Simulation Result for Partially Linear Model

Dongxiao Wu

1 Simulation Design

Let

$$y_i = d_i + x_i'(c_y\theta_0) + u_i,$$
$$d_i = \frac{exp\{x_i'(c_d\theta_0)\}}{1 + exp\{x_i'(c_d\theta_0)\}} + v_i,$$

where $x_i, v_i \sim N(0,1)$, u_i and v_i are independent, $p = dim(x_i) = 250$, the covariates $x_i \sim N(0,\Sigma)$ with $\Sigma_{kj} = (0.5)^{|j-k|}$, and sample size $n = 200$. θ_0 is a $p\times 1$ vector with elements set as $\theta_{0,j} = (1/j)^2$ for $j = 1, ..., p$. c_d and c_y are scalars that control the strength of the relationship between the controls, the outcome, and the treatment variables d_i. We can try several different combinations of c_d and c_y, setting

$$c_d = \sqrt{\frac{(\pi^2/3)R_d^2}{(1-R_d^2)\theta_0'\Sigma\theta_0}}, c_y = \sqrt{\frac{R_y^2}{(1-R_y^2)\theta_0'\Sigma\theta_0}},$$

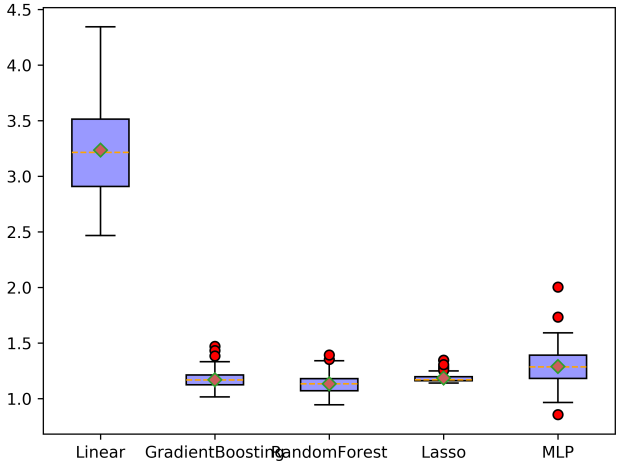
for different combinations of $R_d^2, R_y^2 \in \{0, 0.1, 0.5, 0.9\}$.

2 Results

We set $R_d^2 = R_y^2 = 0.1$.

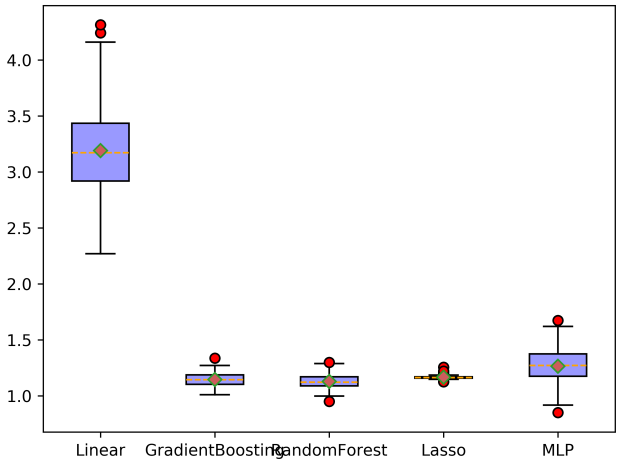
2 folds:

Method	DML1/DML2	$\check{\theta}$	$95\%CI$	$\hat{\theta}$	true θ
LinearRegression	DML1	3.19176	[3.16102,3.22249]	104.27364	1
LinearRegression	DML2	3.17220	[3.12759,3.21671]	102.29749	1
GradientBoostingRegressor	DML1	1.17759	[1.15991,1.19528]	80.22819	1
GradientBoostingRegressor	DML2	1.10664	[1.09871,1.11458]	78.15933	1
RandomForestRegressor	DML1	1.17748	[1.12351,1.15795]	80.28174	1
RandomForestRegressor	DML2	1.19961	[1.18952,1.20970]	80.13725	1
Lasso	DML1	1.17748	[1.15791,1.19705]	88.37562	1
Lasso	DML2	1.154701	[1.14808,1.16134]	87.97575	1
MLPRegressor	DML1	1.29603	[1.28245,1.30962]	88.55359	1
MLPRegressor	DML2	1.25332	[1.24761,1.25904]	86.45241	1



K=2.png

Figure 1: K=2,DML1

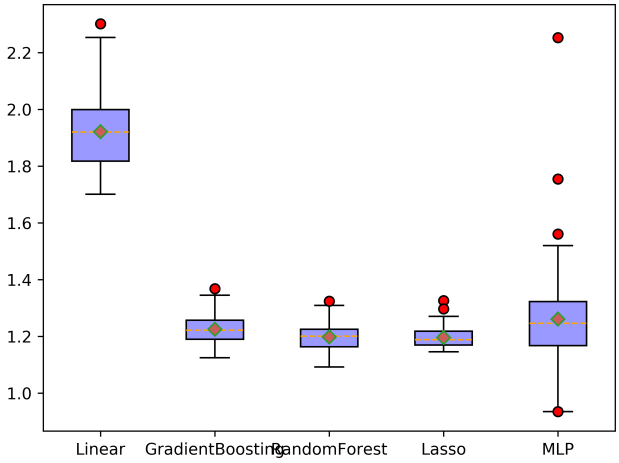


K=2.png

Figure 2: K=2,DML2

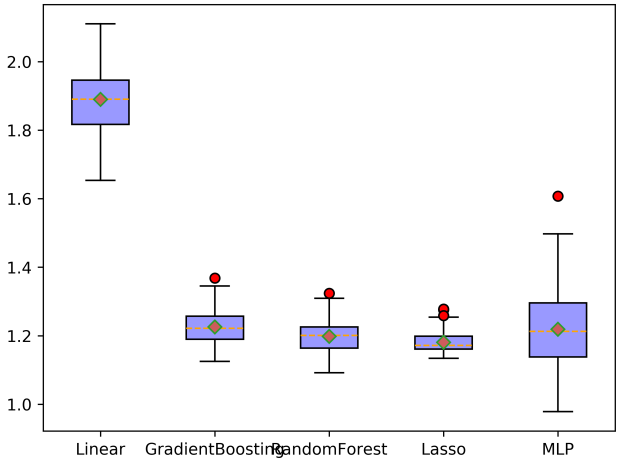
5 folds:

Method	DML1/DML2	$\check{\theta}$	$95\%CI$	$\hat{\theta}$	true θ
LinearRegression	DML1	1.91196	[1.86687,1.95705]	150.23473	1
LinearRegression	DML2	1.89027	[1.87604,1.90449]	148.84999	1
GradientBoostingRegressor	DML1	1.31641	[1.28118,1.35163]	145.15192	1
GradientBoostingRegressor	DML2	1.20509	[1.19359,1.21659]	145.85326	1
RandomForestRegressor	DML1	1.26268	[1.22493,1.30043]	145.41695	1
RandomForestRegressor	DML2	1.21355	[1.20242,1.22467]	145.00807	1
Lasso	DML1	1.29486	[1.25086,1.33887]	144.14305	1
Lasso	DML2	1.17356	[1.15417,1.19295]	140.20278	1
MLPRegressor	DML1	1.32824	[1.26351,1.39297]	133.92810	1
MLPRegressor	DML2	1.22952	[1.17064,1.28841]	132.74614	1



K=5.png

Figure 3: K=5,DML1



K=5.png

Figure 4: K=5,DML2