Notes: Double/Debiased Machine Learning

Ziyang Gong

Consider the partical linear regression model

$$\begin{aligned} Y_i &= D_i\theta + g(\mathbf{X}_i) + U_i, \quad E\left(U_i \mid \mathbf{X}_i, D_i\right) = 0 \\ D_i &= m(\mathbf{X}_i) + V_i, \quad E\left(V_i \mid \mathbf{X}_i\right) = 0 \end{aligned} \tag{1}$$

where $\mathbf{X}_i \in \mathbb{R}^p$, $i = 1, 2, \ldots, n$.

Consider the following score functions for partical linear regression, that,

1. IV-type: The score function of IV-type is

$$\psi\left(\mathbf{W}; \theta, \boldsymbol{\eta}\right) := \left[Y - D\theta - g(\mathbf{X})\right]\left[D - m(\mathbf{X})\right] \tag{2}$$

   where $\boldsymbol{\eta} = (g, m)$. Suppose $\hat{g}$ and $\hat{m}$ are estimated using the auxiliary samples indexed by I^c, Thus

$$\check{\theta} = \left(\frac{1}{n}\sum_{i \in \mathrm{I}} \widehat{V}_i D_i\right)^{-1} \frac{1}{n}\sum_{i \in \mathrm{I}} \hat{V}_i \left(Y_i - \hat{g}\left(X_i\right)\right) \tag{3}$$

   where $\hat{V} = D - \hat{m}(\mathbf{X})$.

2. Partialling-Out: The score function of partialling-out is

$$\psi\left(\mathbf{W}; \theta, \boldsymbol{\eta}\right) := \left[Y - \ell\left(\mathbf{X}\right) - \theta\left(D - m(\mathbf{X})\right)\right]\left[D - m(\mathbf{X})\right] \tag{4}$$

   where $\boldsymbol{\eta} = (\ell, m)$. Suppose $\hat{\ell}$ and $\hat{m}$ are estimated using the auxiliary samples indexed by I^c, Thus

$$\check{\theta} = \left(\frac{1}{n}\sum_{i \in \mathrm{I}} \widehat{V}_i^2\right)^{-1} \frac{1}{n}\sum_{i \in \mathrm{I}} \hat{V}_i \left(Y_i - \hat{\ell}\left(X_i\right)\right) \tag{5}$$

   where $\hat{V} = D - \hat{m}(\mathbf{X})$.

In this note, we use samples $n = 200$ and repeat 1000 times to evaluate the estimation of DoubleML method with different machine learning models (`lasso` and `Random Forest`), and the code is avaliable at https://github.com/SignorinoY/ultralisk.

1 Linear Models

Suppose

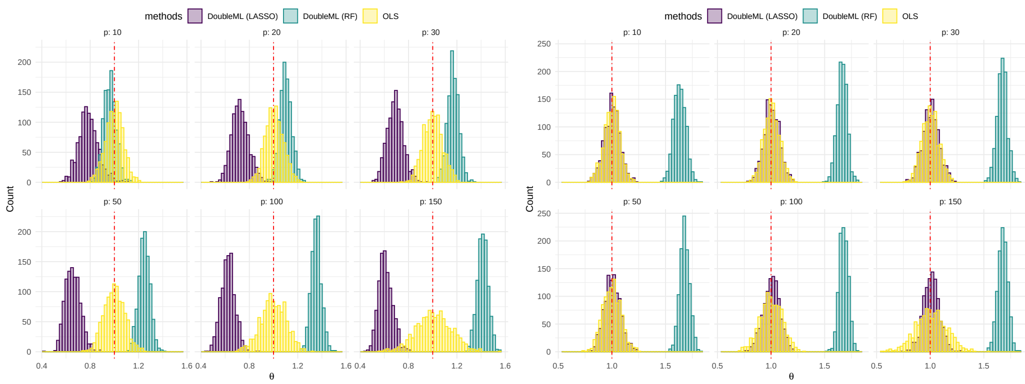
$$
\begin{gathered}
m(\mathbf{X}_i) = \textstyle\sum_{j=1}^{3} \alpha_j x_{i,j}, \quad g(\mathbf{X}_i) = \sum_{j=1}^{3} \beta_j x_{i,j} \\
\boldsymbol{\alpha} = (3,2,1)', \quad \boldsymbol{\beta} = (1,2,3)'
\end{gathered} \tag{6}
$$

