbf{x}_i)\}] - E(Y_1) \\
& + \big(1 - t_i[1 + \exp\{-\widehat{\eta}(\boldsymbol{\alpha}^{\mathrm{T}} \mathbf{x}_i)\}]\big) \widehat{m}_1(\widehat{\boldsymbol{\beta}}_1^{\mathrm{T}} \mathbf{x}_i) \Big\} \\
& + \left[n^{-1} \sum_{i=1}^{n} t_i \{ y_{1i} - \widehat{m}_1(\widehat{\boldsymbol{\beta}}_1^{\mathrm{T}} \mathbf{x}_i)\} \frac{\partial \exp\{-\widehat{\eta}(\boldsymbol{\alpha}^{\mathrm{T}} \mathbf{x}_i)\}}{\partial \mathrm{vecl}(\boldsymbol{\alpha})^{\mathrm{T}}} + o_p(1) \right] \sqrt{n} \mathrm{vecl}(\widehat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}) \\
= \; & n^{-1/2} \sum_{i=1}^{n} \Big(\{ y_{1i} - \widehat{m}_1(\widehat{\boldsymbol{\beta}}_1^{\mathrm{T}} \mathbf{x}_i)\} t_i [1 + \exp\{-\widehat{\eta}(\boldsymbol{\alpha}^{\mathrm{T}} \mathbf{x}_i)\}] + \widehat{m}_1(\widehat{\boldsymbol{\beta}}_1^{\mathrm{T}} \mathbf{x}_i) - E(Y_1) \Big) \\
& - E\left[\{ Y_{1i} - m_1(\boldsymbol{\beta}_1^{\mathrm{T}} \mathbf{X}_i)\} T_i \exp\{-\eta(\boldsymbol{\alpha}^{\mathrm{T}} \mathbf{X}_i)\} \mathrm{vec}\{\mathbf{X}_{Li} \boldsymbol{\eta}'(\boldsymbol{\alpha}^{\mathrm{T}} \mathbf{X}_i)^{\mathrm{T}}\} \right] \\
& \times \sqrt{n} \mathrm{vecl}(\widehat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}) + o_p(1) \\
= \; & n^{-1/2} \sum_{i=1}^{n} \big(\{ y_{1i} - \widehat{m}_1(\boldsymbol{\beta}_1^{\mathrm{T}} \mathbf{x}_i)\} t_i [1 + \exp\{-\widehat{\eta}(\boldsymbol{\alpha}^{\mathrm{T}} \mathbf{x}_i)\}] + \widehat{m}_1(\boldsymbol{\beta}_1^{\mathrm{T}} \mathbf{x}_i) - E(Y_1) \big) \\
& + \left[E\left\{ \frac{\partial m_1(\boldsymbol{\beta}_1^{\mathrm{T}} \mathbf{X}_i)}{\partial \mathrm{vecl}(\boldsymbol{\beta}_1)^{\mathrm{T}}} (1 - T_i[1 + \exp\{-\eta(\boldsymbol{\alpha}^{\mathrm{T}} \mathbf{X}_i)\}]) \right\} + o_p(1) \right] \\
& \times \sqrt{n} \mathrm{vecl}(\widehat{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}_1) - \mathbf{D}_1 \sqrt{n} \mathrm{vecl}(\widehat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}) + o_p(1) \\
= \; & n^{-1/2} \sum_{i=1}^{n} \big(\{ y_{1i} - \widehat{m}_1(\boldsymbol{\beta}_1^{\mathrm{T}} \mathbf{x}_i)\} \big(t_i [1 + \exp\{-\widehat{\eta}(\boldsymbol{\alpha}^{\mathrm{T}} \mathbf{x}_i)\}] - 1 \big) + y_{1i} - E(Y_1) \big) \\
& + \mathbf{C}_1 \sqrt{n} \mathrm{vecl}(\widehat{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}_1) - \mathbf{D}_1 \sqrt{n} \mathrm{vecl}(\widehat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}) + o_p(1)
\end{aligned}
$$

$$
\begin{aligned}
= \quad & n^{-1/2}\sum_{i=1}^{n}\Big\{\{y_{1i}-m_1(\boldsymbol\beta_1^{\mathrm T}\mathbf x_i)\}\left(t_i[1+\exp\{-\eta(\boldsymbol\alpha^{\mathrm T}\mathbf x_i)\}]-1\right)+y_{1i}-E(Y_1)\Big\} \\
& -\mathbf C_1\mathbf B_1 n^{-1/2}\sum_{i=1}^{n} t_i\{y_{1i}-m_1(\boldsymbol\beta_1^{\mathrm T}\mathbf x_i)\}\mathrm{vec}[\mathbf m_1'(\boldsymbol\beta_1^{\mathrm T}\mathbf x_i)\otimes\{\mathbf x_{Li}-E(\mathbf X_{Li}\mid\boldsymbol\beta_1^{\mathrm T}\mathbf x_i)\}] \\
& +\mathbf D_1\mathbf B n^{-1/2}\sum_{i=1}^{n}(t_i-p_i)\mathrm{vec}[\{\mathbf x_{Li}-E(\mathbf X_{Li}\mid\boldsymbol\alpha^{\mathrm T}\mathbf x_i)\}\boldsymbol\eta'(\boldsymbol\alpha^{\mathrm T}\mathbf x_i)^{\mathrm T}]+o_p(1) \\
= \quad & n^{-1/2}\sum_{i=1}^{n}\Big(\{y_{1i}-m_1(\boldsymbol\beta_1^{\mathrm T}\mathbf x_i)\}t_i[1+\exp\{-\eta(\boldsymbol\alpha^{\mathrm T}\mathbf x_i)\}]+m_1(\boldsymbol\beta_1^{\mathrm T}\mathbf x_i)-E(Y_1) \\
& -\mathbf C_1\mathbf B_1 t_i\{y_{1i}-m_1(\boldsymbol\beta_1^{\mathrm T}\mathbf x_i)\}\mathrm{vec}[\mathbf m_1'(\boldsymbol\beta_1^{\mathrm T}\mathbf x_i)\otimes\{\mathbf x_{Li}-E(\mathbf X_{Li}\mid\boldsymbol\beta_1^{\mathrm T}\mathbf x_i)\}] \\
& +\mathbf D_1\mathbf B(t_i-p_i)\mathrm{vec}[\{\mathbf x_{Li}-E(\mathbf X_{Li}\mid\boldsymbol\alpha^{\mathrm T}\mathbf x_i)\}\boldsymbol\eta'(\boldsymbol\alpha^{\mathrm T}\mathbf x_i)^{\mathrm T}]\Big)+o_p(1).
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
& \sqrt n\{\widehat E(Y_0)-E(Y_0)\} \\
= \quad & n^{-1/2}\sum_{i=1}^{n}\Big\{(1-t_i)y_i[1+\exp\{\widehat\eta(\widehat{\boldsymbol\alpha}^{\mathrm T}\mathbf x_i)\}]-E(Y_0) \\
& +\Big(1-(1-t_i)[1+\exp\{\widehat\eta(\widehat{\boldsymbol\alpha}^{\mathrm T}\mathbf x_i)\}]\Big)\widehat m_0(\widehat{\boldsymbol\beta}_0^{\mathrm T}\mathbf x_i)\Big\} \\
= \quad & n^{-1/2}\sum_{i=1}^{n}\big\{(1-t_i)y_{0i}[1+\exp\{\widehat\eta(\boldsymbol\alpha^{\mathrm T}\mathbf x_i)\}]-E(Y_0) \\
& +\big(1-(1-t_i)[1+\exp\{\widehat\eta(\boldsymbol\alpha^{\mathrm T}\mathbf x_i)\}]\big)\widehat m_0(\widehat{\boldsymbol\beta}_0^{\mathrm T}\mathbf x_i)\Big\} \\
& +\left[n^{-1}\sum_{i=1}^{n}(1-t_i)\{y_{0i}-\widehat m_0(\widehat{\boldsymbol\beta}_0^{\mathrm T}\mathbf x_i)\}\frac{\partial\exp\{\widehat\eta(\boldsymbol\alpha^{\mathrm T}\mathbf x_i)\}}{\partial\mathrm{vecl}(\boldsymbol\alpha)^{\mathrm T}}+o_p(1)\right] \\
& \times\sqrt n\,\mathrm{vecl}(\widehat{\boldsymbol\alpha}-\boldsymbol\alpha)
\end{aligned}
$$

$$
\begin{aligned}
&= n^{-1/2}\sum_{i=1}^{n}\Big(\{y_{0i}-\widehat{m}_0(\widehat{\boldsymbol\beta}_0^{\mathrm T}\mathbf{x}_i)\}(1-t_i)[1+\exp\{\widehat{\eta}(\boldsymbol\alpha^{\mathrm T}\mathbf{x}_i)\}]+\widehat{m}_0(\widehat{\boldsymbol\beta}_0^{\mathrm T}\mathbf{x}_i)-E(Y_0)\Big)\\
&\quad +E\left[\{Y_{0i}-m_0(\boldsymbol\beta_0^{\mathrm T}\mathbf{X}_i)\}(1-T_i)\exp\{\eta(\boldsymbol\alpha^{\mathrm T}\mathbf{X}_i)\}\mathrm{vec}\{\mathbf{X}_{Li}\boldsymbol\eta'(\boldsymbol\alpha^{\mathrm T}\mathbf{X}_i)^{\mathrm T}\}\right]\\
&\quad \times\sqrt{n}\mathrm{vecl}(\widehat{\boldsymbol\alpha}-\boldsymbol\alpha)+o_p(1)\\
&= n^{-1/2}\sum_{i=1}^{n}\big(\{y_{0i}-\widehat{m}_0(\boldsymbol\beta_0^{\mathrm T}\mathbf{x}_i)\}(1-t_i)[1+\exp\{\widehat{\eta}(\boldsymbol\alpha^{\mathrm T}\mathbf{x}_i)\}]+\widehat{m}_0(\boldsymbol\beta_0^{\mathrm T}\mathbf{x}_i)\\
&\quad -E(Y_0)\big)+\left[E\left\{\frac{\partial m_0(\boldsymbol\beta_0^{\mathrm T}\mathbf{X}_i)}{\partial\mathrm{vecl}(\boldsymbol\beta_0)^{\mathrm T}}(1-(1-T_i)[1+\exp\{\eta(\boldsymbol\alpha^{\mathrm T}\mathbf{X}_i)\}])\right\}+o_p(1)\right]\\
&\quad \times\sqrt{n}\mathrm{vecl}(\widehat{\boldsymbol\beta}_1-\boldsymbol\beta_1)+\mathbf{D}_0\sqrt{n}\mathrm{vecl}(\widehat{\boldsymbol\alpha}-\boldsymbol\alpha)+o_p(1)\\
&= n^{-1/2}\sum_{i=1}^{n}\big(\{y_{0i}-\widehat{m}_0(\boldsymbol\beta_0^{\mathrm T}\mathbf{x}_i)\}\big((1-t_i)[1+\exp\{\widehat{\eta}(\boldsymbol\alpha^{\mathrm T}\mathbf{x}_i)\}]-1\big)+y_{0i}\\
&\quad -E(Y_0)\big)+\mathbf{C}_0\sqrt{n}\mathrm{vecl}(\widehat{\boldsymbol\beta}_0-\boldsymbol\beta_0)+\mathbf{D}_0\sqrt{n}\mathrm{vecl}(\widehat{\boldsymbol\alpha}-\boldsymbol\alpha)+o_p(1)\\
&= n^{-1/2}\sum_{i=1}^{n}\big\{\{y_{0i}-m_0(\boldsymbol\beta_0^{\mathrm T}\mathbf{x}_i)\}\big((1-t_i)[1+\exp\{\eta(\boldsymbol\alpha^{\mathrm T}\mathbf{x}_i)\}]-1\big)+y_{0i}\\
&\quad -E(Y_0)\big\}-\mathbf{C}_0\mathbf{B}_0n^{-1/2}\sum_{i=1}^{n}\big((1-t_i)\{y_{0i}-m_0(\boldsymbol\beta_0^{\mathrm T}\mathbf{x}_i)\}\\
&\quad \times\mathrm{vec}[\mathbf{m}_0'(\boldsymbol\beta_0^{\mathrm T}\mathbf{x}_i)\otimes\{\mathbf{x}_{Li}-E(\mathbf{X}_{Li}\mid\boldsymbol\beta_0^{\mathrm T}\mathbf{x}_i)\}]\big)\\
&\quad -\mathbf{D}_0\mathbf{B}n^{-1/2}\sum_{i=1}^{n}(t_i-p_i)\mathrm{vec}[\{\mathbf{x}_{Li}-E(\mathbf{X}_{Li}\mid\boldsymbol\alpha^{\mathrm T}\mathbf{x}_i)\}\boldsymbol\eta'(\boldsymbol\alpha^{\mathrm T}\mathbf{x}_i)^{\mathrm T}]+o_p(1)\\
&= n^{-1/2}\sum_{i=1}^{n}\big(\{y_{0i}-m_0(\boldsymbol\beta_0^{\mathrm T}\mathbf{x}_i)\}(1-t_i)[1+\exp\{\eta(\boldsymbol\alpha^{\mathrm T}\mathbf{x}_i)\}]+m_0(\boldsymbol\beta_0^{\mathrm T}\mathbf{x}_i)\\
&\quad -E(Y_0)-\mathbf{C}_0\mathbf{B}_0(1-t_i)\{y_{0i}-m_0(\boldsymbol\beta_0^{\mathrm T}\mathbf{x}_i)\}\\
&\quad \times\mathrm{vec}[\mathbf{m}_0'(\boldsymbol\beta_0^{\mathrm T}\mathbf{x}_i)\otimes\{\mathbf{x}_{Li}-E(\mathbf{X}_{Li}\mid\boldsymbol\beta_0^{\mathrm T}\mathbf{x}_i)\}]\\
&\quad -\mathbf{D}_0\mathbf{B}(t_i-p_i)\mathrm{vec}[\{\mathbf{x}_{Li}-E(\mathbf{X}_{Li}\mid\boldsymbol\alpha^{\mathrm T}\mathbf{x}_i)\}\boldsymbol\eta'(\boldsymbol\alpha^{\mathrm T}\mathbf{x}_i)^{\mathrm T}]\big)+o_p(1).
\end{aligned}
$$

