Supplementary Materials for Semiparametric efficient estimation of genetic relatedness with machine learning methods

Xu Guo[1], Yiyuan Qian[1], Hongwei Shi[1], Weichao Yang[1] and Niwen Zhou[2*]

[1] School of Statistics, Beijing Normal University, Beijing, China

[2] Faculty of Arts and Sciences, Beijing Normal University, Zhuhai, China

Notation: For two positive sequences a_n and b_n, $a_n \lesssim b_n$ means $a_n \leq Cb_n$ for all n, $a_n \gtrsim b_n$ if $b_n \lesssim a_n$ and $a_n \asymp b_n$ if $a_n \lesssim b_n$ and $a_n \gtrsim b_n$. C is used to denote generic positive constants that may vary from place to place. What's more, we denote $a_n \lesssim_P b_n$ to represent $a_n \lesssim b_n$ in probability.

Proof of Theorem 1: Denote $\mathcal{P}$ as the distribution of $O = (X^\top, Y^*, Z^*, T_y, T_z)^\top$. For better illustration, we can rewrite I as

$$I = \Psi(\mathcal{P}) = E_{\mathcal{P}}\{[E_{\mathcal{P}}(Y|X) - E_{\mathcal{P}}(Y)][E_{\mathcal{P}}(Z|X) - E_{\mathcal{P}}(Z)]\}.$$

Note that the indicator variables T_y and T_z are independent of (X, Y, Z). Thus we have $E_{\mathcal{P}}(Y|X) = E_{\mathcal{P}}(Y^*|X, T_y = 1)$ and $E_{\mathcal{P}}(Y) = E_{\mathcal{P}}(Y^*|T_y = 1)$. The case for Z is similar. Thus, it follows that

$$I = \Psi(\mathcal{P}) = E_{\mathcal{P}}\{[E_{\mathcal{P}}(Y^*|X, T_y = 1) - E_{\mathcal{P}}(Y^*|T_y = 1)][E_{\mathcal{P}}(Z^*|X, T_z = 1) - E_{\mathcal{P}}(Z^*|T_z = 1)]\}.$$

Consider the following parametric submodel indexed by t, i.e.

$$\mathcal{P}_t = t\tilde{\mathcal{P}} + (1-t)\mathcal{P},$$

where $t \in [0,1]$, and $\tilde{\mathcal{P}}$ is a point mass at a single observation $\tilde{o} = (\tilde{x}^\top, \tilde{y}^*, \tilde{z}^*, \tilde{t}_y, \tilde{t}_z)^\top$. As mentioned in Hines et al. (2022), the efficient influence function (EIF) for I at observation $\tilde{o}$ directly as

$$\phi(\tilde{o}, \mathcal{P}) = \frac{d\Psi(\mathcal{P}_t)}{dt}\bigg|_{t=0},$$

where $\Psi(\mathcal{P}_t) = E_{\mathcal{P}_t}\{[E_{\mathcal{P}_t}(Y^*|X, T_y = 1) - E_{\mathcal{P}_t}(Y^*|T_y = 1)][E_{\mathcal{P}_t}(Z^*|X, T_z = 1) - E_{\mathcal{P}_t}(Z^*|T_z = 1)]\}$.

Denote $m_t^*(X) = E_{\mathcal{P}_t}\{[E_{\mathcal{P}_t}(Y^*|X, T_y = 1) - E_{\mathcal{P}_t}(Y^*|T_y = 1)]$ and $h_t^*(X) = [E_{\mathcal{P}_t}(Z^*|X, T_z = 1) - E_{\mathcal{P}_t}(Z^*|T_z = 1)]$. Further, the operator, ∂_t, applied to an arbitrary function $g(t)$, is defined as

$$\partial_t g(t) = \frac{dg(t)}{dt}\Big|_{t=0}.$$

Simple calculation entails that

$$\partial_t \Psi(\mathcal{P}_t) = E_{\mathcal{P}}\{\partial_t[m_t^*(X)h_t^*(X)]\} + m_0^*(\tilde{x})h_0^*(\tilde{x}) - \Psi(\mathcal{P}).$$

By the facts that

$$\partial_t E_{\mathcal{P}_t}(Y^*|X, T_y = 1) = \frac{\mathbb{I}_{(\tilde{x},\tilde{t}_y)}(X, T_y = 1)}{f(X, T_y = 1)}[\tilde{y}^* - E(Y^*|X, T_y = 1)],$$

and

$$\partial_t E_{\mathcal{P}_t}(Y|T_y = 1) = \frac{\mathbb{I}_{\tilde{t}_y}(T_y = 1)}{\Pr(T_y = 1)}[\tilde{y}^* - E(Y^*|T_y = 1)],$$

we have

$$\begin{aligned} E_{\mathcal{P}}\{\partial_t[m_t^*(X)]h_0^*(X)\} &= \int \frac{\mathbb{I}_{(\tilde{x},\tilde{t}_y)}(X, T_y = 1)}{f(X, T_y = 1)}[\tilde{y}^* - E(Y^*|X, T_y = 1)]h_0^*(X)f(X)dX \\ &= \frac{\mathbb{I}_{\tilde{t}_y}(T_y = 1)}{\Pr(T_y = 1|\tilde{x})}[\tilde{y}^* - E(Y^*|\tilde{x}, T_y = 1)]h_0^*(\tilde{x}), \end{aligned}$$

where the first equation holds because $E[h_0^*(X)] = 0$. Note that $T_y \perp X$, we have $\Pr(T_y = 1|X) = \Pr(T_y = 1)$, which implies that

$$E_{\mathcal{P}}\{\partial_t[m_t^*(X)]h_0^*(X)\} = \frac{\mathbb{I}_{\tilde{t}_y}(T_y = 1)}{\Pr(T_y = 1)}[\tilde{y}^* - E(Y^*|\tilde{x}, T_y = 1)]h_0^*(\tilde{x}).$$

Hence, by the chain rule and the quotient rule for derivatives, we obtain

$$\begin{aligned} E_{\mathcal{P}}\{\partial_t[m_t^*(X)h_t^*(X)]\} = &\frac{\mathbb{I}_{\tilde{t}_y}(T_y = 1)}{\Pr(T_y = 1)}[\tilde{y}^* - E(Y^*|\tilde{x}, T_y = 1)][E(Z^*|\tilde{x}, T_z = 1) - E(Z^*|T_z = 1)] \\ &+ \frac{\mathbb{I}_{\tilde{t}_z}(T_z = 1)}{\Pr(T_z = 1)}[\tilde{z} - E(Z^*|\tilde{x}, T_z = 1)][E(Y^*|\tilde{x}, T_y = 1) - E(Y^*|T_y = 1)]. \end{aligned}$$

Note that $m_0^*(\tilde{x})h_0^*(\tilde{x}) = [E(Y^*|\tilde{x}, T_y = 1) - E(Y^*|T_y = 1)][E(Z^*|\tilde{x}, T_z = 1) - E(Z^*|T_z = 1)]$. It implies that

$$\begin{aligned}\phi(\tilde{o}, \mathcal{P}) = & \frac{\mathbb{I}_{\tilde{t}_y}(T_y = 1)}{\Pr(T_y = 1)}[\tilde{y}^* - E(Y^*|\tilde{x}, T_y = 1)][E(Z^*|\tilde{x}, T_z = 1) - E(Z^*|T_z = 1)] \\ & + \frac{\mathbb{I}_{\tilde{t}_z}(T_z = 1)}{\Pr(T_z = 1)}[\tilde{z}^* - E(Z^*|\tilde{x}, T_z = 1)][E(Y^*|\tilde{x}, T_y = 1) - E(Y^*|T_y = 1)] \\ & + [E(Y^*|\tilde{x}, T_y = 1) - E(Y^*|T_y = 1)][E(Z^*|\tilde{x}, T_z = 1) - E(Z^*|T_z = 1)] - \Psi(\mathcal{P}).\end{aligned} \tag{S.1}$$

Since $\phi(O, \mathcal{P})$ has finite variance, we conclude that the EIF for I is

$$\begin{aligned}\phi(O, \mathcal{P}) = & \frac{\mathbb{I}_{T_y}(1)}{\Pr(T_y = 1)}[Y^* - E(Y^*|X, T_y = 1)][E(Z^*|X, T_z = 1) - E(Z^*|T_z = 1)] \\ & + \frac{\mathbb{I}_{T_z}(1)}{\Pr(T_z = 1)}[Z^* - E(Z^*|X, T_z = 1)][E(Y^*|X, T_y = 1) - E(Y^*|T_y = 1)] \\ & + [E(Y^*|X, T_y = 1) - E(Y^*|T_y = 1)][E(Z^*|X, T_z = 1) - E(Z^*|T_z = 1)] - \Psi(\mathcal{P}).\end{aligned}$$