or

$$
\begin{gathered}
m(\mathbf{X}_i) = \textstyle\sum_{j=1}^{6} \alpha_j x_{i,j}, \quad g(\mathbf{X}_i) = \sum_{j=1}^{6} \beta_j x_{i,j} \\
\boldsymbol{\alpha} = (6,5,4,3,2,1)', \quad \boldsymbol{\beta} = (1,2,3,4,5,6)'
\end{gathered} \tag{7}
$$

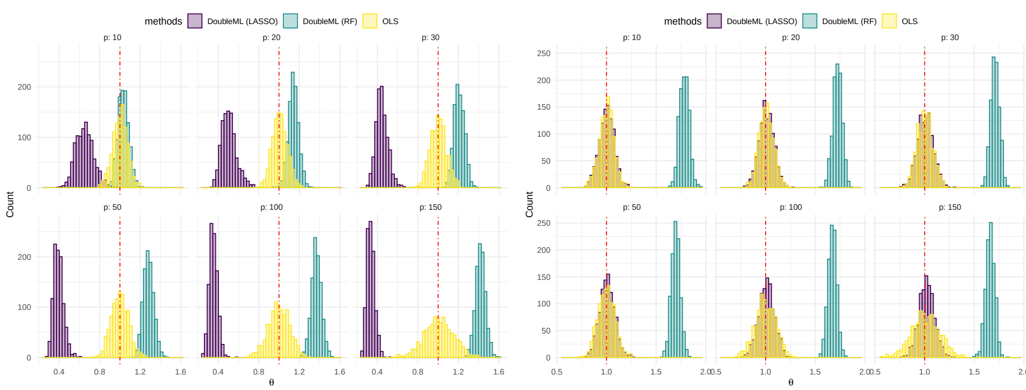
where

$$\mathbf{X}_i \sim N\left(\mathbf{0}, \mathrm{I}_p\right) \text{ and } U_i, V_i \sim N\left(0,1\right), \quad i = 1,2,\ldots,n.$$



(a) IV-type **(b)** Partialling-Out

Figure 1: Linear Models for (6)



(a) IV-type **(b)** Partialling-Out

Figure 2: Linear Models for (7)

2 Relu Models

Suppose

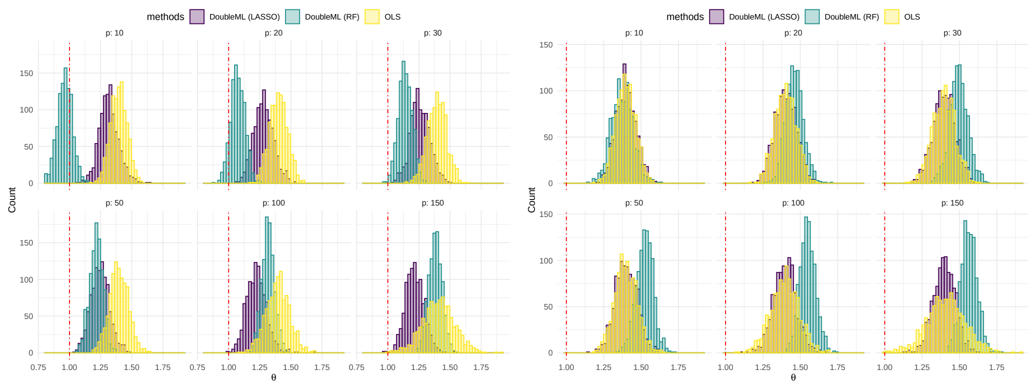
$$\begin{array}{c} m(\mathbf{X}_i) = \sum_{j=1}^{3} \alpha_j \max\left(x_{i,j}, 0\right), \quad g(\mathbf{X}_i) = \sum_{j=1}^{3} \beta_j \max\left(x_{i,j}, 0\right) \\ \boldsymbol{\alpha} = (3,2,1)', \quad \boldsymbol{\beta} = (1,2,3)' \end{array} \tag{8}$$