Combining the above results, we get

$$
\begin{aligned}
&\sqrt{n}[\{\widehat{E}(Y_1)-\widehat{E}(Y_0)\}-\{E(Y_1)-E(Y_0)\}]\\
= \quad & n^{-1/2}\sum_{i=1}^{n}\big(\{y_{1i}-m_1(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)\}t_i[1+\exp\{-\eta(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_i)\}]+m_1(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)-E(Y_1)\\
&-\mathbf{C}_1\mathbf{B}_1 t_i\{y_{1i}-m_1(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)\}\mathrm{vec}[\mathbf{m}_1'(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)\otimes\{\mathbf{x}_{Li}-E(\mathbf{X}_{Li}\mid\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)\}]\\
&+\mathbf{D}_1\mathbf{B}(t_i-p_i)\mathrm{vec}[\{\mathbf{x}_{Li}-E(\mathbf{X}_{Li}\mid\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_i)\}\boldsymbol{\eta}'(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_i)^{\mathrm{T}}]\\
&-\{y_{0i}-m_0(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)\}(1-t_i)[1+\exp\{\eta(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_i)\}]-m_0(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)+E(Y_0)\\
&+\mathbf{C}_0\mathbf{B}_0(1-t_i)\{y_{0i}-m_0(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)\}\mathrm{vec}[\mathbf{m}_0'(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)\otimes\{\mathbf{x}_{Li}-E(\mathbf{X}_{Li}\mid\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)\}]\\
&+\mathbf{D}_0\mathbf{B}(t_i-p_i)\mathrm{vec}[\{\mathbf{x}_{Li}-E(\mathbf{X}_{Li}\mid\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_i)\}\boldsymbol{\eta}'(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_i)^{\mathrm{T}}]\big)+o_p(1).
\end{aligned}
$$

$\square$

S3 Proof of Properties of IMP

Using similar analysis as before, we get

$$
\begin{aligned}
& \sqrt{n}\{\widehat{E}(Y_1) - E(Y_1)\} \\
= \quad & n^{-1/2} \sum_{i=1}^{n} \Big\{ t_i y_i + (1-t_i)\widehat{m}_1(\widehat{\boldsymbol{\beta}}_1^{\mathrm{T}} \mathbf{x}_i) - E(Y_1) \Big\} \\
= \quad & n^{-1/2} \sum_{i=1}^{n} \big\{ t_i y_{1i} + (1-t_i)\widehat{m}_1(\boldsymbol{\beta}_1^{\mathrm{T}} \mathbf{x}_i) - E(Y_1) \big\} \\
& + \left\{ n^{-1} \sum_{i=1}^{n} (1-t_i) \frac{\partial \widehat{m}_1(\boldsymbol{\beta}_1^{\mathrm{T}} \mathbf{x}_i)}{\partial \mathrm{vecl}(\boldsymbol{\beta}_1)^{\mathrm{T}}} + o_p(1) \right\} \sqrt{n}\mathrm{vecl}(\widehat{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}_1) \\
= \quad & n^{-1/2} \sum_{i=1}^{n} \big\{ t_i y_{1i} + (1-t_i)\widehat{m}_1(\boldsymbol{\beta}_1^{\mathrm{T}} \mathbf{x}_i) - E(Y_1) \big\} \\
& + E[(1-P_i)\mathrm{vec}\{\mathbf{X}_{Li} \mathbf{m}_1'(\boldsymbol{\beta}_1^{\mathrm{T}} \mathbf{X}_i)^{\mathrm{T}}\}]^{\mathrm{T}} \sqrt{n}\mathrm{vecl}(\widehat{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}_1) + o_p(1).
\end{aligned}
$$

We further have that

$$
\begin{aligned}
& E[(1-P_i)\mathrm{vec}\{\mathbf{X}_{Li} \mathbf{m}_1'(\boldsymbol{\beta}_1^{\mathrm{T}} \mathbf{X}_i)^{\mathrm{T}}\}]^{\mathrm{T}} \sqrt{n}\mathrm{vecl}(\widehat{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}_1) \\
= \quad & -n^{-1/2} \sum_{i=1}^{n} E[(1-P_i)\mathrm{vec}\{\mathbf{X}_{Li} \mathbf{m}_1'(\boldsymbol{\beta}_1^{\mathrm{T}} \mathbf{X}_i)^{\mathrm{T}}\}]^{\mathrm{T}} \mathbf{B}_1 \\
& \times t_i \{y_{1i} - m_1(\boldsymbol{\beta}_1^{\mathrm{T}} \mathbf{x}_i)\} \mathrm{vec}[\mathbf{m}_1'(\boldsymbol{\beta}_1^{\mathrm{T}} \mathbf{x}_i) \otimes \{\mathbf{x}_{Li} - E(\mathbf{X}_{Li} \mid \boldsymbol{\beta}_1^{\mathrm{T}} \mathbf{x}_i)\}] + o_p(1).
\end{aligned}
$$

On the other hand,

$$
\begin{aligned}
& n^{-1/2}\sum_{i=1}^{n}\left\{t_i y_{1i}+(1-t_i)\widehat{m}_1(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)-E(Y_1)\right\}\\
= & n^{-1/2}\sum_{i=1}^{n}\left\{t_i y_{1i}+(1-t_i)\frac{n^{-1}\sum_{j=1}^{n}t_j y_{1j}K_h(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_j-\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)}{n^{-1}\sum_{j=1}^{n}t_j K_h(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_j-\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)}-E(Y_1)\right\}\\
= & n^{-1/2}\sum_{i=1}^{n}\left[t_i y_{1i}+(1-t_i)\left\{\frac{E(T_iY_{1i}\mid\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)f(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)}{E(T_i\mid\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)f(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)}\right.\right.\\
& +\frac{n^{-1}\sum_{j=1}^{n}t_j y_{1j}K_h(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_j-\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)-E(T_iY_{1i}\mid\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)f(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)}{E(T_i\mid\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)f(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)}\\
& \left.-E(T_iY_{1i}\mid\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)f(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)\frac{n^{-1}\sum_{j=1}^{n}t_j K_h(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_j-\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)-E(T_i\mid\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)f(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)}{\{E(T_i\mid\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)f(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)\}^2}\right\}\\
& \left.-E(Y_1)\right]+O_p(n^{1/2}h^{2m}+n^{-1/2}h^{-d})\\
= & n^{-1/2}\sum_{i=1}^{n}\left[t_i y_{1i}+(1-t_i)\left\{m_1(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)\right.\right.\\
& \left.\left.+\frac{n^{-1}\sum_{j=1}^{n}t_j\{y_{1j}-m_1(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)\}K_h(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_j-\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)}{p_i f(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)}\right\}-E(Y_1)\right]+o_p(1)\\
= & n^{-1/2}\sum_{i=1}^{n}\left\{t_i y_{1i}+(1-t_i)m_1(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)-E(Y_1)\right\}\\
& +n^{-3/2}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{(1-t_i)t_j\{y_{1j}-m_1(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)\}K_h(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_j-\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)}{p_i f(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)}+o_p(1)\\
= & n^{-1/2}\sum_{i=1}^{n}\left\{t_i y_{1i}+(1-t_i)m_1(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)-E(Y_1)\right\}\\
& +n^{-1/2}\sum_{i=1}^{n}E\left\{\frac{(1-t_i)T_j\{Y_{1j}-m_1(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)\}K_h(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{X}_j-\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)}{p_i f(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)}\right\}\\
& +n^{-1/2}\sum_{j=1}^{n}E\left\{\frac{(1-T_i)t_j\{y_{1j}-m_1(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{X}_i)\}K_h(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_j-\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{X}_i)}{P_i f(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{X}_i)}\right\}\\
& -n^{1/2}E\left\{\frac{(1-T_i)T_j\{Y_{1j}-m_1(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{X}_i)\}K_h(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{X}_j-\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{X}_i)}{P_i f(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{X}_i)}\right\}+o_p(1)
\end{aligned}
$$

$$
\begin{aligned}
&= n^{-1/2}\sum_{i=1}^{n}\left\{t_i y_{1i} + (1-t_i)m_1(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i) - E(Y_1)\right\} \\
&\quad + n^{-1/2}\sum_{j=1}^{n} E[\exp\{-\eta(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{X}_j)\} \mid \boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_j] t_j\{y_{1j} - m_1(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_j)\} + o_p(1).
\end{aligned}
$$