Proof of Theorem 2: Let $\epsilon = Y - m(X)$ and $\eta = Z - h(X)$. Note that

$$\begin{aligned}G_1 = & \frac{1}{N_y}\sum_{i=1}^{N_y}[Y_i - \hat{m}(X_i)][\hat{h}(X_i) - \bar{Z}_N] \\ = & \frac{1}{N_y}\sum_{i=1}^{N_y}[\epsilon_i + m(X_i) - \tilde{m}(X_i) + \tilde{m}(X_i) - \hat{m}(X_i)][\hat{h}(X_i) - \tilde{h}(X_i) + \tilde{h}(X_i) - EZ + EZ - \bar{Z}_N] \\ = & \frac{1}{N_y}\sum_{i=1}^{N_y}[\epsilon_i + m(X_i) - \tilde{m}(X_i)][\tilde{h}(X_i) - EZ] + \frac{1}{N_y}\sum_{i=1}^{N_y}[\tilde{m}(X_i) - \hat{m}(X_i)][\tilde{h}(X_i) - EZ] \\ & + \frac{1}{N_y}\sum_{i=1}^{N_y}\epsilon_i[\hat{h}(X_i) - \tilde{h}(X_i)] + \frac{1}{N_y}\sum_{i=1}^{N_y}[m(X_i) - \tilde{m}(X_i)][\hat{h}(X_i) - \tilde{h}(X_i)] \\ & + \frac{1}{N_y}\sum_{i=1}^{N_y}[\tilde{m}(X_i) - \hat{m}(X_i)][\hat{h}(X_i) - \tilde{h}(X_i)] + [EZ - \bar{Z}_N]\frac{1}{N_y}\sum_{i=1}^{N_y}[\epsilon_i + m(X_i) - \tilde{m}(X_i)] \\ & + [EZ - \bar{Z}_N]\frac{1}{N_y}\sum_{i=1}^{N_y}[\tilde{m}(X_i) - \hat{m}(X_i)] := \sum_{i=1}^{8} G_{1i}.\end{aligned}$$

For the term G_{12}, by Cauchy-Schwartz inequality, we have

$$\begin{aligned}& \frac{1}{N_y}\sum_{i=1}^{N_y}[\tilde{m}(X_i) - \hat{m}(X_i)][\tilde{h}(X_i) - EZ] \\ \leq & \left(\frac{1}{N_y}\sum_{i=1}^{N_y}[\tilde{m}(X_i) - \hat{m}(X_i)]^2\right)^{1/2}\left(\frac{1}{N_y}\sum_{i=1}^{N_y}[\tilde{h}(X_i) - EZ]^2\right)^{1/2} = o_p(1),\end{aligned}$$

where the last equation holds under the conditions $E[\tilde{h}(X_i) - EZ]^2 < \infty$ and $E[\hat{h}(X_i) - h(X_i)]^2 = o(1)$. Similarly, due to the consistencies of $\hat{m}(X_i), \hat{h}(X_i)$ to $\tilde{m}(X_i), \tilde{h}(X_i)$, respectively, we conclude that

$$\begin{aligned} G_1 &= \frac{1}{N_y}\sum_{i=1}^{N_y}[\epsilon_i + m(X_i) - \tilde{m}(X_i)][\tilde{h}(X_i) - EZ] + o_p(1) \\ &= E\{[m(X) - \tilde{m}(X)][\tilde{h}(X) - EZ]\} + o_p(1), \end{aligned}$$

The last equation follows from the fact that $E\{\epsilon[\tilde{h}(X_i) - EZ]\} = 0$. Similarly for the second term of $\tilde{I}_N$, we have $G_2 = E\{[\tilde{m}(X) - EY][h(X) - \tilde{h}(X)]\} + o_p(1)$.

For the third term of $\tilde{I}_N$, note that

$$\begin{aligned} G_3 &= \frac{1}{N}\sum_{i=1}^{N}[\hat{m}(X_i) - \bar{Y}_N][\hat{h}(X_i) - \bar{Z}_N] \\ &= \frac{1}{N}\sum_{i=1}^{N}[\hat{m}(X_i) - \tilde{m}(X_i) + \tilde{m}(X_i) - EY + EY - \bar{Y}_N][\hat{h}(X_i) - \tilde{h}(X_i) + \tilde{h}(X_i) - EZ + EZ - \bar{Z}_N] \\ &= \frac{1}{N}\sum_{i=1}^{N}[\tilde{m}(X_i) - EY][\tilde{h}(X_i) - EZ] + o_p(1) \\ &= E[\tilde{m}(X) - EY][\tilde{h}(X) - EZ] + o_p(1). \end{aligned}$$

The third equation holds due to the consistencies of the estimators.

In sum, we have

$$\tilde{I}_N = G_1 + G_2 + G_3 = E[\tilde{m}(X_i) - m(X_i)][h(X_i) - \tilde{h}(X_i)] + I + o_p(1).$$

Then the Theorem follows.

Proof of Theorem 3: Note that the proof of this theorem is a special case of that of Theorem 6 with $g_1(x) = g_2(x) = x$. Hence we omit the detail.

Proof of Theorem 4: If $(\hat{\sigma}^2 - \sigma^2)/\sigma^2 = o_p(1)$ holds, note that

$$\frac{I_n^* - I}{\sigma} \xrightarrow{d} N\Big(0, 1\Big),$$

which implies

$$\frac{I_n^* - I}{\hat{\sigma}} \xrightarrow{d} N\Big(0, 1\Big).$$

Hence we can finish the proof.

In the following, we aim to show that $(\hat{\sigma}^2 - \sigma^2)/\sigma^2 = o_p(1)$. Denote

$$\delta_i - n^{-1}I = \frac{1}{n_y}\epsilon_i^* I(i \in \mathcal{D}_{2y}) + \frac{1}{n_z}\eta_i^* I(i \in \mathcal{D}_{2z}) + \frac{1}{n}\xi_i^*,$$

and $\tilde{K}_1 = \sum_{i\in\mathcal{D}_2}(\delta_i - n^{-1}I)^2$, where $\epsilon_i^* = \epsilon_i[h(X_i) - EZ]$, $\eta_i^* = \eta_i[m(X_i) - EY]$, and $\xi_i^* = [m(X_i) - EY][h(X_i) - EZ] - I$. Similarly, we can definite $\tilde{K}_2$. Denote

$$\tilde{\sigma}^2 = \frac{\tilde{K}_1 + \tilde{K}_2}{4}.$$

We have

$$\frac{\hat{\sigma}^2 - \sigma^2}{\sigma^2} = \frac{\tilde{\sigma}^2 - \sigma^2}{\sigma^2} + \frac{\hat{\sigma}^2 - \tilde{\sigma}^2}{\sigma^2}. \tag{S.2}$$

We first consider the first term. Note that $E(\tilde{\sigma}^2 - \sigma^2) = 0$, and

$$\mathrm{Var}(\tilde{\sigma}^2 - \sigma^2) \lesssim \mathrm{Var}(\tilde{K}_1) + \mathrm{Var}(\tilde{K}_2).$$

We have

$$\mathrm{Var}(\tilde{K}_1) \lesssim E[\sum_{i\in\mathcal{D}_2}(\delta_i - n^{-1}I)^4] \lesssim \frac{1}{n_y^3}E(\epsilon_i^{*4}) + \frac{1}{n_z^3}E(\eta_i^{*4}) + \frac{1}{n^3}E(\xi_i^{*4}),$$

and $\mathrm{Var}(\tilde{K}_2) \lesssim \frac{1}{n_y^3}E(\epsilon_i^{*4}) + \frac{1}{n_z^3}E(\eta_i^{*4}) + \frac{1}{n^3}E(\xi_i^{*4})$, which implies that

$$\tilde{\sigma}^2 - \sigma^2 = O_p\Big(\frac{1}{n_y^{3/2}} + \frac{1}{n_z^{3/2}}\Big).$$

Since

$$\sigma^2 = \frac{1}{2n_y}E(\epsilon_i^{*2}) + \frac{1}{2n_z}E(\eta_i^{*2}) + \frac{1}{2n}E(\xi_i^{*2}) + \frac{N_0}{N_y N_z}E[\epsilon^*\eta^*] = O_p\left(\frac{1}{n_y} + \frac{1}{n_z}\right),$$