or

$$\begin{array}{c} m(\mathbf{X}_i) = \sum_{j=1}^{6} \alpha_j \max\left(x_{i,j}, 0\right), \quad g(\mathbf{X}_i) = \sum_{j=1}^{6} \beta_j \max\left(x_{i,j}, 0\right) \\ \boldsymbol{\alpha} = (6,5,4,3,2,1)', \quad \boldsymbol{\beta} = (1,2,3,4,5,6)' \end{array} \tag{9}$$

where

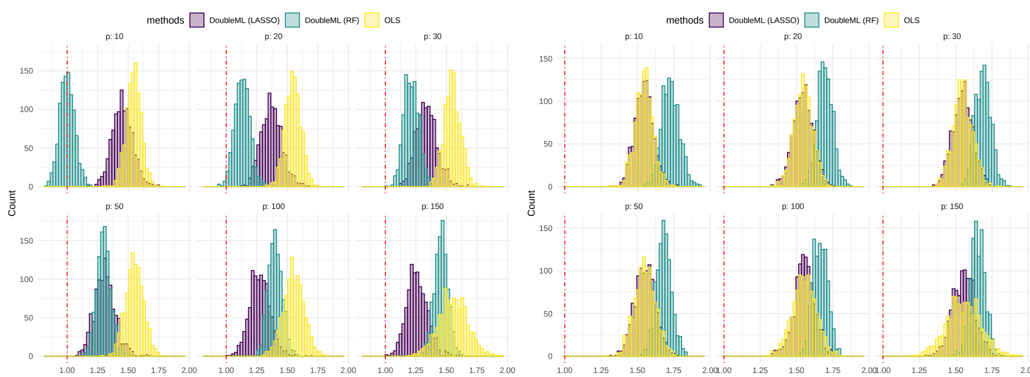
$$\mathbf{X}_i \sim N\left(\mathbf{0}, \mathrm{I}_p\right) \text{ and } U_i, V_i \sim N\left(0,1\right), \quad i = 1, 2, \ldots, n.$$



(a) IV-type

(b) Partialling-Out

Figure 3: Relu Models for (8)



(a) IV-type

(b) Partialling-Out

Figure 4: Relu Models for (9)

3 Polynomial Models

Suppose

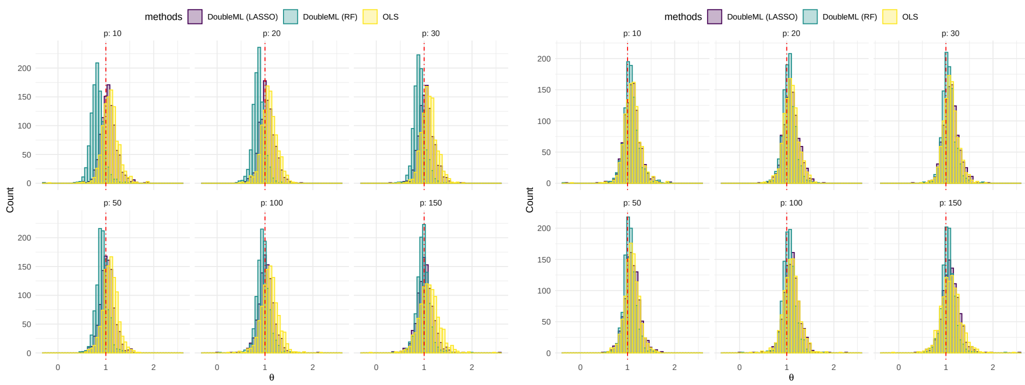
$$\begin{aligned} m(\mathbf{X}_i) =& x_{i,1}^3 + x_{i,2}^2 + x_{i,1}x_{i,2} + x_{i,2}x_{i,3} \\ g(\mathbf{X}_i) =& x_{i,2}^3 + x_{i,3}^2 + x_{i,1}x_{i,3} + x_{i,2}x_{i,3} \end{aligned} \tag{10}$$

