

TABLE S4 Results of ATE Estimates from the Single Covariate Simulation Study Setting 4:
 $\{\gamma_0, \gamma_1, \gamma_2, \gamma_3\} = \{1, 0.15, 0.7, 0\}$.

Method	Sample Size	RMSE	MAE	Bias	Rc	SE_{avg}	SE_{emp}
GPMatch	100	0.226	0.140	0.022	0.97	0.238	0.226
	200	0.164	0.105	0.024	0.98	0.169	0.163
	400	0.104	0.073	0.009	0.96	0.114	0.104
QNT_PS	100	0.579	0.477	0.437	0.90	0.452	0.382
	200	0.369	0.312	0.279	0.87	0.308	0.244
	400	0.300	0.268	0.245	0.78	0.199	0.174
LM	100	0.323	0.236	0.175	0.96	0.362	0.273
	200	0.292	0.214	0.187	0.90	0.259	0.226
	400	0.213	0.170	0.165	0.87	0.174	0.136
AIPTW1	100	0.596	0.456	0.461	0.78	0.377	0.380
	200	0.523	0.457	0.473	0.48	0.262	0.223
	400	0.474	0.459	0.433	0.32	0.210	0.192
AIPTW2	100	0.541	0.452	0.438	0.76	0.335	0.319
	200	0.488	0.451	0.446	0.48	0.231	0.199
	400	0.418	0.407	0.393	0.31	0.166	0.143
LM_PS1	100	0.398	0.264	0.256	0.84	0.266	0.307
	200	0.376	0.297	0.288	0.64	0.191	0.244
	400	0.340	0.305	0.303	0.35	0.132	0.155
LM_PS2	100	0.563	0.421	0.444	0.81	0.352	0.349
	200	0.498	0.471	0.453	0.50	0.253	0.207
	400	0.411	0.400	0.388	0.32	0.175	0.135
LM_sp(PS1)	100	0.226	0.140	0.004	0.970	0.237	0.227
	200	0.138	0.097	0.007	0.98	0.165	0.138
	400	0.104	0.067	0.006	0.97	0.116	0.104
LM_sp(PS2)	100	0.235	0.147	0.009	0.95	0.240	0.236
	200	0.144	0.112	0.012	0.96	0.167	0.144
	400	0.401	0.361	0.361	0.39	0.165	0.174
Bart	100	0.262	0.183	0.105	0.97	0.271	0.241
	200	0.168	0.117	0.076	0.97	0.190	0.150
	400	0.130	0.097	0.064	0.96	0.136	0.114

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 inverted 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 $logitA \sim X$; and AIPTW2, LM_PS2 & LM_sp(PS2) use PS estimated using logistic model $logitA \sim X^{1/3}$. QNT_PS using either PS estimates produces identical results.