It follows that

$$\frac{\tilde{\sigma}^2 - \sigma^2}{\sigma^2} = o_p(1).$$

Next, we turn to consider the bound for $(\hat{\sigma}^2 - \tilde{\sigma}^2)/\sigma^2$. Note that

$$\frac{\hat{\sigma}^2 - \tilde{\sigma}^2}{\sigma^2} = \frac{K_1 - \tilde{K}_1}{4\sigma^2} + \frac{K_2 - \tilde{K}_2}{4\sigma^2}.$$

We can rewrite $K_1 - \tilde{K}_1$ as

$$
\begin{aligned}
K_1 - \tilde{K}_1 &= \sum_{i\in\mathcal{D}_2}(\hat{\delta}_i - n^{-1}I_n)^2 - \sum_{i\in\mathcal{D}_2}(\delta_i - n^{-1}I)^2 \\
&= \sum_{i\in\mathcal{D}_2}(\hat{\delta}_i - n^{-1}I_n - \delta_i + n^{-1}I)^2 + 2\sum_{i\in\mathcal{D}_2}(\delta_i - n^{-1}I)\sum_{i\in\mathcal{D}_2}(\hat{\delta}_i - n^{-1}I_n - \delta_i + n^{-1}I) \\
&\le \sum_{i\in\mathcal{D}_2}(\hat{\delta}_i - n^{-1}I_n - \delta_i + n^{-1}I)^2 + 2\sqrt{\tilde{K}_1}\sqrt{\sum_{i\in\mathcal{D}_2}(\hat{\delta}_i - n^{-1}I_n - \delta_i + n^{-1}I)^2} \\
&=: K_{11} + 2K_{12}.
\end{aligned}
$$

For the term K_{11}, we have

$$
\sum_{i\in\mathcal{D}_2}(\hat{\delta}_i - n^{-1}I_n - \delta_i + n^{-1}I)^2 \le 2n^{-1}(I_n - I)^2 + 2\sum_{i\in\mathcal{D}_2}(\hat{\delta}_i - \delta_i)^2.
$$

By the asymptotic normality of I_n, we have $(I_n - I)^2 = O_p(2\sigma^2)$, which implies that $n^{-1}(I_n - I)^2/4\sigma^2 = o_p(1)$.

Hence, it is left to show that $\sum_{i\in\mathcal{D}_2}(\hat{\delta}_i - \delta_i)^2/4\sigma^2 = o_p(1)$. Note that

$$
\begin{aligned}
\sum_{i\in\mathcal{D}_2}(\hat{\delta}_i - \delta_i)^2 &\lesssim \frac{1}{n_y^2}\sum_{i\in\mathcal{D}_{2y}}\{[Y_i - \hat{m}_{\mathcal{D}_1}(X_i)][\hat{h}_{\mathcal{D}_1}(X_i) - \bar{Z}_n] - [Y_i - m(X_i)][h(X_i) - EZ]\}^2 \\
&\quad + \frac{1}{n_z^2}\sum_{i\in\mathcal{D}_{2z}}\{[\hat{m}_{\mathcal{D}_1}(X_i) - \bar{Y}_n][Z_i - \hat{h}_{\mathcal{D}_1}(X_i)] - [m(X_i) - EY][Z_i - h(X_i)]\}^2 \\
&\quad + \frac{1}{n^2}\sum_{i\in\mathcal{D}_2}\{[\hat{m}_{\mathcal{D}_1}(X_i) - \bar{Y}_n][\hat{h}_{\mathcal{D}_1}(X_i) - \bar{Z}_n] - [m(X_i) - EY][h(X_i) - EZ]\}^2 \\
&=: K_{111} + K_{112} + K_{113}.
\end{aligned}
$$

We first consider the bound of the term K_{111}. Note that

$$
\begin{aligned}
K_{111} &= \frac{1}{n_y^2}\sum_{i\in\mathcal{D}_{2y}}\{[Y_i - \hat{m}_{\mathcal{D}_1}(X_i)][\hat{h}_{\mathcal{D}_1}(X_i) - \bar{Z}_n] - [Y_i - m(X_i)][h(X_i) - EZ]\}^2 \\
&\lesssim_P \frac{1}{n_y^2}\sum_{i\in\mathcal{D}_{2y}}\epsilon_i^2[\hat{h}_{\mathcal{D}_1}(X_i) - h(X_i)]^2 + [EZ - \bar{Z}_n]^2\frac{1}{n_y^2}\sum_{i\in\mathcal{D}_{2y}}\epsilon_i^2 \\
&\quad + \frac{1}{n_y^2}\sum_{i\in\mathcal{D}_{2y}}[\hat{m}_{\mathcal{D}_1}(X_i) - m(X_i)]^2[h(X_i) - EZ]^2 + \frac{1}{n_y^2}\sum_{i\in\mathcal{D}_{2y}}[\hat{m}_{\mathcal{D}_1}(X_i) - m(X_i)]^2[\hat{h}_{\mathcal{D}_1}(X_i) - h(X_i)]^2 \\
&\quad + [EZ - \bar{Z}_n]^2\frac{1}{n_y^2}\sum_{i\in\mathcal{D}_{2y}}[\hat{m}_{\mathcal{D}_1}(X_i) - m(X_i)]^2.
\end{aligned}
$$

Based on the fact that $[EZ - \bar{Z}_n]^2 = O_p(n_z^{-1})$ and under Conditions 2.1–2.4, it is easily to show that

$$E(|K_{111}|) = o_p\left(\frac{1}{n_y}\right),$$

which follows that $K_{111} = o_p\left({n_y}^{-1}\right)$.

Similarly, we can obtain that

$$K_{112} = o_p\left(\frac{1}{n_z}\right), \text{ and } K_{113} = o_p\left(\frac{1}{n}\right).$$

It follows that

$$\frac{\sum_{i\in\mathcal{D}_2}(\hat{\delta}_i - \delta_i)^2}{4\sigma^2} = o_p(1),$$

which implies that $K_{11}/4\sigma^2 = o_p(1)$.

Similarly as the discussion of $\tilde{\sigma}^2$, we can show that

$$\frac{\tilde{K}_1}{4\sigma^2} = o_p(1).$$

Hence, for the term K_{12}, we have

$$\frac{K_{12}}{4\sigma^2} = \sqrt{\frac{\tilde{K}_1}{4\sigma^2}}\sqrt{\frac{K_{11}}{4\sigma^2}} = o_p(1).$$

Based on the above results, it follows that

$$(K_1 - \tilde{K}_1)/4\sigma^2 = o_p(1).$$

Similarly, we can show that

$$(K_2 - \tilde{K}_2)/4\sigma^2 = o_p(1).$$

So that we finish the proof.

Proof of Proposition 3.1: Similarly, we can rewrite ρ as

$$\rho = \Psi_1(\mathcal{P}) = \frac{Cov[m(X), h(X)]}{\sqrt{Var[m(X)]Var[h(X)]}}.$$

Consider the same parametric submodel,

$$\mathcal{P}_t = t\tilde{\mathcal{P}} + (1-t)\mathcal{P},$$

where $t \in [0,1]$. Hence, the efficient influence function(EIF) for ρ at observation $\tilde{o}$ directly as

$$\phi_1(\tilde{o}, \mathcal{P}) = \frac{d\Psi_1(\mathcal{P}_t)}{dt}|_{t=0} = \frac{\phi(\tilde{o}, \mathcal{P})}{\sqrt{B_0^y B_0^z}} - \frac{\Psi_1(\mathcal{P})}{2B_0^y B_0^z}\partial_t(B_t^y B_t^u),$$

where $\phi(\tilde{o}, \mathcal{P})$ is the EIF for I, $B_t^y = E_{\mathcal{P}_t}\{[E_{\mathcal{P}_t}(Y^*|X, T_y = 1) - E_{\mathcal{P}_t}(Y^*|T_y = 1)]^2\}$, and $B_0^y = Var(E[Y^*|X, T_y = 1])$, and the definition of B_t^z is similar. Similarly, as the arguments in Theorem 2, we can show that $\partial_t B_t^u$ equals to

$$\begin{aligned}\phi_u(\tilde{o}, \mathcal{P}) = 2&\frac{\mathbb{I}_{\tilde{t}_u}(T_u = 1)}{P(T_u = 1)}[\tilde{u}^* - E(U^*|\tilde{x}, T_u = 1)][E(U^*|\tilde{x}, T_u = 1) - E(U^*|T_u = 1)] \\ &+ [E(U^*|\tilde{x}, T_u = 1) - E(U^*|T_u = 1)]^2 - B_0^u.\end{aligned}$$

By the chain rule and the quotient rule for derivatives, we obtain

$$\begin{aligned}\phi_1(\tilde{o}, \mathcal{P}) &= \frac{\phi(\tilde{o}, \mathcal{P})}{\sqrt{B_0^y B_0^z}} - \frac{\Psi_1(\mathcal{P})}{2B_0^y B_0^z}[\phi_y(\tilde{o}, \mathcal{P})B_0^z + B_0^y \phi_z(\tilde{o}, \mathcal{P})] \\ &= \frac{S_{yz}}{\sqrt{B_0^y B_0^z}} - \rho\frac{S_{yy}}{2B_0^y} - \rho\frac{S_{zz}}{2B_0^z}.\end{aligned}$$

Note that the variance of $\phi_1(O, \mathcal{P})$ is finite, which implies that the EIF for ρ is $\phi_1(O, \mathcal{P})$.