Combining the above results, we get

$$
\begin{aligned}
&\sqrt{n}\{\widehat{E}(Y_1) - E(Y_1)\} \\
&= n^{-1/2}\sum_{i=1}^{n}\left\{t_i y_{1i} + (1-t_i)m_1(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i) - E(Y_1)\right\} \\
&\quad + n^{-1/2}\sum_{i=1}^{n} E[\exp\{-\eta(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{X}_i)\} \mid \boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i] t_i\{y_{1i} - m_1(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)\} \\
&\quad - n^{-1/2}\sum_{i=1}^{n} E[(1-P_i)\mathrm{vec}\{\mathbf{X}_{Li}\mathbf{m}_1'(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{X}_i)^{\mathrm{T}}\}]^{\mathrm{T}}\mathbf{B}_1 \\
&\quad \times t_i\{y_{1i} - m_1(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)\}\mathrm{vec}[\mathbf{m}_1'(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i) \otimes \{\mathbf{x}_{Li} - E(\mathbf{X}_{Li} \mid \boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)\}] + o_p(1).
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
\sqrt{n}\{\widehat{E}(Y_0) - E(Y_0)\} &= n^{-1/2}\sum_{i=1}^{n}\left\{(1-t_i)y_i + t_i\widehat{m}_0(\widehat{\boldsymbol{\beta}}_0^{\mathrm{T}}\mathbf{x}_i) - E(Y_0)\right\} \\
&= n^{-1/2}\sum_{i=1}^{n}\left\{(1-t_i)y_{0i} + t_i\widehat{m}_0(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i) - E(Y_0)\right\} \\
&\quad + \left\{n^{-1}\sum_{i=1}^{n} t_i \frac{\partial \widehat{m}_0(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)}{\partial \mathrm{vecl}(\boldsymbol{\beta}_0)^{\mathrm{T}}} + o_p(1)\right\}\sqrt{n}\mathrm{vecl}(\widehat{\boldsymbol{\beta}}_0 - \boldsymbol{\beta}_0) \\
&= n^{-1/2}\sum_{i=1}^{n}\left\{(1-t_i)y_{0i} + t_i\widehat{m}_0(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i) - E(Y_0)\right\} \\
&\quad + E[P_i\mathrm{vec}\{\mathbf{X}_{Li}\mathbf{m}_0'(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i)^{\mathrm{T}}\}]^{\mathrm{T}}\sqrt{n}\mathrm{vecl}(\widehat{\boldsymbol{\beta}}_0 - \boldsymbol{\beta}_0) + o_p(1).
\end{aligned}
$$

We have that

$$
\begin{aligned}
& E[P_i \mathrm{vec}\{\mathbf{X}_{Li}\mathbf{m}_0'(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i)^{\mathrm{T}}\}]^{\mathrm{T}}\sqrt{n}\mathrm{vecl}(\widehat{\boldsymbol{\beta}}_0-\boldsymbol{\beta}_0) \\
= & -n^{-1/2}\sum_{i=1}^n E[P_i \mathrm{vec}\{\mathbf{X}_{Li}\mathbf{m}_0'(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i)^{\mathrm{T}}\}]^{\mathrm{T}}\mathbf{B}_0 \\
& \times(1-t_i)\{y_{0i}-m_0(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)\}\mathrm{vec}[\mathbf{m}_0'(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)\otimes\{\mathbf{x}_{Li}-E(\mathbf{X}_{Li}\mid\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)\}] \\
& +o_p(1).
\end{aligned}
$$

On the other hand,

$$
\begin{aligned}
& n^{-1/2}\sum_{i=1}^n\left\{(1-t_i)y_{0i}+t_i\widehat{m}_0(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)-E(Y_0)\right\} \\
= & n^{-1/2}\sum_{i=1}^n\left\{(1-t_i)y_{0i}+t_i\frac{n^{-1}\sum_{j=1}^n(1-t_j)y_{0j}K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_j-\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)}{n^{-1}\sum_{j=1}^n(1-t_j)K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_j-\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)}-E(Y_0)\right\} \\
= & n^{-1/2}\sum_{i=1}^n\left[(1-t_i)y_{0i}+t_i\left\{\frac{E\{(1-T_i)Y_{0i}\mid\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i\}f(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)}{E\{(1-T_i)\mid\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i\}f(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)}\right.\right. \\
& +\frac{n^{-1}\sum_{j=1}^n(1-t_j)y_{0j}K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_j-\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)-E\{(1-T_i)Y_{0i}\mid\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i\}f(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)}{E(1-T_i\mid\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)f(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)} \\
& -E\{(1-T_i)Y_{0i}\mid\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i\}f(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i) \\
& \left.\times\frac{n^{-1}\sum_{j=1}^n(1-t_j)K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_j-\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)-E(1-T_i\mid\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)f(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)}{\{E(1-T_i\mid\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)f(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)\}^2}\right\} \\
& \left.-E(Y_0)\right]+O_p(n^{1/2}h^{2m}+n^{-1/2}h^{-d}) \\
= & n^{-1/2}\sum_{i=1}^n\left[(1-t_i)y_{0i}+t_i\left\{m_0(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)\right.\right. \\
& \left.+\frac{n^{-1}\sum_{j=1}^n(1-t_j)\{y_{0j}-m_0(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)\}K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_j-\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)}{(1-p_i)f(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)}\right\} \\
& \left.-E(Y_0)\right]+o_p(1)
\end{aligned}
$$

$$
\begin{aligned}
&= n^{-1/2}\sum_{i=1}^{n}\big\{(1-t_i)y_{0i}+t_i m_0(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)-E(Y_0)\big\}\\
&\quad +n^{-3/2}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{t_i(1-t_j)\{y_{0j}-m_0(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)\}K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_j-\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)}{(1-p_i)f(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)}+o_p(1)\\
&= n^{-1/2}\sum_{i=1}^{n}\big\{(1-t_i)y_{0i}+t_i m_0(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)-E(Y_0)\big\}\\
&\quad +n^{-1/2}\sum_{i=1}^{n}E\left\{\frac{t_i(1-T_j)\{Y_{0j}-m_0(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)\}K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_j-\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)}{(1-p_i)f(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)}\right\}\\
&\quad +n^{-1/2}\sum_{j=1}^{n}E\left\{\frac{T_i(1-t_j)\{y_{0j}-m_0(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i)\}K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_j-\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i)}{(1-P_i)f(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i)}\right\}\\
&\quad -n^{1/2}E\left\{\frac{T_i(1-T_j)\{Y_{0j}-m_0(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i)\}K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_j-\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i)}{(1-P_i)f(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i)}\right\}+o_p(1)\\
&= n^{-1/2}\sum_{i=1}^{n}\big\{(1-t_i)y_{0i}+t_i m_0(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)-E(Y_0)\big\}\\
&\quad +n^{-1/2}\sum_{j=1}^{n}E[\exp\{\eta(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{X}_j)\}\mid\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_j](1-t_j)\{y_{0j}-m_0(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_j)\}+o_p(1).
\end{aligned}
$$

Combining the above results, we get

$$
\begin{aligned}
&\sqrt{n}\{\widehat{E}(Y_0)-E(Y_0)\}\\
&= n^{-1/2}\sum_{i=1}^{n}\big\{(1-t_i)y_{0i}+t_i m_0(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)-E(Y_0)\big\}\\
&\quad +n^{-1/2}\sum_{i=1}^{n}E[\exp\{\eta(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{X}_i)\}\mid\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i](1-t_i)\{y_{0i}-m_0(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)\}\\
&\quad -n^{-1/2}\sum_{i=1}^{n}E[P_i\mathrm{vec}\{\mathbf{X}_{Li}\mathbf{m}_0'(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i)^{\mathrm{T}}\}]^{\mathrm{T}}\mathbf{B}_0\\
&\quad \times(1-t_i)\{y_{0i}-m_0(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)\}\mathrm{vec}[\mathbf{m}_0'(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)\otimes\{\mathbf{x}_{Li}-E(\mathbf{X}_{Li}\mid\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)\}]+o_p(1).
\end{aligned}
$$

Now combining the results regarding $\widehat{E}(Y_1)$ and $\widehat{E}(Y_0)$, we get

$$
\begin{aligned}
&\sqrt{n}[\{\widehat{E}(Y_1)-E(Y_1)\}-\{\widehat{E}(Y_0)-E(Y_0)\}]\\
= \;& n^{-1/2}\sum_{i=1}^{n}\big\{t_i y_{1i}-(1-t_i)y_{0i}+(1-t_i)m_1(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)-t_i m_0(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)-E(Y_1)\\
&+E(Y_0)\big\}+n^{-1/2}\sum_{i=1}^{n}E[\exp\{-\eta(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{X}_i)\}\mid\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i]t_i\{y_{1i}-m_1(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)\}\\
&-n^{-1/2}\sum_{i=1}^{n}E[\exp\{\eta(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{X}_i)\}\mid\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i](1-t_i)\{y_{0i}-m_0(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)\}\\
&-n^{-1/2}\sum_{i=1}^{n}E[(1-P_i)\mathrm{vec}\{\mathbf{X}_{Li}\mathbf{m}_1'(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{X}_i)^{\mathrm{T}}\}]^{\mathrm{T}}\mathbf{B}_1\\
&\times t_i\{y_{1i}-m_1(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)\}\mathrm{vec}[\mathbf{m}_1'(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)\otimes\{\mathbf{x}_{Li}-E(\mathbf{X}_{Li}\mid\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)\}]\\
&+n^{-1/2}\sum_{i=1}^{n}E[P_i\mathrm{vec}\{\mathbf{X}_{Li}\mathbf{m}_0'(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i)^{\mathrm{T}}\}]^{\mathrm{T}}\mathbf{B}_0\\
&\times(1-t_i)\{y_{0i}-m_0(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)\}\mathrm{vec}[\mathbf{m}_0'(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)\otimes\{\mathbf{x}_{Li}-E(\mathbf{X}_{Li}\mid\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)\}]+o_p(1)\\
= \;& n^{-1/2}\sum_{i=1}^{n}\big\{m_1(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)-m_0(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)-E(Y_1)+E(Y_0)\big\}\\
&+n^{-1/2}\sum_{i=1}^{n}E[1+\exp\{-\eta(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{X}_i)\}\mid\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i]t_i\{y_{1i}-m_1(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)\}\\
&-n^{-1/2}\sum_{i=1}^{n}E[1+\exp\{\eta(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{X}_i)\}\mid\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i](1-t_i)\{y_{0i}-m_0(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)\}\\
&-n^{-1/2}\sum_{i=1}^{n}E[(1-P_i)\mathrm{vec}\{\mathbf{X}_{Li}\mathbf{m}_1'(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{X}_i)^{\mathrm{T}}\}]^{\mathrm{T}}\mathbf{B}_1\\
&\times t_i\{y_{1i}-m_1(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)\}\mathrm{vec}[\mathbf{m}_1'(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)\otimes\{\mathbf{x}_{Li}-E(\mathbf{X}_{Li}\mid\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)\}]\\
&+n^{-1/2}\sum_{i=1}^{n}E[P_i\mathrm{vec}\{\mathbf{X}_{Li}\mathbf{m}_0'(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i)^{\mathrm{T}}\}]^{\mathrm{T}}\mathbf{B}_0(1-t_i)\\
&\times\{y_{0i}-m_0(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)\}\mathrm{vec}[\mathbf{m}_0'(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)\otimes\{\mathbf{x}_{Li}-E(\mathbf{X}_{Li}\mid\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)\}]+o_p(1).
\end{aligned}
$$

$\square$

S4 Proof of Properties of IMP2

$$
\begin{aligned}
& n^{1/2}\{\widehat{E}(Y_1) - E(Y_1)\} \\
= \; & n^{-1/2}\sum_{i=1}^{n}\left\{\widehat{m}_1(\widehat{\boldsymbol{\beta}}_1^{\mathrm{T}}\mathbf{x}_i) - E(Y_1)\right\} \\
= \; & n^{-1/2}\sum_{i=1}^{n}\left\{\widehat{m}_1(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i) - E(Y_1)\right\} \\
& + \left\{n^{-1}\sum_{i=1}^{n}\frac{\partial\widehat{m}_1(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)}{\partial\mathrm{vecl}(\boldsymbol{\beta}_1)^{\mathrm{T}}} + o_p(1)\right\}\sqrt{n}\mathrm{vecl}(\widehat{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}_1) \\
= \; & n^{-1/2}\sum_{i=1}^{n}\left\{\widehat{m}_1(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i) - E(Y_1)\right\} \\
& + E[\mathrm{vec}\{\mathbf{X}_{Li}\mathbf{m}_1'(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{X}_i)^{\mathrm{T}}\}]^{\mathrm{T}}\sqrt{n}\mathrm{vecl}(\widehat{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}_1) + o_p(1).
\end{aligned}
$$