or

$$m(\mathbf{X}_i) = \left(\sum_{j=1}^{3} x_{i,j}\right)^3 + \left(\sum_{j=4}^{6} x_{i,j}\right)^2, \quad g(\mathbf{X}_i) = \left(\sum_{j=1}^{3} x_{i,j}\right)^2 + \left(\sum_{j=4}^{6} x_{i,j}\right)^3 \tag{11}$$

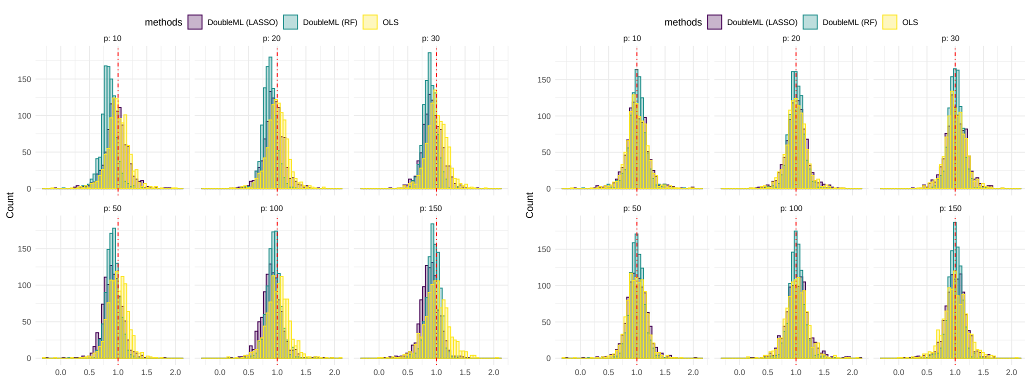
where

$$\mathbf{X}_i \sim N\left(\mathbf{0}, \mathrm{I}_p\right) \text{ and } U_i, V_i \sim N\left(0, 1\right), \quad i = 1, 2, \ldots, n.$$



(a) IV-type **(b)** Partialling-Out

Figure 5: Polynomial Models for (10)



(a) IV-type **(b)** Partialling-Out

Figure 6: Polynomial Models for (11)

4 Logistic-like Models

Suppose

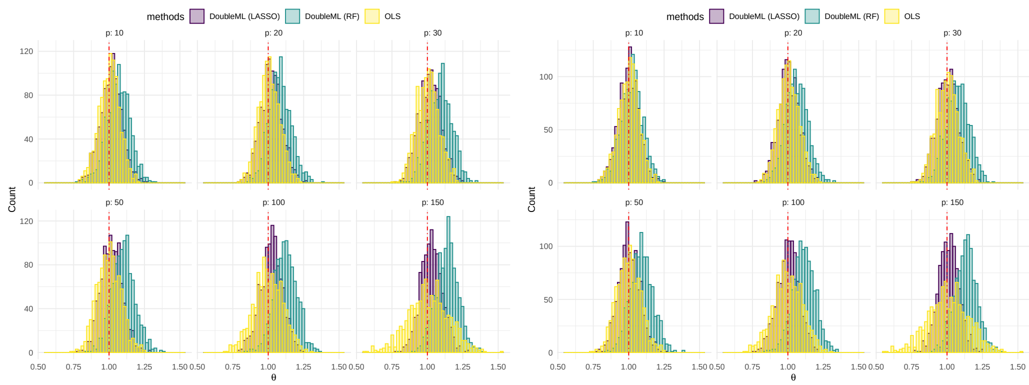
$$
\begin{aligned}
m\left(\mathbf{X}_i\right) &= \frac{\exp\left(x_{i,1}+x_{i,2}\right)}{1+\exp\left(x_{i,1}+x_{i,2}\right)} + \frac{1}{4}\cdot x_{i,3} \\
g\left(\mathbf{X}_i\right) &= x_{i,1} + \frac{1}{4}\cdot \frac{\exp\left(x_{i,2}+x_{i,3}\right)}{1+\exp\left(x_{i,2}+x_{i,3}\right)}
\end{aligned} \tag{12}
$$