Proof of Theorem 5: Before giving the proof, a necessary lemma is presented.

Lemma 1.1 *Suppose Conditions 2.1–2.2 are satisfied. The asymptotically linear representation for $\hat{B}_0^y$ and $\hat{B}_0^z$ respectively are*

$$\begin{aligned}\hat{B}_0^y - B_0^y &= \sum_{i \in \mathcal{D}_2} S_{yy,i}^2 + o_p\left(\frac{1}{\sqrt{n}} + \frac{1}{\sqrt{n_y}}\right), \\ \hat{B}_0^z - B_0^z &= \sum_{i \in \mathcal{D}_2} S_{zz,i}^2 + o_p\left(\frac{1}{\sqrt{n}} + \frac{1}{\sqrt{n_z}}\right).\end{aligned}$$

Here

$$S_{yy,i}^2 = \frac{2}{n_y}\epsilon_i[m(X_i) - EY]\mathbb{I}(T_{y,i} = 1) + \frac{1}{n}\{[m(X_i) - EY]^2 - B_0^y\},$$

and

$$S_{zz,i}^2 = \frac{2}{n_z}\eta_i[h(X_i) - EZ]\mathbb{I}(T_{z,i} = 1) + \frac{1}{n}\{[h(X_i) - EZ]^2 - B_0^z\}.$$

Proof of Lemma 1.1: The proof of this lemma is similar to that of Theorem 3 and thus omitted here.

Note that

$$\begin{aligned}
\rho_n - \rho &= \Big(\frac{I_n}{\sqrt{\hat{B}_0^y \hat{B}_0^z}} - \frac{I}{\sqrt{B_0^y B_0^z}}\Big) \\
&= \frac{(I_n - I)}{\sqrt{B_0^y B_0^z}} + I\Big(\frac{1}{\sqrt{\hat{B}_0^y \hat{B}_0^z}} - \frac{1}{\sqrt{B_0^y B_0^z}}\Big) + (I_n - I)\Big(\frac{1}{\sqrt{\hat{B}_0^y \hat{B}_0^z}} - \frac{1}{\sqrt{B_0^y B_0^z}}\Big) \\
&=: C_1 + C_2 + C_3.
\end{aligned}$$

According to the proof of Theorem 3, we have

$$C_1 = \sum_{i \in \mathcal{D}_2} \frac{S_{yz,i}}{\sqrt{B_0^y B_0^z}} + o_p\left(\frac{1}{\sqrt{n_z}} + \frac{1}{\sqrt{n_y}}\right),$$

where

$$S_{yz,i} = \frac{1}{n_y}\epsilon_i[h(X_i) - EZ]\mathbb{I}(T_{y,i} = 1) + \frac{1}{n_z}\eta_i[m(X_i) - EY]I(D_{z,i} = 1) + \frac{1}{n}[m(X_i) - EY][h(X_i) - EZ].$$

Through the Taylor series expansion, we have

$$\begin{aligned}
\frac{1}{\sqrt{\hat{B}_0^y \hat{B}_0^z}} - \frac{1}{\sqrt{B_0^y B_0^z}} = &-\frac{1}{2}(B_0^y B_0^z)^{-3/2} B_0^z(\hat{B}_0^y - B_0^y) - \frac{1}{2}(B_0^y B_0^z)^{-3/2} B_0^y(\hat{B}_0^z - B_0^z) \\
&+ o_p(|\hat{B}_0^y - B_0^y|^2 + |\hat{B}_0^z - B_0^z|^2).
\end{aligned}$$

From Lemma 1.1, it follows that

$$\begin{aligned}
\frac{1}{\sqrt{\hat{B}_0^y \hat{B}_0^z}} - \frac{1}{\sqrt{B_0^y B_0^z}} &= -\frac{1}{2}(B_0^y B_0^z)^{-3/2} B_0^z(\hat{B}_0^y - B_0^y) - \frac{1}{2}(B_0^y B_0^z)^{-3/2} B_0^y(\hat{B}_0^z - B_0^z) + o_p\left(\frac{1}{\sqrt{n_z}} + \frac{1}{\sqrt{n_y}}\right) \\
&= -\frac{(B_0^y B_0^z)^{-3/2}}{2} \sum_{i \in \mathcal{D}_2} \{B_0^z S_{yy,i} + B_0^y S_{zz,i}\} + o_p\left(\frac{1}{\sqrt{n_z}} + \frac{1}{\sqrt{n_y}}\right).
\end{aligned}$$

Thus,

$$C_2 = -\frac{\rho}{2B_0^y B_0^z} \sum_{i \in \mathcal{D}_2} \{B_0^z S_{yy,i} + B_0^y S_{zz,i}\} + o_p\left(\frac{1}{\sqrt{n_z}} + \frac{1}{\sqrt{n_y}}\right).$$

Also note that $I_n - I = o_p(1)$. Hence it can be easily obtained that $C_3 = o_p\left(\frac{1}{\sqrt{n_z}} + \frac{1}{\sqrt{n_y}}\right)$. Combining the above results, we then conclude that

$$\rho_n - \rho = \sum_{i \in \mathcal{D}_2} \frac{S_{yz,i}}{\sqrt{B_0^y B_0^z}} - \rho\frac{S_{yy,i}}{2B_0^y} - \rho\frac{S_{zz,i}}{2B_0^z} + o_p\left(\frac{1}{\sqrt{n_z}} + \frac{1}{\sqrt{n_y}}\right).$$

Similarly for ρ_n', we have

$$\rho_n' - \rho = \sum_{i \in \mathcal{D}_1} \frac{S_{yz,i}}{\sqrt{B_0^y B_0^z}} - \rho \frac{S_{yy,i}}{2B_0^y} - \rho \frac{S_{zz,i}}{2B_0^z} + o_p\left(\frac{1}{\sqrt{n_z}} + \frac{1}{\sqrt{n_y}}\right).$$

Clearly, ρ_n and ρ_n' are asymptotically independent. Then we conclude that

$$\rho_n^* - \rho = \sum_{i \in \mathcal{D}} \frac{S_{yz,i}^*}{\sqrt{B_0^y B_0^z}} - \rho \frac{S_{yy,i}^*}{2B_0^y} - \rho \frac{S_{zz,i}^*}{2B_0^z} + o_p\left(\frac{1}{\sqrt{N_z}} + \frac{1}{\sqrt{N_y}}\right).$$

Here

$$\begin{aligned} S_{yz,i}^* = {} & \frac{1}{N_y}\epsilon_i[h(X_i) - EZ]\mathbb{I}(T_{y,i} = 1) + \frac{1}{N_z}\eta_i[m(X_i) - EY]I(D_{z,i} = 1) \\ & + \frac{1}{N}\{[m(X_i) - EY][h(X_i) - EZ] - I\}, \end{aligned}$$

and $S_{yy,i}^*$ and $S_{zz,i}^*$ are similarly defined.