We further have that

$$
\begin{aligned}
& E[\mathrm{vec}\{\mathbf{X}_{Li}\mathbf{m}_1'(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{X}_i)^{\mathrm{T}}\}]^{\mathrm{T}}\sqrt{n}\mathrm{vecl}(\widehat{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}_1) \\
= \; & -n^{-1/2}\sum_{i=1}^{n}E[\mathrm{vec}\{\mathbf{X}_{Li}\mathbf{m}_1'(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{X}_i)^{\mathrm{T}}\}]^{\mathrm{T}}\mathbf{B}_1 \\
& \times t_i\{y_{1i} - m_1(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)\}\mathrm{vec}[\mathbf{m}_1'(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)\otimes\{\mathbf{x}_{Li} - E(\mathbf{X}_{Li} \mid \boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)\}] + o_p(1).
\end{aligned}
$$

On the other hand,

$$
\begin{aligned}
& n^{-1/2}\sum_{i=1}^{n}\left\{\widehat{m}_1(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)-E(Y_1)\right\} \\
= \;& n^{-1/2}\sum_{i=1}^{n}\left\{\frac{n^{-1}\sum_{j=1}^{n}t_jy_{1j}K_h(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_j-\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)}{n^{-1}\sum_{j=1}^{n}t_jK_h(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_j-\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)}-E(Y_1)\right\} \\
= \;& n^{-1/2}\sum_{i=1}^{n}\left\{\frac{E(T_iY_{1i}\mid\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)f(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)}{E(T_i\mid\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)f(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)}\right. \\
& +\frac{n^{-1}\sum_{j=1}^{n}t_jy_{1j}K_h(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_j-\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)-E(T_iY_{1i}\mid\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)f(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)}{E(T_i\mid\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)f(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)} \\
& -E(T_iY_{1i}\mid\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)f(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)\frac{n^{-1}\sum_{j=1}^{n}t_jK_h(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_j-\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)-E(T_i\mid\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)f(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)}{\{E(T_i\mid\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)f(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)\}^2} \\
& \left.-E(Y_1)\right\}+O_p(n^{1/2}h^{2m}+n^{-1/2}h^{-d}) \\
= \;& n^{-1/2}\sum_{i=1}^{n}\left\{m_1(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)+\frac{n^{-1}\sum_{j=1}^{n}t_j\{y_{1j}-m_1(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)\}K_h(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_j-\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)}{p_if(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)}\right. \\
& \left.-E(Y_1)\right\}+o_p(1) \\
= \;& n^{-1/2}\sum_{i=1}^{n}\left\{m_1(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)-E(Y_1)\right\} \\
& +n^{-3/2}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{t_j\{y_{1j}-m_1(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)\}K_h(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_j-\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)}{p_if(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)}+o_p(1) \\
= \;& n^{-1/2}\sum_{i=1}^{n}\left\{m_1(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)-E(Y_1)\right\} \\
& +n^{-1/2}\sum_{i=1}^{n}E\left\{\frac{T_j\{Y_{1j}-m_1(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)\}K_h(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{X}_j-\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)}{p_if(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)}\right\} \\
& +n^{-1/2}\sum_{j=1}^{n}E\left\{\frac{t_j\{y_{1j}-m_1(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{X}_i)\}K_h(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_j-\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{X}_i)}{P_if(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{X}_i)}\right\} \\
& -n^{1/2}E\left\{\frac{T_j\{Y_{1j}-m_1(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{X}_i)\}K_h(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{X}_j-\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{X}_i)}{P_if(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{X}_i)}\right\}+o_p(1) \\
= \;& n^{-1/2}\sum_{i=1}^{n}\left\{m_1(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)-E(Y_1)\right\} \\
& +n^{-1/2}\sum_{j=1}^{n}E[1+\exp\{-\eta(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{X}_j)\}\mid\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_j]t_j\{y_{1j}-m_1(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_j)\}+o_p(1).
\end{aligned}
$$

Combining the above results, we get

$$
\begin{aligned}
& \sqrt{n}\{\widehat{E}(Y_1) - E(Y_1)\} \\
= \ & n^{-1/2}\sum_{i=1}^{n}\Big\{m_1(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i) - E(Y_1)\Big\} + n^{-1/2}\sum_{i=1}^{n}\Big(E[1+\exp\{-\eta(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{X}_i)\} \mid \boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i] \\
& \times t_i\{y_{1i} - m_1(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)\}\Big) - n^{-1/2}\sum_{i=1}^{n} E[\mathrm{vec}\{\mathbf{X}_{Li}\mathbf{m}_1'(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{X}_i)^{\mathrm{T}}\}]^{\mathrm{T}}\mathbf{B}_1 \\
& \times t_i\{y_{1i} - m_1(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)\}\mathrm{vec}[\mathbf{m}_1'(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)\otimes\{\mathbf{x}_{Li} - E(\mathbf{X}_{Li} \mid \boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)\}] + o_p(1).
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
& \sqrt{n}\{\widehat{E}(Y_0) - E(Y_0)\} \\
= \ & n^{-1/2}\sum_{i=1}^{n}\Big\{\widehat{m}_0(\widehat{\boldsymbol{\beta}}_0^{\mathrm{T}}\mathbf{x}_i) - E(Y_0)\Big\} \\
= \ & n^{-1/2}\sum_{i=1}^{n}\Big\{\widehat{m}_0(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i) - E(Y_0)\Big\} + \left\{n^{-1}\sum_{i=1}^{n}\frac{\partial\widehat{m}_0(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)}{\partial\mathrm{vecl}(\boldsymbol{\beta}_0)^{\mathrm{T}}} + o_p(1)\right\}\sqrt{n}\mathrm{vecl}(\widehat{\boldsymbol{\beta}}_0 - \boldsymbol{\beta}_0) \\
= \ & n^{-1/2}\sum_{i=1}^{n}\Big\{\widehat{m}_0(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i) - E(Y_0)\Big\} + E[\mathrm{vec}\{\mathbf{X}_{Li}\mathbf{m}_0'(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i)^{\mathrm{T}}\}]^{\mathrm{T}}\sqrt{n}\mathrm{vecl}(\widehat{\boldsymbol{\beta}}_0 - \boldsymbol{\beta}_0) + o_p(1).
\end{aligned}
$$

We further have that

$$
\begin{aligned}
& E[\mathrm{vec}\{\mathbf{X}_{Li}\mathbf{m}_0'(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i)^{\mathrm{T}}\}]^{\mathrm{T}}\sqrt{n}\mathrm{vecl}(\widehat{\boldsymbol{\beta}}_0 - \boldsymbol{\beta}_0) \\
= \ & -n^{-1/2}\sum_{i=1}^{n}E[\mathrm{vec}\{\mathbf{X}_{Li}\mathbf{m}_0'(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i)^{\mathrm{T}}\}]^{\mathrm{T}}\mathbf{B}_0 \\
& \times(1-t_i)\{y_{0i} - m_0(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)\}\mathrm{vec}[\mathbf{m}_0'(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)\otimes\{\mathbf{x}_{Li} - E(\mathbf{X}_{Li} \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)\}] + o_p(1).
\end{aligned}
$$

On the other hand,

$$
\begin{aligned}
& n^{-1/2}\sum_{i=1}^{n}\left\{\widehat{m}_0(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)-E(Y_0)\right\}\\
= \;& n^{-1/2}\sum_{i=1}^{n}\left\{\frac{n^{-1}\sum_{j=1}^{n}(1-t_j)y_{0j}K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_j-\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)}{n^{-1}\sum_{j=1}^{n}(1-t_j)K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_j-\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)}-E(Y_0)\right\}\\
= \;& n^{-1/2}\sum_{i=1}^{n}\left\{\frac{E\{(1-T_i)Y_{0i}\mid\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i\}f(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)}{E\{(1-T_i)\mid\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i\}f(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)}\right.\\
& +\frac{n^{-1}\sum_{j=1}^{n}(1-t_j)y_{0j}K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_j-\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)-E\{(1-T_i)Y_{0i}\mid\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i\}f(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)}{E(1-T_i\mid\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)f(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)}\\
& -E\{(1-T_i)Y_{0i}\mid\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i\}f(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)\\
& \times\frac{n^{-1}\sum_{j=1}^{n}(1-t_j)K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_j-\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)-E(1-T_i\mid\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)f(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)}{\{E(1-T_i\mid\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)f(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)\}^2}\\
& \left.-E(Y_0)\right\}+O_p(n^{1/2}h^{2m}+n^{-1/2}h^{-d})\\
= \;& n^{-1/2}\sum_{i=1}^{n}\left\{m_0(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)\right.\\
& \left.+\frac{n^{-1}\sum_{j=1}^{n}(1-t_j)\{y_{0j}-m_0(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)\}K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_j-\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)}{(1-p_i)f(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)}-E(Y_0)\right\}+o_p(1)\\
= \;& n^{-1/2}\sum_{i=1}^{n}\left\{m_0(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)-E(Y_0)\right\}\\
& +n^{-3/2}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{(1-t_j)\{y_{0j}-m_0(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)\}K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_j-\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)}{(1-p_i)f(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)}+o_p(1)\\
= \;& n^{-1/2}\sum_{i=1}^{n}\left\{m_0(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)-E(Y_0)\right\}\\
& +n^{-1/2}\sum_{i=1}^{n}E\left\{\frac{(1-T_j)\{Y_{0j}-m_0(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)\}K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_j-\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)}{(1-p_i)f(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)}\right\}\\
& +n^{-1/2}\sum_{j=1}^{n}E\left\{\frac{(1-t_j)\{y_{0j}-m_0(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i)\}K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_j-\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i)}{(1-P_i)f(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i)}\right\}\\
& -n^{1/2}E\left\{\frac{(1-T_j)\{Y_{0j}-m_0(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i)\}K_h(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_j-\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i)}{(1-P_i)f(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i)}\right\}+o_p(1)
\end{aligned}
$$

$$
\begin{aligned}
&= n^{-1/2}\sum_{i=1}^{n}\left\{m_0(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i) - E(Y_0)\right\} \\
&\quad + n^{-1/2}\sum_{j=1}^{n} E[1+\exp\{\eta(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{X}_j)\} \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_j](1-t_j)\{y_{0j} - m_0(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_j)\} \\
&\quad + o_p(1).
\end{aligned}
$$

Combining the above results, we get

$$
\begin{aligned}
& \sqrt{n}\{\widehat{E}(Y_0) - E(Y_0)\} \\
&= n^{-1/2}\sum_{i=1}^{n}\left\{m_0(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i) - E(Y_0)\right\} \\
&\quad + n^{-1/2}\sum_{i=1}^{n} E[1+\exp\{\eta(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{X}_i)\} \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i](1-t_i)\{y_{0i} - m_0(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)\} \\
&\quad - n^{-1/2}\sum_{i=1}^{n} E[\mathrm{vec}\{\mathbf{X}_{Li}\mathbf{m}_0'(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i)^{\mathrm{T}}\}]^{\mathrm{T}}\mathbf{B}_0 \\
&\quad \times(1-t_i)\{y_{0i} - m_0(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)\}\mathrm{vec}[\mathbf{m}_0'(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)\otimes\{\mathbf{x}_{Li} - E(\mathbf{X}_{Li} \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)\}] \\
&\quad + o_p(1).
\end{aligned}
$$

