

Notes: Double/Debiased Machine Learning

Ziyang Gong

Consider the partial linear regression model

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

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

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

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

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

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

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

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

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

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

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

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

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

In this note, we use samples $n = 200$ and repeat 1000 times to evaluate the estimation effects of different models, and the code is available at <https://github.com/SignorinoY/ultralisk>.

1 Linear Models

Suppose

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

or

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

where

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

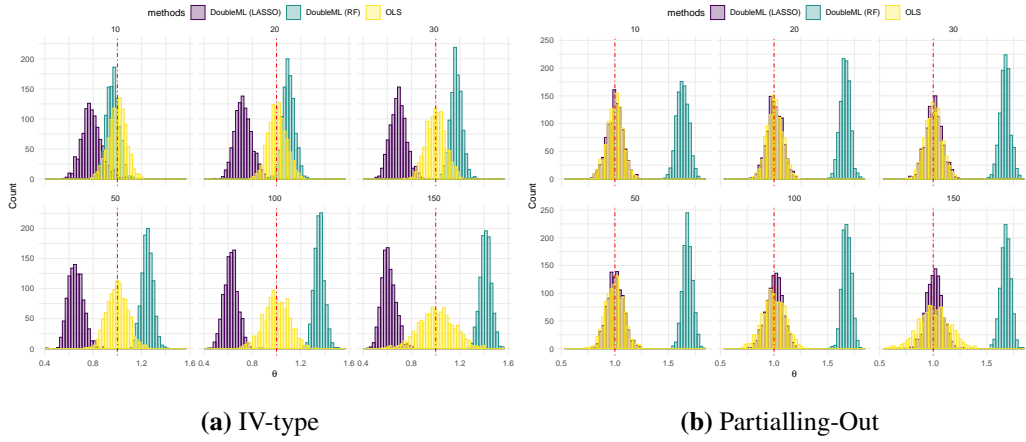


Figure 1: Linear Models for (6)

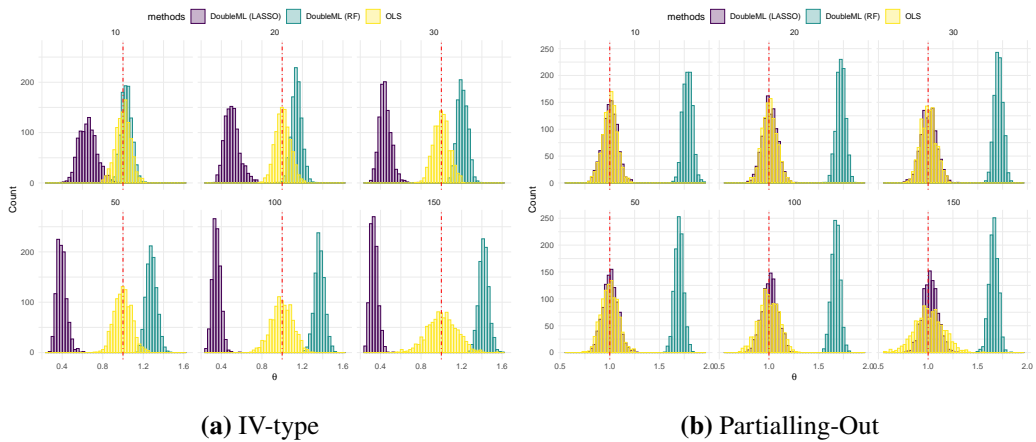


Figure 2: Linear Models for (7)