Proof of Proposition 4.1: We first rewrite I as

$$I = \Psi(\mathcal{P}) = E_{\mathcal{P}}\{g_1^*(X)g_2^*(X)\},$$

where $g_1^*(X) = g_1[E_{\mathcal{P}}(Y^*|X, T_y = 1)] - E_{\mathcal{P}}\{g_1[E_{\mathcal{P}}(Y^*|X, T_y = 1)]\}$ and $g_2^*(X) = g_2[E_{\mathcal{P}}(Z^*|X, T_z = 1)] - E_{\mathcal{P}}\{g_2[E_{\mathcal{P}}(Z^*|X, T_z = 1)]\}$. Consider the same parametric submodel indexed by t, i.e.

$$\mathcal{P}_t = t\tilde{\mathcal{P}} + (1 - t)\mathcal{P},$$

where $t \in [0, 1]$, and $\tilde{\mathcal{P}}$ is a point mass at a single observation $\tilde{o}$. So that the efficient influence function (EIF) for I at observation $\tilde{o}$ directly as

$$\phi(\tilde{o}, \mathcal{P}) = \frac{d\Psi(\mathcal{P}_t)}{dt}|_{t=0} = E_{\mathcal{P}}\{\partial_t[g_{1t}^*(X)g_{2t}^*(X)]\} + \partial_t E_{\mathcal{P}_t}\{g_1^*(X)g_2^*(X)\}.$$

where $g_{1t}^*(X) = g_1[E_{\mathcal{P}_t}(Y^*|X, T_y = 1)] - E_{\mathcal{P}_t}\{g_1[E_{\mathcal{P}_t}(Y^*|X, T_y = 1)]\}$ and $g_{2t}^*(X) = g_2[E_{\mathcal{P}_t}(Z^*|X, T_z = 1)] - E_{\mathcal{P}_t}\{g_2[E_{\mathcal{P}_t}(Z^*|X, T_z = 1)]\}$.

Firstly, we focus on the term $\partial_t g_{1t}^*$. Note that

$$\begin{aligned} \partial_t g_{1t}^*(X) = {} & \partial_t\Big[g_1[E_{\mathcal{P}_t}(Y^*|X, T_y = 1)]\Big] - \partial_t\Big(\int g_1[E_{\mathcal{P}_t}(Y^*|X, T_y = 1)]f_t(X)dX\Big) \\ = {} & g_1'[E_{\mathcal{P}}(Y^*|X, T_y = 1)]\partial_t E_{\mathcal{P}_t}(Y^*|X, T_y = 1)] - g_1[E_{\mathcal{P}}(Y^*|\tilde{x}, T_y = 1)] \\ & + E_{\mathcal{P}}(g_1[E_{\mathcal{P}}(Y^*|X, T_y = 1)]) - E_{\mathcal{P}}\left(\partial_t\Big[g_1[E_{\mathcal{P}_t}(Y^*|X, T_y = 1)]\Big]\right) =: \psi_1(\tilde{o}, \mathcal{P}). \end{aligned}$$

Similarly, we can get

$$
\begin{aligned}
\partial_t g_{2t}^*(X) &= \partial_t \Big[g_2[E_{\mathcal{P}_t}(Z^*|X, T_z = 1)] \Big] - \partial_t \Big(\int g_2[E_{\mathcal{P}_t}(Z^*|X, T_z = 1)] f_t(X) dX \Big) \\
&= g_2'[E_{\mathcal{P}}(Z^*|X, T_z = 1)] \partial_t E_{\mathcal{P}_t}(Z^*|X, T_z = 1)] - g_2[E_{\mathcal{P}}(Z^*|\tilde{x}, T_z = 1)] \\
&\quad + E_{\mathcal{P}}(g_2[E_{\mathcal{P}}(Z^*|X, T_z = 1)]) - E_{\mathcal{P}} \Big(\partial_t \Big[g_1[E_{\mathcal{P}_t}(Z^*|X, T_z = 1)] \Big] \Big) =: \psi_2(\tilde{o}, \mathcal{P}).
\end{aligned}
$$

By the chain rule and the quotient rule for derivatives, it follows that

$$
\begin{aligned}
E_{\mathcal{P}}\{\partial_t [g_{1t}^*(X) g_{2t}^*(X)]\} &= E_{\mathcal{P}}\{\psi_1(\tilde{o}, \mathcal{P}) g_2^*(X)\} + E_{\mathcal{P}}\{g_1^*(X) \psi_2(\tilde{o}, \mathcal{P})\} \\
&= E_{\mathcal{P}}\{g_1'[E_{\mathcal{P}}(Y^*|X, T_y = 1)] \partial_t E_{\mathcal{P}_t}(Y^*|X, T_y = 1)] g_2^*(X)\} \\
&\quad + E_{\mathcal{P}}\{g_1^*(X) g_2'[E_{\mathcal{P}}(Z^*|X, T_z = 1)] \partial_t E_{\mathcal{P}_t}(Z^*|X, D_z)]\}.
\end{aligned}
$$

The last equality holds because $E_{\mathcal{P}}[g_i^*(X)] = 0$, $i = 1, 2$. According to Theorem 1, we have

$$
\partial_t E_{\mathcal{P}_t}(Y^*|X, T_y = 1) = \frac{\mathbb{I}_{(\tilde{x}, \tilde{t}_y)}(X, T_y = 1)}{f(X, T_y = 1)} [\tilde{y}^* - E(Y^*|X, T_y = 1)],
$$

which implies that

$$
\begin{aligned}
&E_{\mathcal{P}}\{g_1'[E_{\mathcal{P}_t}(Y^*|X, T_y = 1)] \partial_t E_{\mathcal{P}_t}(Y^*|X, T_y = 1)] g_2^*(X)\} \\
=& \frac{\mathbb{I}_{\tilde{t}_y}(T_y = 1)}{P(T_y = 1)} g_1'[E_{\mathcal{P}}(Y^*|\tilde{x}, T_y = 1)][\tilde{y}^* - E(Y^*|\tilde{x}, T_y = 1)] g_2^*(\tilde{x}),
\end{aligned}
$$

where the last equation holds because $T_y \perp X$.

Similarly, we have

$$
\begin{aligned}
&E_{\mathcal{P}}\{g_2'[E_{\mathcal{P}_t}(Z^*|X, T_z = 1)] \partial_t E_{\mathcal{P}_t}(Z^*|X, T_z = 1)] g_1^*(X)\} \\
=& \frac{\mathbb{I}_{\tilde{t}_z}(T_z = 1)}{P(T_z = 1)} g_2'[E_{\mathcal{P}}(Z^*|\tilde{x}, T_z = 1)][\tilde{z}^* - E(Z^*|\tilde{x}, T_z = 1)] g_1^*(\tilde{x}).
\end{aligned}
$$

Further, we have

$$
\partial_t E_{\mathcal{P}_t}\{g_1^*(X) g_2^*(X)\} = g_1^*(\tilde{x}) g_2^*(\tilde{x}) - \Psi(\mathcal{P}).
$$

Based on the above results, we obtain

$$
\begin{aligned}
\phi(\tilde{o}, \mathcal{P}) &= \frac{d\Psi(\mathcal{P}_t)}{dt}|_{t=0} \\
&= \frac{\mathbb{I}_{\tilde{t}_y}(T_y = 1)}{P(T_y = 1)} g_1'[E_{\mathcal{P}}(Y^*|\tilde{x}, T_y = 1)][\tilde{y}^* - E(Y^*|\tilde{x}, T_y = 1)] g_2^*(\tilde{x}) \\
&\quad + \frac{\mathbb{I}_{\tilde{t}_z}(T_z = 1)}{P(T_z = 1)} g_2'[E_{\mathcal{P}}(Z^*|\tilde{x}, T_z = 1)][\tilde{z}^* - E(Z^*|\tilde{x}, T_z = 1)] g_1^*(\tilde{x}) \\
&\quad + g_1^*(\tilde{x}) g_2^*(\tilde{x}) - \Psi(\mathcal{P}).
\end{aligned}
$$

Note that the variance of $\phi(O, \mathcal{P})$ is finite, which implies that the EIF for I is $\phi(O, \mathcal{P})$.

Given the EIF of I, we can get the results of ρ by using the similar arguments of the proof of Theorem 3.1 and thus omitted here.