Now combining the results regarding $\widehat{E}(Y_1)$ and $\widehat{E}(Y_0)$, we get

$$
\begin{aligned}
& \sqrt{n}[\{\widehat{E}(Y_1) - E(Y_1)\} - \{\widehat{E}(Y_0) - E(Y_0)\}] \\
= \quad & n^{-1/2}\sum_{i=1}^{n}\left\{m_1(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i) - m_0(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i) - E(Y_1) + E(Y_0)\right\} \\
& + n^{-1/2}\sum_{i=1}^{n} E[1+\exp\{-\eta(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{X}_i)\} \mid \boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i] t_i\{y_{1i} - m_1(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)\} \\
& - n^{-1/2}\sum_{i=1}^{n} E[1+\exp\{\eta(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{X}_i)\} \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i](1-t_i)\{y_{0i} - m_0(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)\} \\
& - n^{-1/2}\sum_{i=1}^{n} E[\mathrm{vec}\{\mathbf{X}_{Li}\mathbf{m}_1'(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{X}_i)^{\mathrm{T}}\}]^{\mathrm{T}}\mathbf{B}_1 \\
& \times t_i\{y_{1i} - m_1(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)\}\mathrm{vec}[\mathbf{m}_1'(\boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i) \otimes \{\mathbf{x}_{Li} - E(\mathbf{X}_{Li} \mid \boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x}_i)\}] \\
& + n^{-1/2}\sum_{i=1}^{n} E[\mathrm{vec}\{\mathbf{X}_{Li}\mathbf{m}_0'(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{X}_i)^{\mathrm{T}}\}]^{\mathrm{T}}\mathbf{B}_0 \\
& \times (1-t_i)\{y_{0i} - m_0(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)\}\mathrm{vec}[\mathbf{m}_0'(\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i) \otimes \{\mathbf{x}_{Li} - E(\mathbf{X}_{Li} \mid \boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{x}_i)\}] \\
& + o_p(1).
\end{aligned}
$$

$\square$

S5 Results for Simulation Study, when n=500

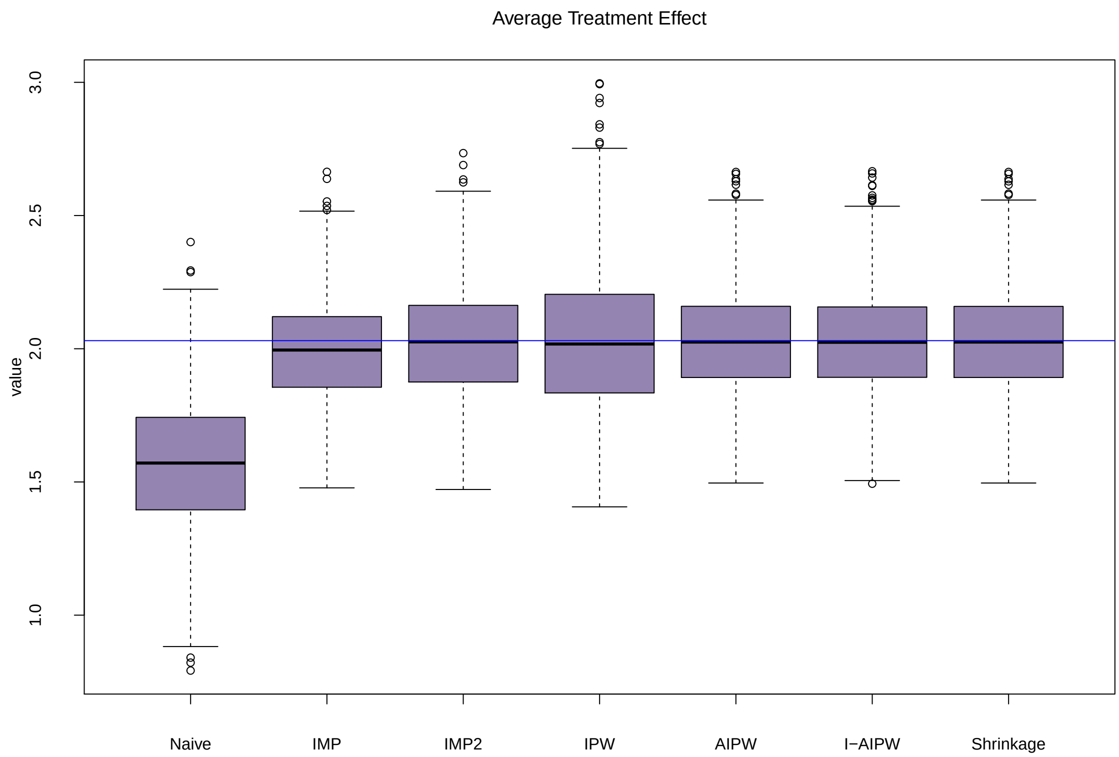

Figure 1: Boxplot of Naive, IMP, IMP2, IPW, AIPW, IAIPW and Shrinkage estimators under the setup for Study 1, when $n = 500$. The blue horizontal line is the true average causal effect, here 2.030.

Table 1: Results for Study 1 based on 500 replicates, where Full gives the average causal effect and corresponding standard deviation (sd) based on all potential responses, i.e. including the counterfactual ones not observable in practice, and Naive the same statistics based only on the observed potential responses. For the different estimators, we also compute the mean of the estimated sd (based on asymptotics, column $\widehat{\text{sd}}$), the empirical coverage obtained with confidence intervals based on these estimated sd (95% cvg), and finally the mean squared error (mse).

Estimators	Full	Naive	IMP	IMP2	IPW	AIPW	IAIPW	Shrinkage
mean	2.030	1.569	1.994	2.024	2.030	2.031	2.030	2.031
sd	0.185	0.253	0.197	0.213	0.275	0.196	0.196	0.196
$\widehat{\text{sd}}$	-	-	0.201	0.197	0.265	0.201	0.201	0.200
95% cvg	-	-	93.8%	92%	93.9%	96.1%	96%	96%
mse	-	-	0.040	0.046	0.076	0.039	0.039	0.039

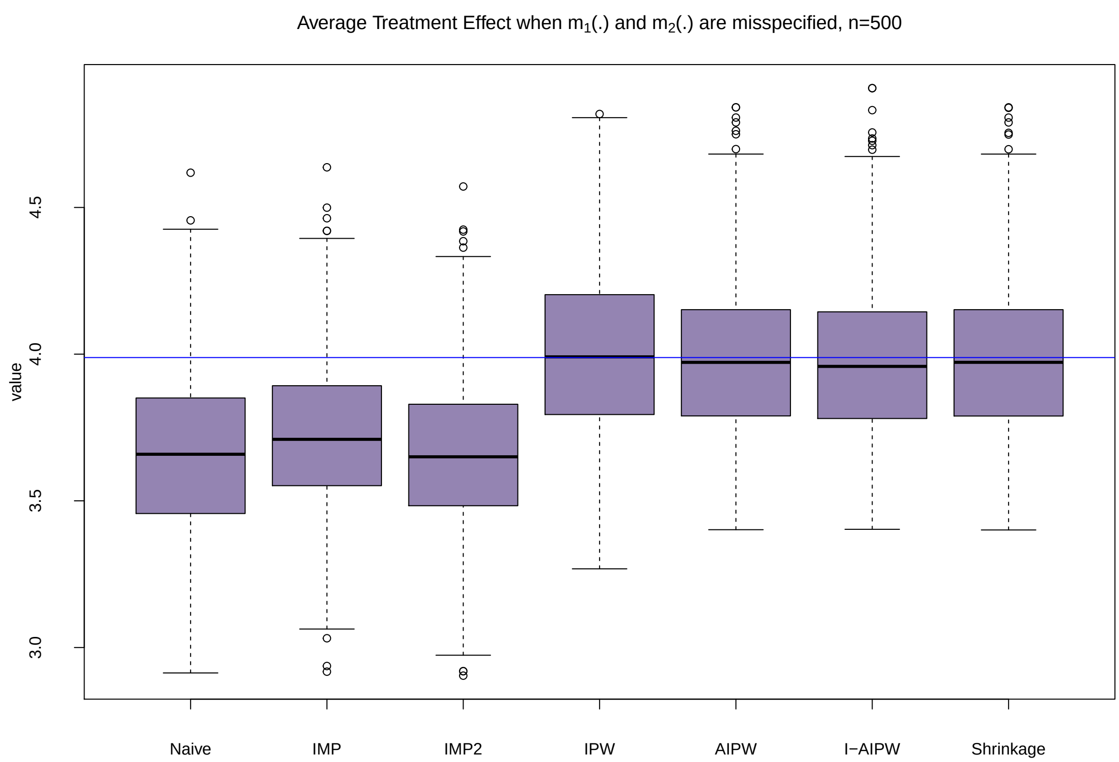

Figure 2: Boxplot of Naive, IMP, IMP2, IPW, AIPW, IAIPW and Shrinkage estimators under the setup for Study 2, where $m_1(\cdot)$ and $m_0(\cdot)$ are misspecified and $n = 500$. The blue horizontal line is the true average causal effect, here 3.988.

Table 2: Results for Study 2, where $m_1(\cdot)$ and $m_0(\cdot)$ are misspecified; see also caption of Table 1.

Estimators	Full	Naive	IMP	IMP2	IPW	AIPW	IAIPW	Shrinkage
mean	3.988	3.661	3.723	3.659	4.004	3.977	3.971	3.977
sd	0.189	0.282	0.254	0.258	0.293	0.266	0.269	0.266
$\widehat{\text{sd}}$	-	-	0.267	0.275	0.296	0.264	0.264	0.263
95% cvg	-	-	85.7%	81.1%	95.5%	95.8%	95.9%	95.7%
mse	-	-	0.135	0.175	0.086	0.071	0.072	0.071

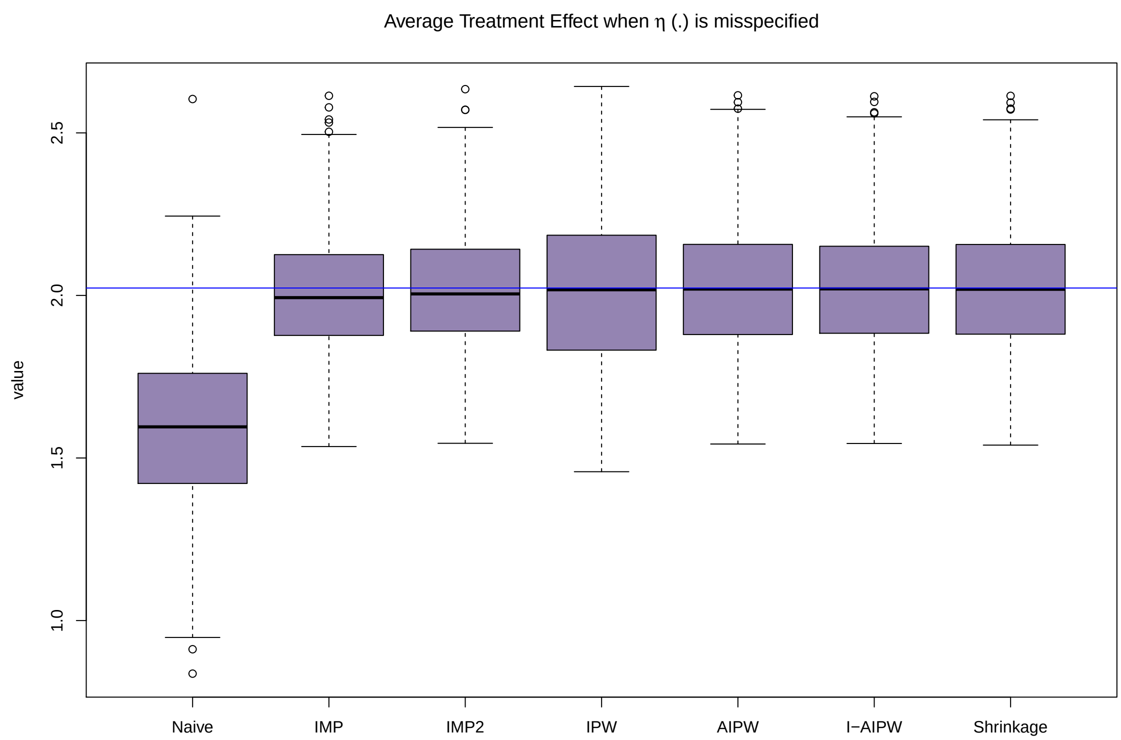

Figure 3: Boxplot of Naive, IMP, IMP2, IPW, AIPW, IAIPW and Shrinkage estimators under the setup for Study 3, where $\eta(\cdot)$ is misspecified and $n = 500$. The blue horizontal line is the true average causal effect, here 2.023.