2 Relu Models

Suppose

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

or

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

where

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

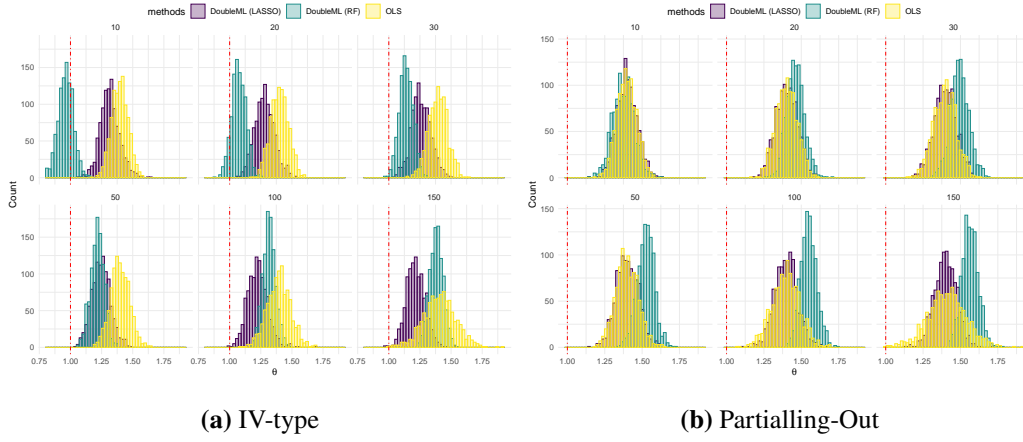


Figure 3: Relu Models for (8)

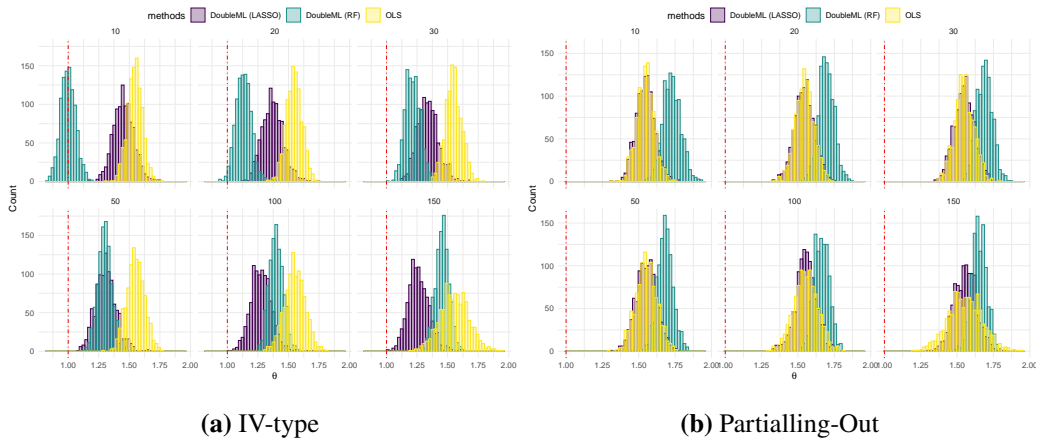


Figure 4: Relu Models for (9)

3 Polynomial Models

Suppose

$$m(\mathbf{X}_i) = x_{i,1}^3 + x_{i,2}^2 + x_{i,1}x_{i,2} + x_{i,2}x_{i,3} \quad (10)$$

$$g(\mathbf{X}_i) = x_{i,2}^3 + x_{i,3}^2 + x_{i,1}x_{i,3} + x_{i,2}x_{i,3}$$

or

$$m(\mathbf{X}_i) = x_{i,1}^3 + x_{i,3}^2 + x_{i,6} + x_{i,1}x_{i,2} + x_{i,4}x_{i,5} + x_{i,2}x_{i,3}x_{i,4} \quad (11)$$

$$g(\mathbf{X}_i) = x_{i,1} + x_{i,2}^2 + x_{i,3}^3 + x_{i,2}x_{i,3} + x_{i,5}x_{i,6} + x_{i,2}x_{i,3}x_{i,4}$$