Proof of Theorem 6: Let

$$I_n = \frac{1}{n_y}\sum_{i\in\mathcal{D}_{2y}} \hat{g}'_{1i}[Y_i - \hat{m}_i][\hat{g}_{2i} - \bar{g}_2] + \frac{1}{n_z}\sum_{i\in\mathcal{D}_{2z}} \hat{g}'_{2i}[Z_i - \hat{h}_i][\hat{g}_{1i} - \bar{g}_1] + \frac{1}{n}\sum_{i\in\mathcal{D}_2}[\hat{g}_{2i} - \bar{g}_2][\hat{g}_{1i} - \bar{g}_1]$$

$$=: I_{n1} + I_{n2} + I_{n3}.$$

For the term I_{n1}, we have

$$\begin{aligned} I_{n1} &= \frac{1}{n_y}\sum_{i\in\mathcal{D}_{2y}} \hat{g}'_{1i}[Y_i - \hat{m}_i][\hat{g}_{2i} - \bar{g}_2] \\ &= \frac{1}{n_y}\sum_{i\in\mathcal{D}_{2y}} \hat{g}'_{1i}[\epsilon_i + m_i - \hat{m}_i][\hat{g}_{2i} - g_{2i} + g_{2i} - Eg_2 + Eg_2 - \bar{g}_2] \\ &= \frac{1}{n_y}\sum_{i\in\mathcal{D}_{2y}} \hat{g}'_{1i}\epsilon_i(\hat{g}_{2i} - g_{2i}) + \frac{1}{n_y}\sum_{i\in\mathcal{D}_{2y}} \hat{g}'_{1i}\epsilon_i(g_{2i} - Eg_2) + \frac{1}{n_y}\sum_{i\in\mathcal{D}_{2y}} \hat{g}'_{1i}\epsilon_i(Eg_2 - \bar{g}_2) \\ &\quad + \frac{1}{n_y}\sum_{i\in\mathcal{D}_{2y}} \hat{g}'_{1i}(m_i - \hat{m}_i)(\hat{g}_{2i} - \bar{g}_2) =: \sum_{i=1}^{4} D_i. \end{aligned}$$

In the following part, we will show that both D_1 and D_3 are negligible, while both D_2 and D_4 contain the leading terms.

For the term D_1, it can be rewritten as

$$D_1 = \frac{1}{n_y}\sum_{i\in\mathcal{D}_{2y}} (\hat{g}'_{1i} - g'_{1i})\epsilon_i(\hat{g}_{2i} - g_{2i}) + \frac{1}{n_y}\sum_{i\in\mathcal{D}_{2y}} g'_{1i}\epsilon_i(\hat{g}_{2i} - g_{2i}) =: D_{11} + D_{12}.$$

For the first term, we have $E(D_{11}) = 0$, and

$$E(D_{11}^2|\mathcal{D}_1) = \frac{1}{n_y^2}\sum_{i\in\mathcal{D}_{2y}} E\left[\sigma^2(X)(\hat{g}'_{1i} - g'_{1i})^2(\hat{g}_{2i} - g_{2i})^2|\mathcal{D}_1\right].$$

Under the conditions $|g'_1(m_i) - g'_1(\hat{m}_i)| \le L|m_i - \hat{m}_i|$ and $g'_2(\cdot) < C < \infty$, we have

$$E(D_{11}^2) \lesssim \frac{1}{n_y}E\left[\sigma^2(X)(\hat{m}_i - m_i)^2(\hat{h}_i - h_i)^2\right] = o\left(\frac{1}{n_y}\right),$$

where the last equation holds under the condition $E\left[\sigma^2(X)(\hat{m}_i - m_i)^2(\hat{h}_i - h_i)^2\right] = o(1)$. Further, we can similarly show that $E(D_{12}) = 0$, and

$$E(D_{12}^2) \lesssim \frac{1}{n_y}E\left[\sigma^2(X)(\hat{h}_i - h_i)^2\right] = o\left(\frac{1}{n_y}\right),$$

when the conditions $E[(\hat{h}_i - h_i)^4] = o(n_y^{-1})$ and $g_2'(\cdot) < C < \infty$ are satisfied. It follows that

$$D_1 = D_{11} + D_{12} = o_p\left(\frac{1}{\sqrt{n_y}}\right).$$

For the term D_2, it can be rewritten as

$$D_2 = \frac{1}{n_y}\sum_{i\in\mathcal{D}_{2y}} (\hat{g}_{1i}' - g_{1i}')\epsilon_i(g_{2i} - Eg_2) + \frac{1}{n_y}\sum_{i\in\mathcal{D}_{2y}} g_{1i}'\epsilon_i(g_{2i} - Eg_2)$$

Under the conditions $|g_1'(m_i) - g_1'(\hat{m}_i)| \le L|m_i - \hat{m}_i|$ and $E[\sigma^2(X)(\hat{m}_i - m_i)^2(g_2 - Eg_2)^2] = o(1)$, it can similarly obtained that

$$D_2 = \frac{1}{n_y}\sum_{i\in\mathcal{D}_{2y}} g_{1i}'\epsilon_i(g_{2i} - Eg_2) + o_p\left(\frac{1}{\sqrt{n_y}}\right). \tag{S.3}$$

Now we turn to consider $Eg_2 - \bar{g}_2$. Under conditions that $g_2'(\cdot) \le C < \infty$ and $E[(h_i - \hat{h}_i)^4] = o_p(n_z^{-1})$, we have

$$\begin{aligned} Eg_2 - \bar{g}_2 &= Eg_2 - \frac{1}{n_z}\sum_{i\in\mathcal{D}_{2z}} g_2(h_i) + \frac{1}{n_z}\sum_{i\in\mathcal{D}_{2z}} g_2(h_i) - g_2(\hat{h}_i) \\ &= Eg_2 - \frac{1}{n_z}\sum_{i\in\mathcal{D}_{2z}} g_2(h_i) + \frac{1}{n_z}\sum_{i\in\mathcal{D}_{2z}} g_2'(\tilde{h}_i)(h_i - \hat{h}_i) = o_p(n_z^{-1/4}), \end{aligned}$$

where $\tilde{h}_i$ is between h_i and $\hat{h}_i$. Similarly, we can get $Eg_1 - \bar{g}_1 = o_p(n_y^{-1/4})$.

For the term D_3, it can be rewritten as

$$D_3 = (Eg_2 - \bar{g}_2)\frac{1}{n_y}\sum_{i\in\mathcal{D}_{2y}} \hat{g}_{1i}'\epsilon_i.$$

Given $\mathcal{D}_{1y}$, it can be easily to show that

$$\frac{1}{n_y}\sum_{i\in\mathcal{D}_{2y}} \hat{g}_{1i}'\epsilon_i = O_p\left(\frac{1}{\sqrt{n_y}}\right).$$

Also based on the fact $Eg_2 - \bar{g}_2 = o_p(n_z^{-1/4})$, we easily get that $D_3 = o_p(n_y^{-1/2})$.

So far, we have

$$I_{n1} = \frac{1}{n_y}\sum_{i\in\mathcal{D}_{2y}} g_{1i}'\epsilon_i(g_{2i} - Eg_2) + D_4 + o_p\left(\frac{1}{\sqrt{n_y}}\right).$$

Next, we consider the bound for D_4. Note that

$$\begin{aligned} D_4 &= \frac{1}{n_y}\sum_{i\in\mathcal{D}_{2y}} \hat{g}_{1i}'(m_i - \hat{m}_i)(\hat{g}_{2i} - \bar{g}_2) \\ &= \frac{1}{n_y}\sum_{i\in\mathcal{D}_{2y}} (\hat{g}_{1i}' - g_{1i}')(m_i - \hat{m}_i)(\hat{g}_{2i} - \bar{g}_2) + \frac{1}{n_y}\sum_{i\in\mathcal{D}_{2y}} g_{1i}'(m_i - \hat{m}_i)(\hat{g}_{2i} - \bar{g}_2) \\ &=: D_{41} + D_{42}. \end{aligned}$$

For the term D_{41}, we have

$$D_{41} = \frac{1}{n_y} \sum_{i \in \mathcal{D}_{2y}} (\hat{g}'_{1i} - g'_{1i})(m_i - \hat{m}_i)(\hat{g}_{2i} - g_{2i}) + \frac{1}{n_y} \sum_{i \in \mathcal{D}_{2y}} (\hat{g}'_{1i} - g'_{1i})(m_i - \hat{m}_i)(g_{2i} - Eg_2)$$
$$+ (Eg_2 - \bar{g}_2) \frac{1}{n_y} \sum_{i \in \mathcal{D}_{2y}} (\hat{g}'_{1i} - g'_{1i})(m_i - \hat{m}_i).$$