Table 3: Results for Study 3, where $\eta(\cdot)$ is misspecified; see also caption of Table 1.

Estimators	Full	Naive	IMP	IMP2	IPW	AIPW	IAIPW	Shrinkage
mean	2.023	1.597	1.999	2.014	2.010	2.021	2.020	2.020
sd	0.175	0.236	0.181	0.182	0.244	0.195	0.192	0.194
$\widehat{\mathrm{sd}}$	-	-	0.189	0.189	0.223	0.195	0.195	0.193
95% cvg	-	-	94.6%	95.6%	92.4%	95.1%	95%	94.9%
mse	-	-	0.033	0.033	0.060	0.038	0.037	0.038

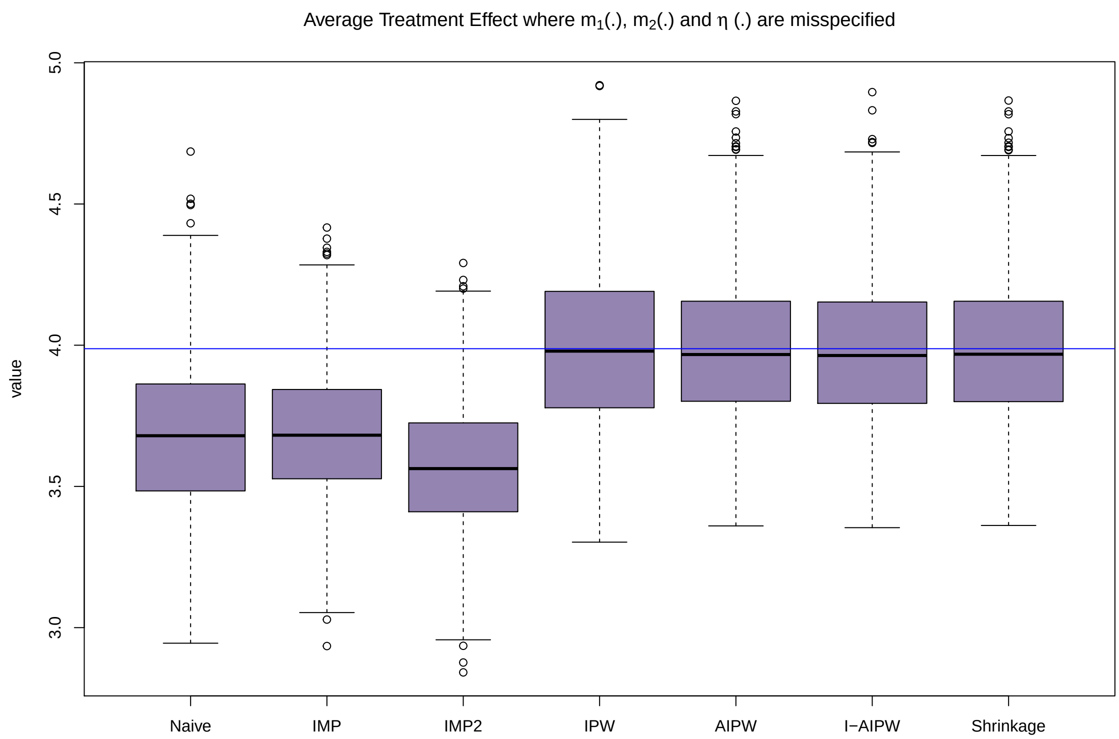

Figure 4: Boxplot of Naive, IMP, IMP2, IPW, AIPW, IAIPW and Shrinkage estimators under the setup for Study 4, where $m_1(\cdot)$, $m_0(\cdot)$ and $\eta(\cdot)$ is misspecified and $n = 500$. The blue horizontal line is the true average causal effect, here 3.988.

Table 4: Results for Study 4, where $m_1(\cdot)$, $m_0(\cdot)$, and $\eta(\cdot)$ are misspecified; see also caption of Table 1.

Estimators	Full	Naive	IMP	IMP2	IPW	AIPW	IAIPW	Shrinkage
mean	3.988	3.682	3.687	3.566	3.996	3.983	3.977	3.983
sd	0.189	0.280	0.237	0.233	0.290	0.264	0.261	0.264
$\widehat{\text{sd}}$	-	-	0.272	0.287	0.294	0.261	0.261	0.261
95% cvg	-	-	83.9%	75.8%	94.5%	94.9%	94.8%	94.9%
mse	-	-	0.147	0.232	0.084	0.070	0.068	0.070

S6 Results for Simulation Study, when n=200

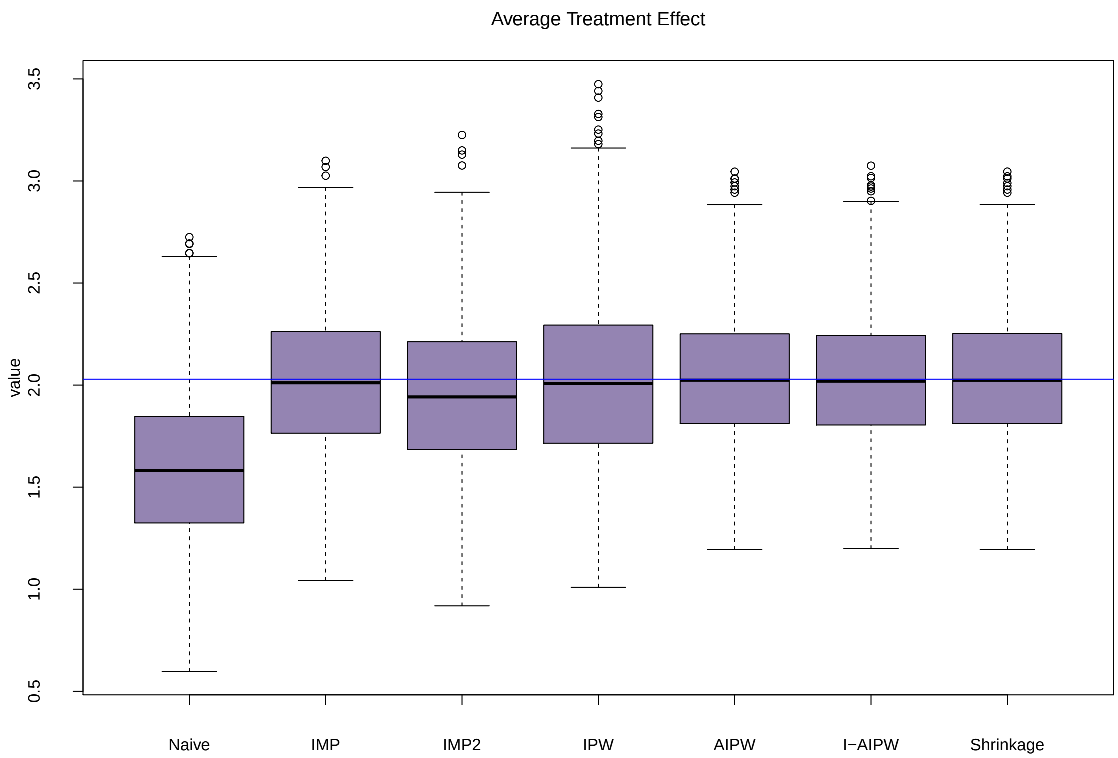

Figure 5: Boxplot of Naive, IMP, IMP2, IPW, AIPW, IAIPW and Shrinkage estimators under the setup for Study 1, when $n = 200$. The blue horizontal line is the true average causal effect, here 2.029.

Table 5: Results for Study 1 based on 200 replicates, where Full gives the average causal effect and corresponding standard deviation (sd) based on all potential responses, i.e. including the counterfactual ones not observable in practice, and Naive the same statistics based only on the observed potential responses. For the different estimators, we also compute the mean of the estimated sd (based on asymptotics, column $\widehat{\text{sd}}$), the empirical coverage obtained with confidence intervals based on these estimated sd (95% cvg), and finally the mean squared error (mse).

Estimators	Full	Naive	IMP	IMP2	IPW	AIPW	IAIPW	Shrinkage
mean	2.029	1.587	2.019	1.955	2.028	2.037	2.035	2.037
sd	0.288	0.388	0.356	0.393	0.439	0.324	0.324	0.324
$\widehat{\text{sd}}$	-	-	0.328	0.325	0.424	0.328	0.328	0.327
95% cvg	-	-	92.6%	87.8%	92.7%	94.8%	94.7%	94.7%
mse	-	-	0.127	0.160	0.193	0.105	0.105	0.105

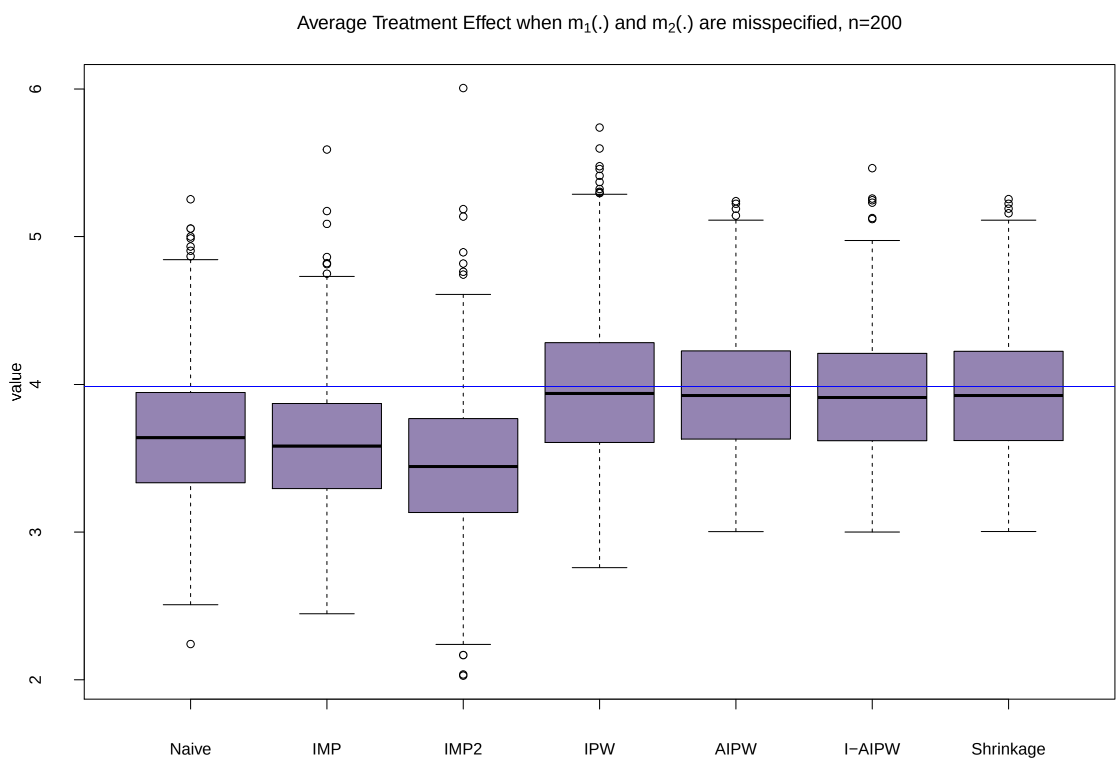

Figure 6: Boxplot of Naive, IMP, IMP2, IPW, AIPW, IAIPW and Shrinkage estimators under the setup for Study 2, where $m_1(\cdot)$ and $m_0(\cdot)$ are misspecified and $n = 200$. The blue horizontal line is the true average causal effect, here 3.987.