where

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

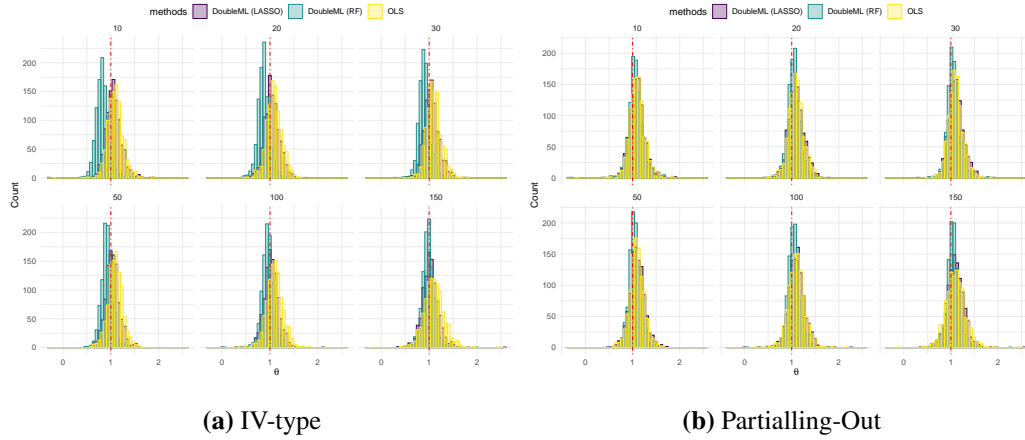


Figure 5: Polynomial Models for (10)

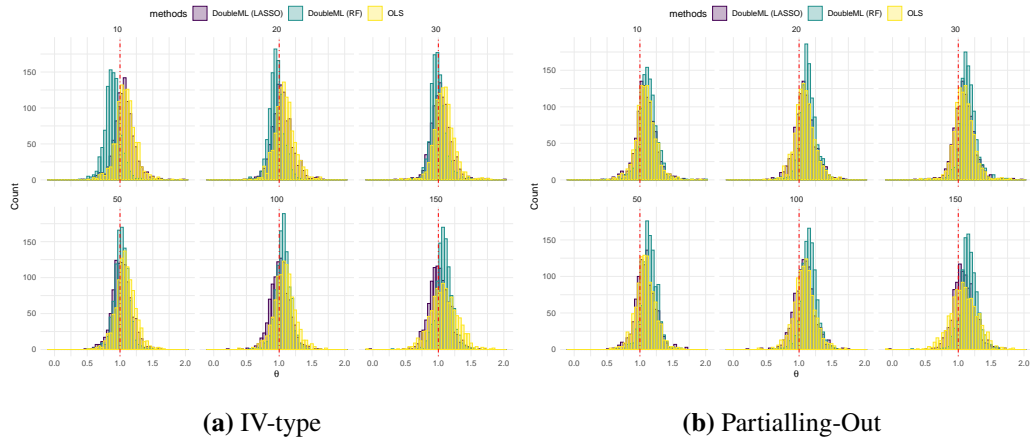


Figure 6: Polynomial Models for (11)

4 Logistic-like Models

Suppose

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

or

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

where

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

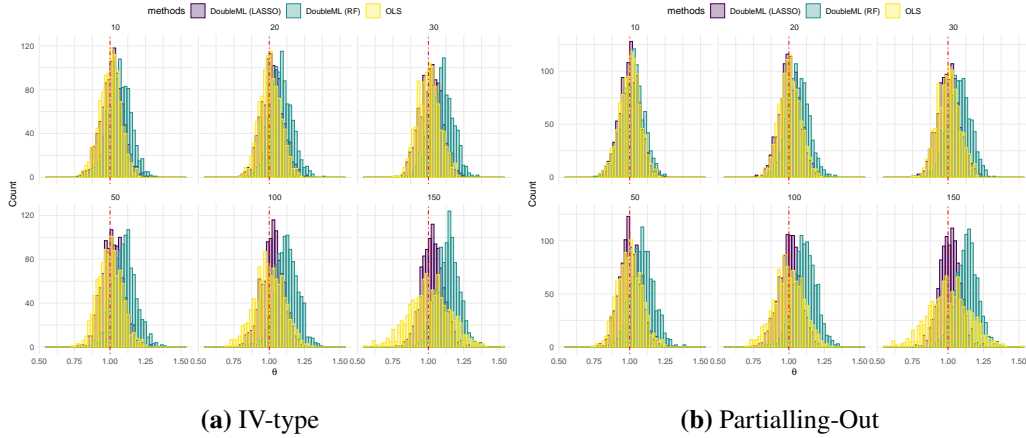


Figure 7: Logistic-like Models for (12)

