

TABLE S2 Results of ATE Estimates from the Single Covariate Simulation Study Setting 2:
 $\{\gamma_0, \gamma_1, \gamma_2, \gamma_3\} = \{1, 0.15, 0, 0\}$.

Method	Sample Size	RMSE	MAE	Bias	Rc	SE_{avg}	SE_{emp}
GPMatch	100	0.237	0.152	0.013	0.97	0.243	0.238
	200	0.175	0.114	0.007	0.94	0.169	0.175
	400	0.117	0.084	0.001	0.96	0.117	0.118
QNT_PS	100	0.436	0.271	0.089	0.95	0.466	0.429
	200	0.301	0.210	0.103	0.98	0.287	0.284
	400	0.254	0.171	0.096	0.88	0.209	0.236
LM	100	0.427	0.255	-0.214	0.93	0.399	0.371
	200	0.348	0.174	-0.164	0.93	0.26	0.309
	400	0.318	0.166	-0.198	0.81	0.191	0.25
AIPTW1	100	0.933	0.378	-0.226	0.96	0.671	0.91
	200	3.853	0.246	-0.478	0.95	0.861	3.842
	400	1.25	0.213	-0.396	0.98	0.565	1.192
AIPTW2	100	0.413	0.306	-0.029	0.96	0.411	0.414
	200	0.345	0.156	-0.001	0.96	0.281	0.346
	400	0.244	0.124	-0.021	0.97	0.221	0.244
LM_PS1	100	0.352	0.265	-0.087	0.97	0.368	0.343
	200	0.273	0.181	-0.055	0.93	0.251	0.269
	400	0.192	0.11	-0.082	0.92	0.188	0.174
LM_PS2	100	0.337	0.251	-0.018	0.98	0.397	0.339
	200	0.252	0.154	-0.004	0.97	0.262	0.253
	400	0.175	0.101	-0.004	0.98	0.192	0.176
LM_sp(PS1)	100	0.237	0.158	-0.006	0.97	0.242	0.238
	200	0.171	0.109	-0.004	0.94	0.169	0.172
	400	0.118	0.083	-0.003	0.96	0.118	0.118
LM_sp(PS2)	100	0.248	0.163	0.002	0.96	0.243	0.249
	200	0.171	0.103	-0.001	0.95	0.169	0.172
	400	0.116	0.087	-0.006	0.96	0.118	0.117
Bart	100	0.286	0.176	0.054	0.95	0.266	0.283
	200	0.182	0.115	0.034	0.96	0.18	0.18
	400	0.161	0.085	0.01	0.93	0.127	0.161

RMSE = root mean square error; MAE = median absolute error; Bias = Estimate-True; Rc = Rate of coverage by the 95% interval estimate; SE_{avg} = average of standard error estimate from all replicate; SE_{emp} = standard error of ATE estimates from all replicate; GPMatch: Bayesian structural model with Gaussian process prior, only treatment effect is included in the mean function; covariance function includes X .

QNT_PS: Propensity score sub-classification by quintiles.

AIPTW1 & AIPTW2: augmented inversed probability of treatment weighting.

LM_PS1 & LM_PS2: linear regression modeling with propensity score adjustment;

LM_sp(PS1) & LM_sp(PS2): linear regression modeling with spline fit propensity score adjustment;

BART: Bayesian additive regression tree.

Propensity scores are estimated using different logistic models, with AIPTW1, LM_PS1 & LM_sp(PS1) use PS estimated using logistic model $\text{logit}A \sim X$; and AIPTW2, LM_PS2 & LM_sp(PS2) use PS estimated using logistic model $\text{logit}A \sim X^{1/3}$. QNT_PS using either PS estimates produces identical results.