or

$$
\begin{aligned}
m\left(\mathbf{X}_i\right) &= \left[\frac{\exp\left(x_{i,1}+x_{i,2}\right)}{1+\exp\left(x_{i,1}+x_{i,2}\right)} + \frac{\exp\left(x_{i,4}+x_{i,5}\right)}{1+\exp\left(x_{i,4}+x_{i,5}\right)}\right] + \frac{1}{4}\left(x_{i,3}+x_{i,6}\right) \\
g\left(\mathbf{X}_i\right) &= \left(x_{i,1}+x_{i,4}\right) + \frac{1}{4}\left[\frac{\exp\left(x_{i,2}+x_{i,3}\right)}{1+\exp\left(x_{i,2}+x_{i,3}\right)} + \frac{\exp\left(x_{i,5}+x_{i,6}\right)}{1+\exp\left(x_{i,5}+x_{i,6}\right)}\right]
\end{aligned} \tag{13}
$$

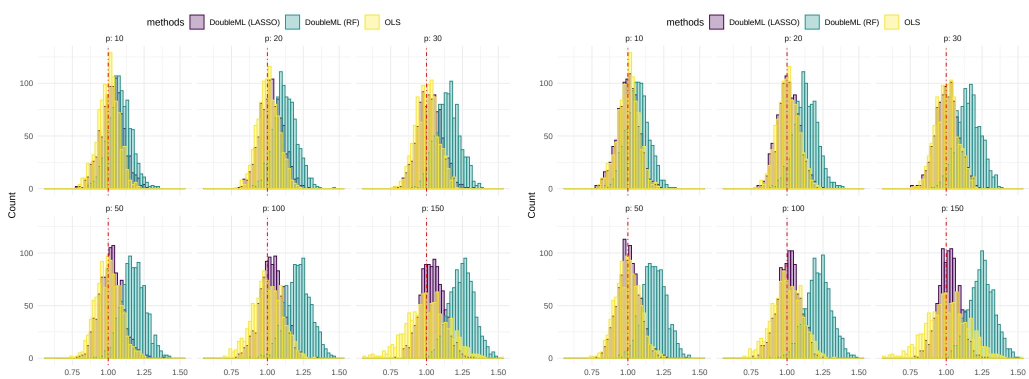
where

$$
\mathbf{X}_i \sim N\left(\mathbf{0}, \mathrm{I}_p\right) \text{ and } U_i, V_i \sim N\left(0,1\right), \quad i = 1, 2, \ldots, n.
$$



(a) IV-type **(b)** Partialling-Out

Figure 7: Logistic-like Models for (12)



(a) IV-type **(b)** Partialling-Out

Figure 8: Logistic-like Models for (13)