For the first term,

$$\begin{aligned} & \frac{1}{n_y} \sum_{i \in \mathcal{D}_{2y}} (\hat{g}'_{1i} - g'_{1i})(m_i - \hat{m}_i)(\hat{g}_{2i} - g_{2i}) \\ \leq & \left[\frac{1}{n_y} \sum_{i \in \mathcal{D}_{2y}} (\hat{g}'_{1i} - g'_{1i})^2 (m_i - \hat{m}_i)^2 \right]^{1/2} \left[\frac{1}{n_y} \sum_{i \in \mathcal{D}_{2y}} (\hat{g}_{2i} - g_{2i})^2 \right]^{1/2} \\ \lesssim & \left[\frac{1}{n_y} \sum_{i \in \mathcal{D}_{2y}} (m_i - \hat{m}_i)^4 \right]^{1/2} \left[\frac{1}{n_y} \sum_{i \in \mathcal{D}_{2y}} (\hat{h}_i - h_i)^2 \right]^{1/2} = o_p\left(\frac{1}{\sqrt{n_y}} \right), \end{aligned}$$

where the last equation holds when $E[(\hat{m}_i - m_i)^4] = o(n_y^{-1})$ and $E[(\hat{h}_i - h_i)^4] = (n_z^{-1})$. Similarly, we can apply the same arguments to the other two terms. Thus, it follows that

$$D_{41} = o_p\left(\frac{1}{\sqrt{n_y}} \right).$$

Next, we turn to the term D_{42}. Note that

$$D_{42} = \frac{1}{n_y} \sum_{i \in \mathcal{D}_{2y}} g'_{1i}(m_i - \hat{m}_i)(\hat{g}_{2i} - \bar{g}_2) = \frac{1}{n_y} \sum_{i \in \mathcal{D}_{2y}} g'_{1i}(m_i - \hat{m}_i)(\hat{g}_{2i} - g_{2i})$$
$$+ \frac{1}{n_y} \sum_{i \in \mathcal{D}_{2y}} g'_{1i}(m_i - \hat{m}_i)(g_{2i} - Eg_2) + (Eg_2 - \bar{g}_2) \frac{1}{n_y} \sum_{i \in \mathcal{D}_{2y}} g'_{1i}(m_i - \hat{m}_i) =: \sum_{i=1}^{3} D_{42i}.$$

For the first term D_{421},

$$\begin{aligned} |D_{421}| \leq & \frac{1}{n_y} \sum_{i \in \mathcal{D}_{2y}} |g'_{1i}(m_i - \hat{m}_i)(\hat{g}_{2i} - g_{2i})| \\ \leq & \left[\frac{1}{n_y} \sum_{i \in \mathcal{D}_{2y}} (g'_{1i})^2 (m_i - \hat{m}_i)^2 \right]^{1/2} \left[\frac{1}{n_y} \sum_{i \in \mathcal{D}_{2y}} (\hat{g}_{2i} - g_{2i})^2 \right]^{1/2} \\ \lesssim & \left[\frac{1}{n_y} \sum_{i \in \mathcal{D}_{2y}} (m_i - \hat{m}_i)^2 \right]^{1/2} \left[\frac{1}{n_y} \sum_{i \in \mathcal{D}_{2y}} (\hat{h}_i - h_i)^2 \right]^{1/2}. \end{aligned}$$

Note that under condition 4.3, we have $E[(\hat{m}_i - m_i)^2] = o(n_y^{-1/2})$ and $E[(\hat{h}_i - h_i)^2] = o(n_z^{-1/2})$. It follows that

$$D_{421} = o_p\left(\frac{1}{n_y^{1/4} n_z^{1/4}}\right) = o_p\left(\frac{1}{\sqrt{n_y}} + \frac{1}{\sqrt{n_z}}\right).$$

Similarly, based on $Eg_2 - \bar{g}_2 = o_p(n_z^{-1/4})$ and $E[(\hat{m}_i - m_i)^2] = o(n_y^{-1/2})$, it can be easily obtained that

$$D_{423} = o_p\left(\frac{1}{\sqrt{n_y}} + \frac{1}{\sqrt{n_z}}\right).$$

It follows that

$$D_4 = \frac{1}{n_y} \sum_{i \in \mathcal{D}_{2y}} g_{1i}'(m_i - \hat{m}_i)(g_{2i} - Eg_2) + o_p\left(\frac{1}{\sqrt{n_y}} + \frac{1}{\sqrt{n_z}}\right).$$

Thus, we can conclude that

$$I_{n1} = \frac{1}{n_y} \sum_{i \in \mathcal{D}_{2y}} g_{1i}'\epsilon_i(g_{2i} - Eg_2) + \frac{1}{n_y} \sum_{i \in \mathcal{D}_{2y}} g_{1i}'(m_i - \hat{m}_i)(g_{2i} - Eg_2) + o_p\left(\frac{1}{\sqrt{n_y}} + \frac{1}{\sqrt{n_z}}\right).$$

Similarly, we can get

$$I_{n2} = \frac{1}{n_z} \sum_{i \in \mathcal{D}_{2z}} g_{2i}'\eta(g_{1i} - Eg_1) + \frac{1}{n_z} \sum_{i \in \mathcal{D}_{2z}} g_{2i}'(h_i - \hat{h}_i)(g_{1i} - Eg_1) + o_p\left(\frac{1}{\sqrt{n_y}} + \frac{1}{\sqrt{n_z}}\right).$$

For the term I_{n3}, we have

$$\begin{aligned}
I_{n3} &= \frac{1}{n} \sum_{i \in \mathcal{D}_2} (\hat{g}_{2i} - g_{2i} + g_{2i} - Eg_2 + Eg_2 - \bar{g}_2)(\hat{g}_{1i} - g_{1i} + g_{1i} - Eg_1 + Eg_1 - \bar{g}_1) \\
&= \frac{1}{n} \sum_{i \in \mathcal{D}_2} (g_{2i} - Eg_2)(g_{1i} - Eg_1) + \frac{1}{n} \sum_{i \in \mathcal{D}_2} (\hat{g}_{2i} - g_{2i})(g_{1i} - Eg_1) + \frac{1}{n} \sum_{i \in \mathcal{D}_2} (\hat{g}_{1i} - g_{1i})(g_{2i} - Eg_2) \\
&\quad + \frac{1}{n} \sum_{i \in \mathcal{D}_2} (\hat{g}_{2i} - g_{2i})(\hat{g}_{1i} - g_{1i}) + (Eg_1 - \bar{g}_1)\frac{1}{n} \sum_{i \in \mathcal{D}_2} (\hat{g}_{2i} - g_{2i}) + (Eg_1 - \bar{g}_1)\frac{1}{n} \sum_{i \in \mathcal{D}_2} (g_{2i} - Eg_2) \\
&\quad + (Eg_1 - \bar{g}_1)(Eg_2 - \bar{g}_2) + (Eg_2 - \bar{g}_2)\frac{1}{n} \sum_{i \in \mathcal{D}_2} (\hat{g}_{1i} - g_{1i}) + (Eg_2 - \bar{g}_2)\frac{1}{n} \sum_{i \in \mathcal{D}_2} (g_{1i} - Eg_1) =: \sum_{i=1}^{9} I_{n3i}.
\end{aligned}$$

For the term I_{n34},

$$\begin{aligned}
|I_{n34}| &\leq \frac{1}{n} \sum_{i \in \mathcal{D}_2} |(\hat{g}_{2i} - g_{2i})(\hat{g}_{1i} - g_{1i})| \leq \left[\frac{1}{n} \sum_{i \in \mathcal{D}_2} (\hat{g}_{2i} - g_{2i})^2\right]^{1/2} \left[\frac{1}{n} \sum_{i \in \mathcal{D}_2} (\hat{g}_{1i} - g_{1i})^2\right]^{1/2} \\
&\lesssim \left[\frac{1}{n} \sum_{i \in \mathcal{D}_2} (\hat{h}_i - h_i)^2\right]^{1/2} \left[\frac{1}{n} \sum_{i \in \mathcal{D}_2} (\hat{m}_i - m_i)^2\right]^{1/2} = o_p\left(\frac{1}{n_y^{1/4} n_z^{1/4}}\right) = o_p\left(\frac{1}{\sqrt{n_y}} + \frac{1}{\sqrt{n_z}}\right).
\end{aligned}$$

Similarly, it can be easy to show that

$$\frac{1}{n}\sum_{i\in\mathcal{D}_2}(\hat{g}_{1i}-g_{1i})=o_p(n_y^{-1/4}),\ \text{ and } \frac{1}{n}\sum_{i\in\mathcal{D}_2}(\hat{g}_{2i}-g_{2i})=o_p(n_z^{-1/4}).$$

Also, based on the facts that $Eg_1-\bar{g}_1=o_p(n_y^{-1/4})$ and $Eg_2-\bar{g}_2=o_p(n_z^{-1/4})$, it can be easily to show that

$$\begin{aligned}I_{n35}+I_{n37}+I_{n38}&=(Eg_1-\bar{g}_1)\frac{1}{n}\sum_{i\in\mathcal{D}_2}(\hat{g}_{2i}-g_{2i})+(Eg_1-\bar{g}_1)(Eg_2-\bar{g}_2)\\&\quad+(Eg_2-\bar{g}_2)\frac{1}{n}\sum_{i\in\mathcal{D}_2}(\hat{g}_{1i}-g_{1i})=o_p\left(\frac{1}{\sqrt{n_y}}+\frac{1}{\sqrt{n_z}}\right).\end{aligned}$$