Table 6: Results for Study 2, where $m_1(\cdot)$ and $m_0(\cdot)$ are misspecified; see also caption of Table 5.

Estimators	Full	Naive	IMP	IMP2	IPW	AIPW	IAIPW	Shrinkage
mean	3.987	3.649	3.601	3.459	3.972	3.943	3.931	3.940
sd	0.294	0.451	0.433	0.472	0.499	0.429	0.420	0.430
$\widehat{\text{sd}}$	-	-	0.431	0.440	0.490	0.424	0.424	0.422
95% cvg	-	-	85.3%	77.7%	93.9%	95.3%	95.6%	95.3%
mse	-	-	0.336	0.502	0.250	0.186	0.179	0.187

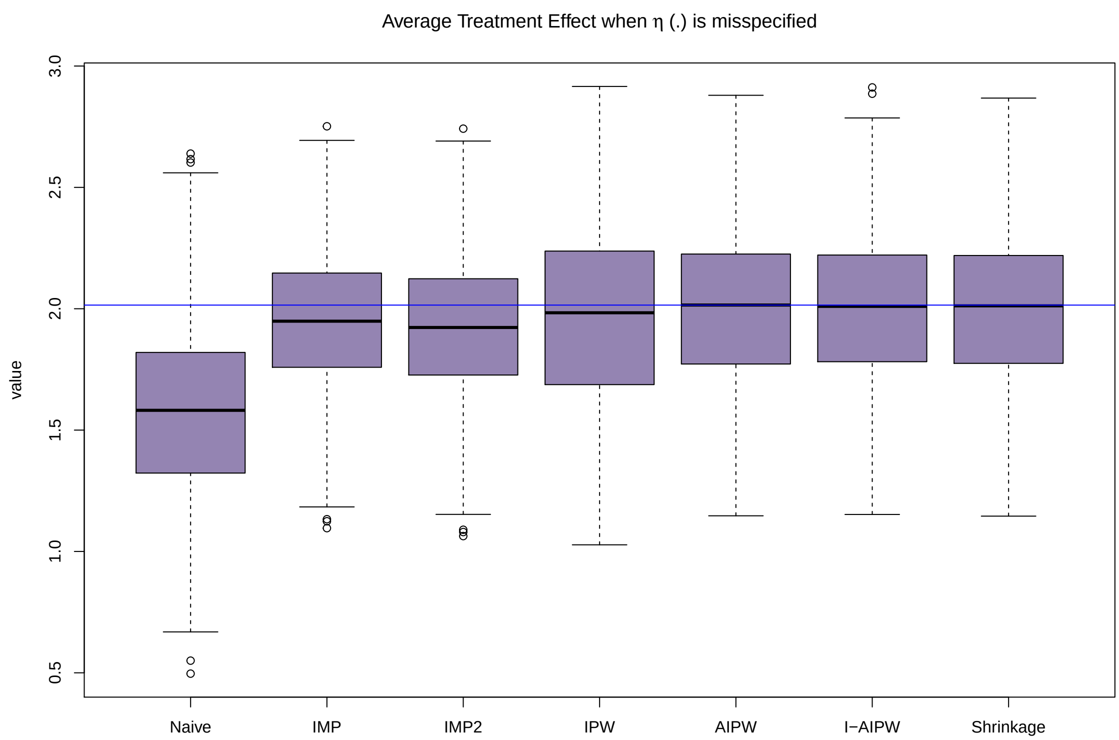

Figure 7: Boxplot of Naive, IMP, IMP2, IPW, AIPW, IAIPW and Shrinkage estimators under the setup for Study 3, where $\eta(\cdot)$ is misspecified and $n = 200$. The blue horizontal line is the true average causal effect, here 2.015.

Table 7: Results for Study 3, where $\eta(\cdot)$ is misspecified; see also caption of Table 5.

Estimators	Full	Naive	IMP	IMP2	IPW	AIPW	IAIPW	Shrinkage
mean	2.015	1.579	1.951	1.921	1.972	2.004	2.005	2.003
sd	0.270	0.359	0.281	0.281	0.386	0.317	0.311	0.315
$\widehat{\text{sd}}$	-	-	0.297	0.297	0.349	0.313	0.313	0.315
95% cvg	-	-	94.3%	93.7%	90.8%	95%	95.1%	94.4%
mse	-	-	0.083	0.088	0.151	0.101	0.097	0.099

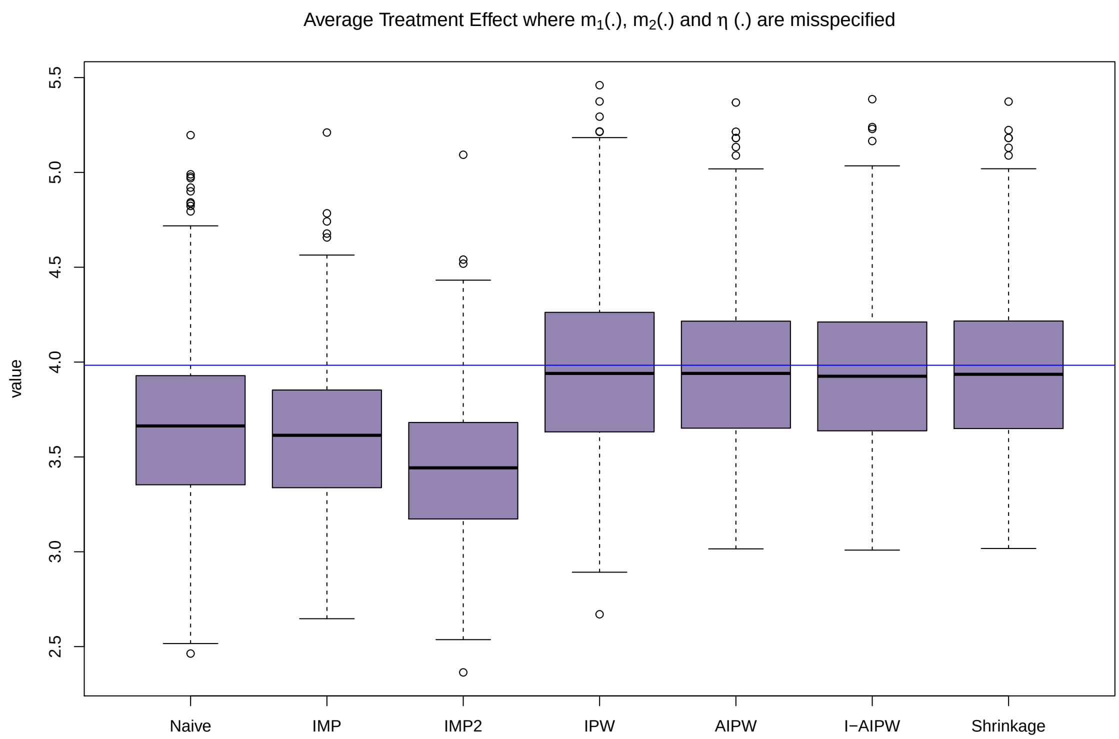

Figure 8: Boxplot of Naive, IMP, IMP2, IPW, AIPW, IAIPW and Shrinkage estimators under the setup for Study 4, where $m_1(\cdot)$, $m_0(\cdot)$ and $\eta(\cdot)$ is misspecified and $n = 200$. The blue horizontal line is the true average causal effect, here 3.983.

Table 8: Results for Study 4, where $m_1(\cdot)$, $m_0(\cdot)$, and $\eta(\cdot)$ are misspecified; see also caption of Table 5.

Estimators	Full	Naive	IMP	IMP2	IPW	AIPW	IAIPW	Shrinkage
mean	3.983	3.656	3.608	3.443	3.961	3.949	3.941	3.948
sd	0.295	0.436	0.377	0.368	0.471	0.422	0.419	0.422
$\widehat{\text{sd}}$	-	-	0.428	0.452	0.471	0.420	0.420	0.417
95% cvg	-	-	87.3%	82.5%	94.1%	94.2%	94.5%	94.3%
mse	-	-	0.283	0.428	0.222	0.179	0.177	0.179

S7 Results for Simulation Study, when n=100

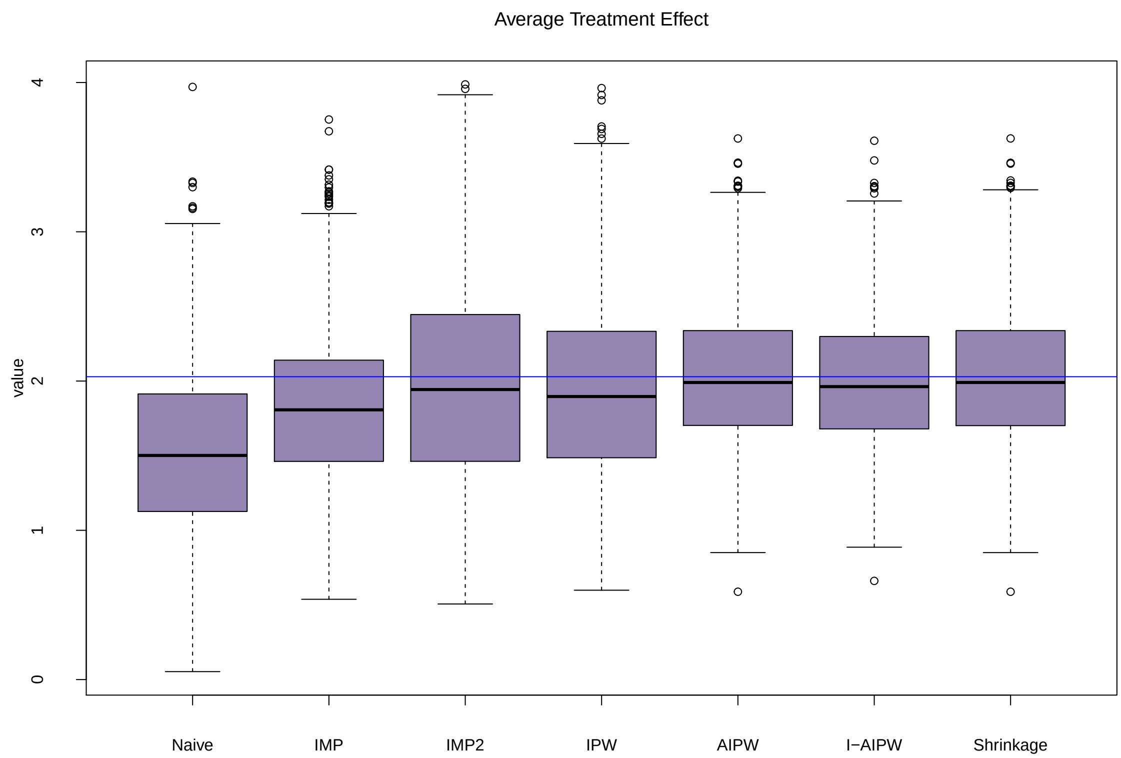

Figure 9: Boxplot of Naive, IMP, IMP2, IPW, AIPW, IAIPW and Shrinkage estimators under the setup for Study 1, when $n = 100$. The blue horizontal line is the true average causal effect, here 2.029.

Table 9: Results for Study 1 based on 100 replicates, where Full gives the average causal effect and corresponding standard deviation (sd) based on all potential responses, i.e. including the counterfactual ones not observable in practice, and Naive the same statistics based only on the observed potential responses. For the different estimators, we also compute the mean of the estimated sd (based on asymptotics, column $\widehat{\text{sd}}$), the empirical coverage obtained with confidence intervals based on these estimated sd (95% cvg), and finally the mean squared error (mse).