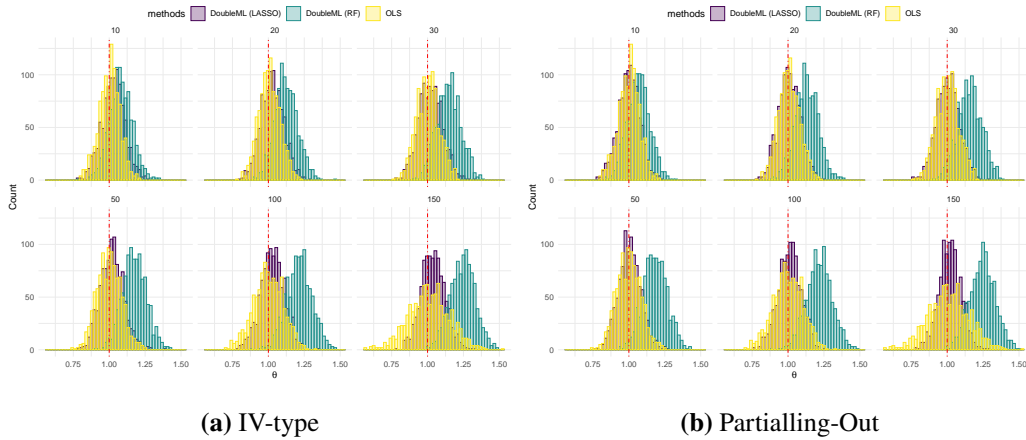


Figure 8: Logistic-like Models for (13)

5 Indicator Models

Suppose

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

or

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

where

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

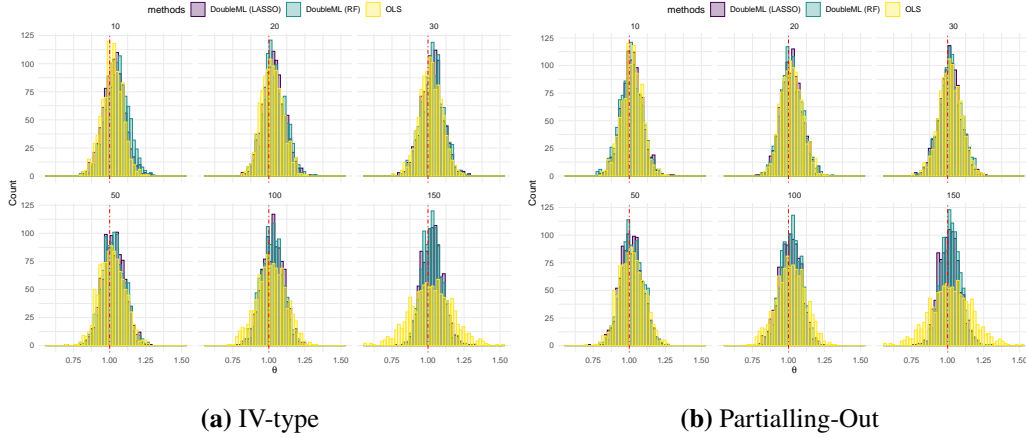


Figure 9: Indicator Models for (14)

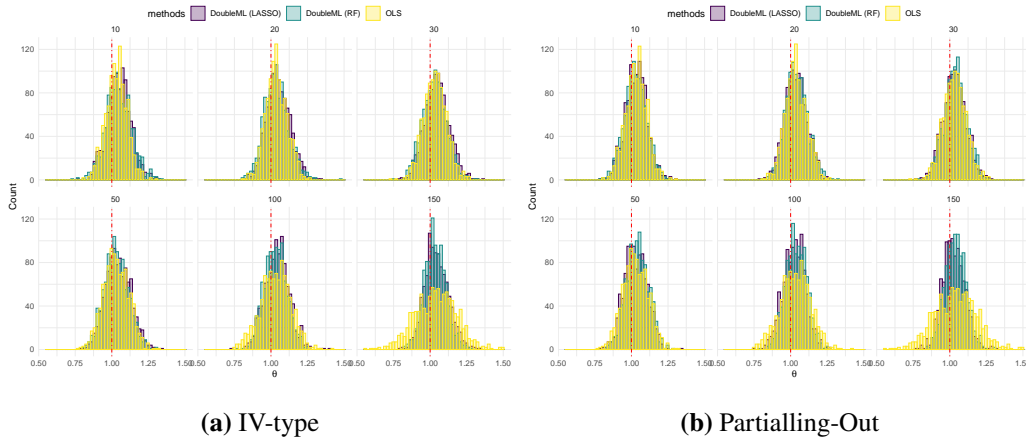


Figure 10: Indicator Models for (15)