5 Indicator Models

Suppose

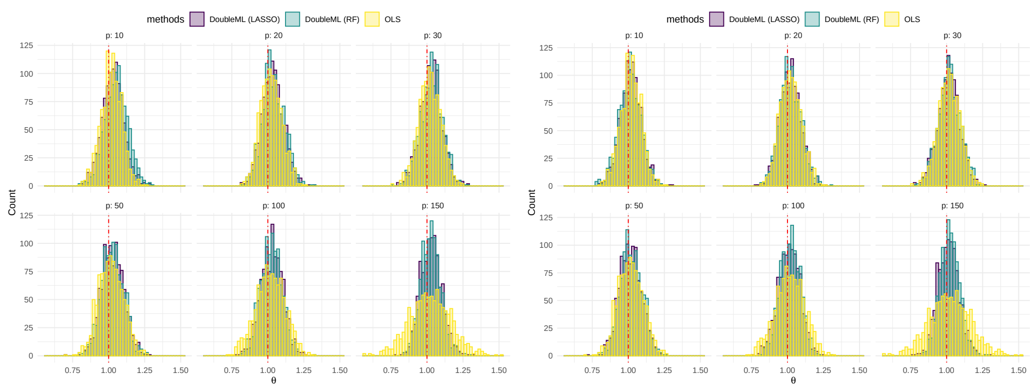
$$\begin{aligned} m(\mathbf{X}_i) =& \mathrm{I}_{x_{i,1}>3} + \mathrm{I}_{x_{i,2}>2} + \mathrm{I}_{x_{i,3}>1} \\ g(\mathbf{X}_i) =& \mathrm{I}_{x_{i,1}>1} + \mathrm{I}_{x_{i,2}>2} + \mathrm{I}_{x_{i,3}>3} \end{aligned} \tag{14}$$

or

$$\begin{aligned} m(\mathbf{X}_i) =& \mathrm{I}_{x_{i,1}>3} + \mathrm{I}_{x_{i,2}>2} + \mathrm{I}_{x_{i,3}>1} + \mathrm{I}_{x_{i,4}>-1} + \mathrm{I}_{x_{i,5}>-2} + \mathrm{I}_{x_{i,6}>-3} + \mathrm{I}_{x_{i,1}x_{i,6}>0} + \mathrm{I}_{x_{i,3}x_{i,4}>1} \\ g(\mathbf{X}_i) =& \mathrm{I}_{x_{i,1}>1} + \mathrm{I}_{x_{i,2}>2} + \mathrm{I}_{x_{i,3}>3} + \mathrm{I}_{x_{i,4}>-3} + \mathrm{I}_{x_{i,5}>-2} + \mathrm{I}_{x_{i,6}>-1} + \mathrm{I}_{x_{i,1}x_{i,6}>2} + \mathrm{I}_{x_{i,3}x_{i,4}>-1} \end{aligned} \tag{15}$$

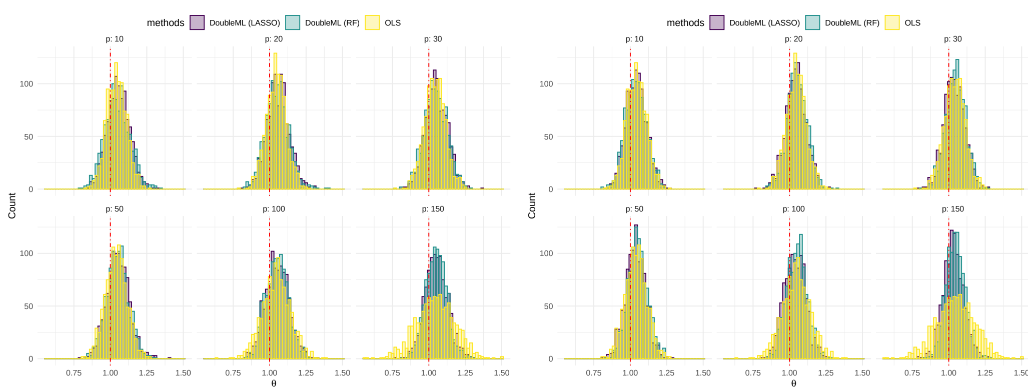
where

$$\mathbf{X}_i \sim N(\mathbf{0}, \mathrm{I}_p) \text{ and } U_i, V_i \sim N(0,1), \quad i = 1, 2, \ldots, n.$$



(a) IV-type **(b)** Partialling-Out

Figure 9: Indicator Models for (14)



(a) IV-type **(b)** Partialling-Out

Figure 10: Indicator Models for (15)