Estimators	Full	Naive	IMP	IMP2	IPW	AIPW	IAIPW	Shrinkage
mean	2.029	1.545	1.833	1.986	1.949	2.029	2.005	2.029
sd	0.399	0.569	0.516	0.708	0.609	0.490	0.479	0.490
$\widehat{\text{sd}}$	-	-	0.493	0.497	0.617	0.493	0.493	0.492
95% cvg	-	-	87.9%	81.4%	92.7%	94.1%	94.6%	94.1%
mse	-	-	0.304	0.503	0.378	0.240	0.230	0.240

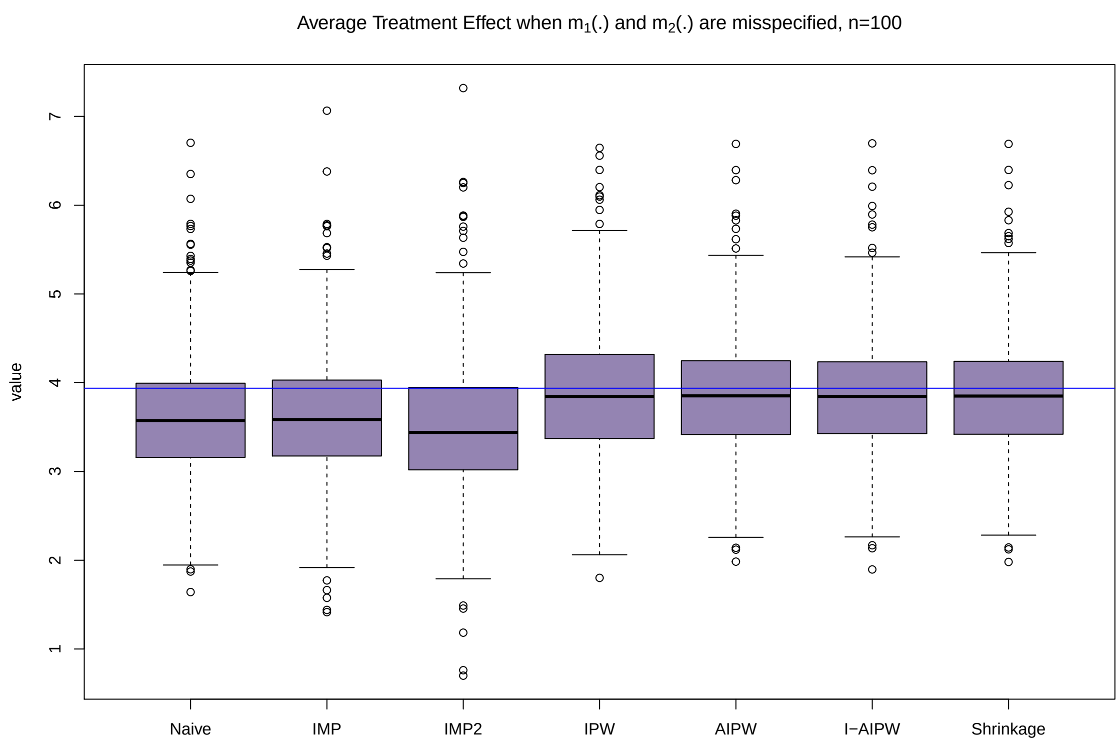

Figure 10: Boxplot of Naive, IMP, IMP2, IPW, AIPW, IAIPW and Shrinkage estimators under the setup for Study 2, where $m_1(\cdot)$ and $m_0(\cdot)$ are misspecified and $n = 100$. The blue horizontal line is the true average causal effect, here 3.938.

Table 10: Results for Study 2, where $m_1(\cdot)$ and $m_0(\cdot)$ are misspecified; see also caption of Table 9.

Estimators	Full	Naive	IMP	IMP2	IPW	AIPW	IAIPW	Shrinkage
mean	3.938	3.590	3.611	3.490	3.876	3.859	3.851	3.855
sd	0.426	0.663	0.667	0.729	0.727	0.642	0.625	0.641
$\widehat{\text{sd}}$	-	-	0.599	0.625	0.666	0.634	0.634	0.606
95% cvg	-	-	88.8%	85.3%	91.6%	92.1%	92.2%	91.4%
mse	-	-	0.552	0.733	0.532	0.419	0.399	0.418

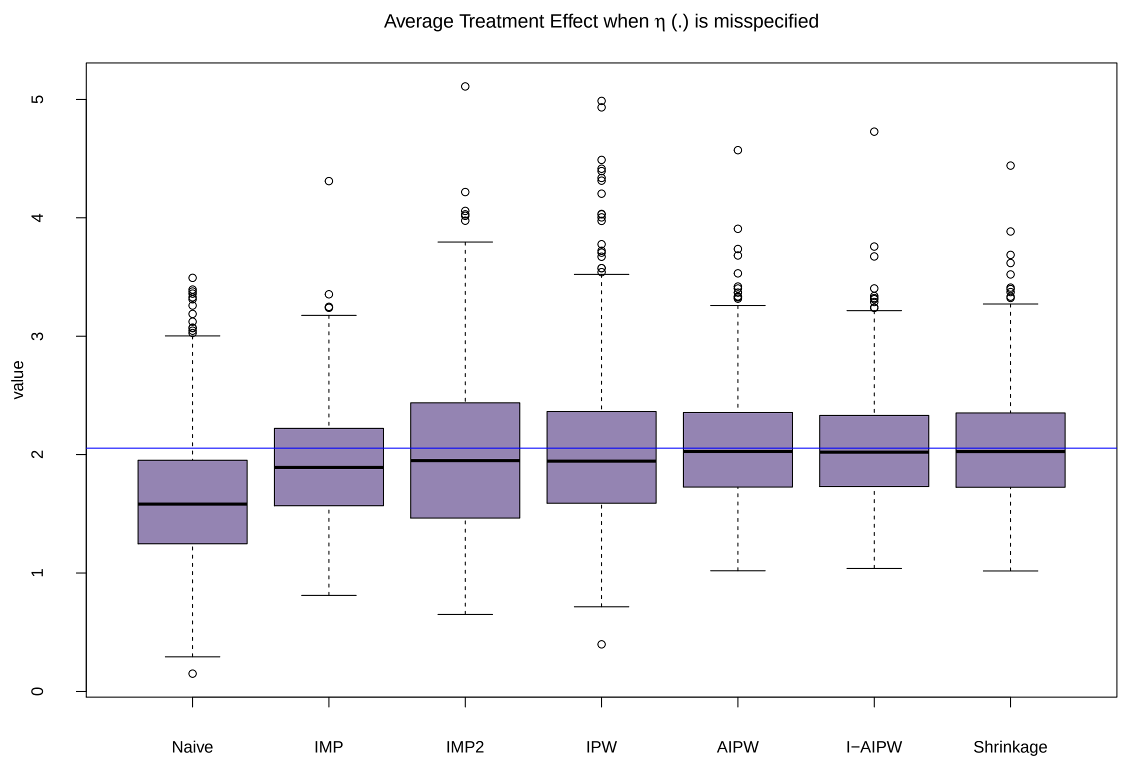

Figure 11: Boxplot of Naive, IMP, IMP2, IPW, AIPW, IAIPW and Shrinkage estimators under the setup for Study 3, where $\eta(\cdot)$ is misspecified and $n = 100$. The blue horizontal line is the true average causal effect, here 2.055.

Table 11: Results for Study 3, where $\eta(\cdot)$ is misspecified; see also caption of Table 9.

Estimators	Full	Naive	IMP	IMP2	IPW	AIPW	IAIPW	Shrinkage
mean	2.055	1.630	1.908	1.994	2.008	2.060	2.050	2.057
sd	0.390	0.535	0.476	0.672	0.610	0.471	0.462	0.468
$\widehat{\text{sd}}$	-	-	0.476	0.656	0.540	0.462	0.462	0.470
95% cvg	-	-	92%	89.9%	90.8%	93.9%	94.3%	93.9%
mse	-	-	0.248	0.455	0.375	0.222	0.213	0.219

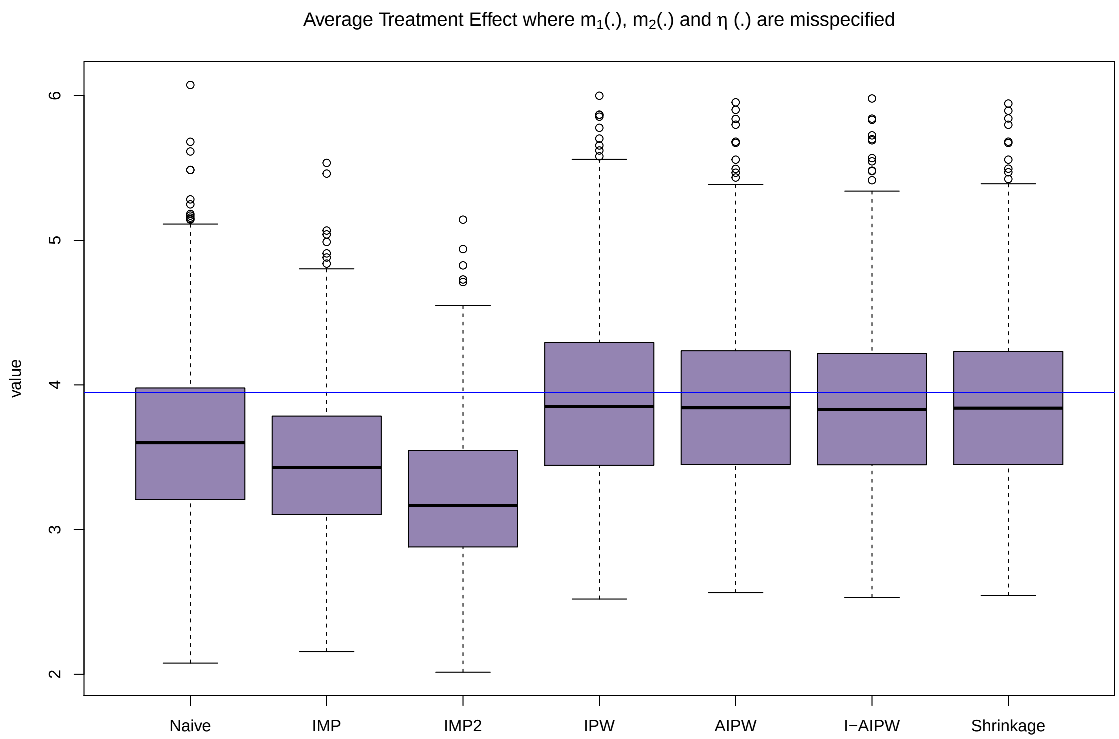

Figure 12: Boxplot of Naive, IMP, IMP2, IPW, AIPW, IAIPW and Shrinkage estimators under the setup for Study 4, where $m_1(\cdot)$, $m_0(\cdot)$ and $\eta(\cdot)$ is misspecified and $n = 100$. The blue horizontal line is the true average causal effect, here 3.948.

Table 12: Results for Study 4, where $m_1(\cdot)$, $m_0(\cdot)$, and $\eta(\cdot)$ are misspecified; see also caption of Table 9.

Estimators	Full	Naive	IMP	IMP2	IPW	AIPW	IAIPW	Shrinkage
mean	3.948	3.629	3.467	3.224	3.893	3.870	3.858	3.869
sd	0.418	0.583	0.512	0.494	0.628	0.582	0.560	0.583
$\widehat{\text{sd}}$	-	-	0.598	0.643	0.613	0.572	0.572	0.567
95% cvg	-	-	88.4%	82.9%	93.9%	92.4%	91.8%	92.1%
mse	-	-	0.493	0.768	0.397	0.344	0.344	0.346