Further, note that

$$\frac{1}{n}\sum_{i\in\mathcal{D}_2}(g_{2i}-Eg_2)=O_p(n^{-1/2}),\ \text{ and } \frac{1}{n}\sum_{i\in\mathcal{D}_2}(g_{1i}-Eg_1)=O_p(n^{-1/2}),$$

which implies that

$$I_{n36}+I_{n39}=(Eg_1-\bar{g}_1)\frac{1}{n}\sum_{i\in\mathcal{D}_2}(g_{2i}-Eg_2)+(Eg_2-\bar{g}_2)\frac{1}{n}\sum_{i\in\mathcal{D}_2}(g_{1i}-Eg_1)=o_p\left(\frac{1}{\sqrt{n_y}}+\frac{1}{\sqrt{n_z}}\right).$$

Note that by the Taylor series expansion

$$\begin{aligned}I_{n32}&=\frac{1}{n}\sum_{i\in\mathcal{D}_2}(\hat{g}_{2i}-g_{2i})(g_{1i}-Eg_1)\\&=\frac{1}{n}\sum_{i\in\mathcal{D}_2}g_2'(h_i)(\hat{h}_i-h_i)(g_{1i}-Eg_1)+O_p\left(\frac{1}{n}\sum_{i\in\mathcal{D}_2}(\hat{h}_i-h_i)^2(g_{1i}-Eg_1)\right)\\&=\frac{1}{n}\sum_{i\in\mathcal{D}_2}g_{2i}'(\hat{h}_i-h_i)(g_{1i}-Eg_1)+o_p\left(\frac{1}{\sqrt{n_z}}\right).\end{aligned}$$

The last equation holds when $E[(\hat{h}_i-h_i)^4]=o(n_z^{-1})$ and condition 4.2 are satisfied. Similarly, we can get

$$\begin{aligned}I_{n33}&=\frac{1}{n}\sum_{i\in\mathcal{D}_2}(\hat{g}_{1i}-g_{1i})(g_{2i}-Eg_2)\\&=\frac{1}{n}\sum_{i\in\mathcal{D}_2}g_{1i}'(\hat{m}_i-m_i)(g_{2i}-Eg_2)+O_p\left(\frac{1}{n}\sum_{i\in\mathcal{D}_2}(\hat{m}_i-m_i)^2(g_{2i}-Eg_2)\right)\\&=\frac{1}{n}\sum_{i\in\mathcal{D}_2}g_{1i}'(\hat{m}_i-m_i)(g_{2i}-Eg_2)+o_p\left(\frac{1}{\sqrt{n_y}}\right).\end{aligned}$$

Based on the above results, we can conclude that

$$
\begin{aligned}
I_n &= \frac{1}{n_y}\sum_{i\in\mathcal{D}_{2y}} g'_{1i}\epsilon_i(g_{2i}-Eg_2) + \frac{1}{n_z}\sum_{i\in\mathcal{D}_{2z}} g'_{2i}\eta(g_{1i}-Eg_1) + \frac{1}{n}\sum_{i\in\mathcal{D}_2}(g_{2i}-Eg_2)(g_{1i}-Eg_1)\\
&+ \frac{1}{n_z}\sum_{i\in\mathcal{D}_{2z}} g'_{2i}(h_i-\hat{h}_i)(g_{1i}-Eg_1) + \frac{1}{n}\sum_{i\in\mathcal{D}_2} g'_{2i}(\hat{h}_i-h_i)(g_{1i}-Eg_1)\\
&+ \frac{1}{n_y}\sum_{i\in\mathcal{D}_{2y}} g'_{1i}(m_i-\hat{m}_i)(g_{2i}-Eg_2) + \frac{1}{n}\sum_{i\in\mathcal{D}_2} g'_{1i}(\hat{m}_i-m_i)(g_{2i}-Eg_2) + o_p\left(\frac{1}{\sqrt{n_y}}+\frac{1}{\sqrt{n_z}}\right)
\end{aligned}
$$

Denote

$$
R_1 = \frac{1}{n_z}\sum_{i\in\mathcal{D}_{2z}} g'_{2i}(h_i-\hat{h}_i)(g_{1i}-Eg_1) + \frac{1}{n}\sum_{i\in\mathcal{D}_2} g'_{2i}(\hat{h}_i-h_i)(g_{1i}-Eg_1).
$$

Further let $l_i = g'_{2i}(\hat{h}_i-h_i)(g_{1i}-Eg_1)$. Note that

$$
R_1 = \frac{1}{n}\sum_{i\in\mathcal{D}_2} l_i - \frac{1}{n_z}\sum_{i\in\mathcal{D}_{2z}} l_i = \frac{n-n_z}{n}\left(\frac{1}{n-n_z}\sum_{i\in\mathcal{D}_2/\mathcal{D}_{2z}} l_i - \frac{1}{n_z}\sum_{i\in\mathcal{D}_{2z}} l_i\right).
$$

Similarly as the discussion in the proof of Theorem 3, it can be obtained that $E(R_1)=0$, and

$$
\mathrm{Var}(R_1) \le \frac{2(n-n_z)^2}{n^2}\left(\frac{1}{n-n_z}+\frac{1}{n_z}\right)E(l_i^2).
$$

Note that $El_i^2 = o(1)$, we then derive that

$$
R_1 = \frac{1}{n_z}\sum_{i\in\mathcal{D}_{2z}} g'_{2i}(h_i-\hat{h}_i)(g_{1i}-Eg_1) + \frac{1}{n}\sum_{i\in\mathcal{D}_2} g'_{2i}(\hat{h}_i-h_i)(g_{1i}-Eg_1) = o_p\left(\frac{1}{\sqrt{n_z}}\right).
$$

Similarly, we can show that

$$
\frac{1}{n_y}\sum_{i\in\mathcal{D}_{2y}} g'_{1i}(m_i-\hat{m}_i)(g_{2i}-Eg_2) + \frac{1}{n}\sum_{i\in\mathcal{D}_2} g'_{1i}(\hat{m}_i-m_i)(g_{2i}-Eg_2) = o_p\left(\frac{1}{\sqrt{n_y}}\right).
$$

Thus, we have

$$
\begin{aligned}
I_n &= \frac{1}{n_y}\sum_{i\in\mathcal{D}_{2y}} g'_{1i}\epsilon_i(g_{2i}-Eg_2) + \frac{1}{n_z}\sum_{i\in\mathcal{D}_{2z}} g'_{2i}\eta(g_{1i}-Eg_1)\\
&+ \frac{1}{n}\sum_{i\in\mathcal{D}_2}(g_{2i}-Eg_2)(g_{1i}-Eg_1) + o_p\left(\frac{1}{\sqrt{n_y}}+\frac{1}{\sqrt{n_z}}\right).
\end{aligned}
$$

Similarly for I'_n, we have:

$$
\begin{aligned}
I'_n &= \frac{1}{n_y}\sum_{i\in\mathcal{D}_{1y}} g'_{1i}\epsilon_i(g_{2i}-Eg_2) + \frac{1}{n_z}\sum_{i\in\mathcal{D}_{1z}} g'_{2i}\eta(g_{1i}-Eg_1)\\
&+ \frac{1}{n}\sum_{i\in\mathcal{D}_1}(g_{2i}-Eg_2)(g_{1i}-Eg_1) + o_p\left(\frac{1}{\sqrt{n_y}}+\frac{1}{\sqrt{n_z}}\right).
\end{aligned}
$$

Based on the above results, we can conclude that

$$I_n^* = \frac{1}{N_y}\sum_{i\in\mathcal{D}_y} g_{1i}'\epsilon_i(g_{2i} - Eg_2) + \frac{1}{N_z}\sum_{i\in\mathcal{D}_z} g_{2i}'\eta(g_{1i} - Eg_1) + \frac{1}{N}\sum_{i\in\mathcal{D}}(g_{2i} - Eg_2)(g_{1i} - Eg_1)$$
$$+ o_p\left(\frac{1}{\sqrt{N_y}} + \frac{1}{\sqrt{N_z}}\right).$$

Given the asymptotic results of I_n^*, we can get the results of ρ_n^* by using similar arguments of the proof of Theorem 5 and thus omitted here.