Supplemental Appendix

Figure S1: Results for ignorable scenarios 1-4 (minimized distance)

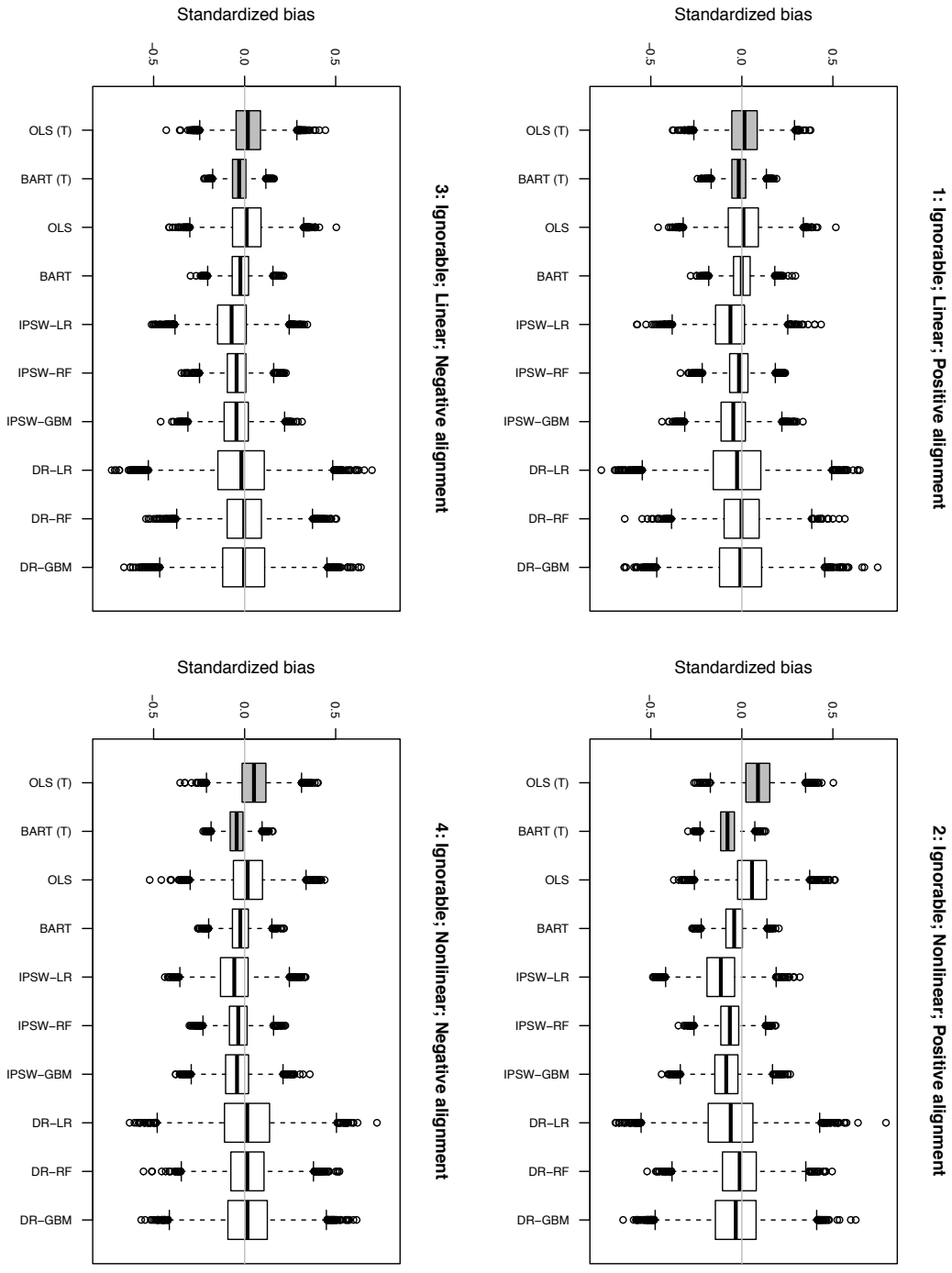

Figure S2: Results for non-ignorable scenarios 5-8 (minimized distance)

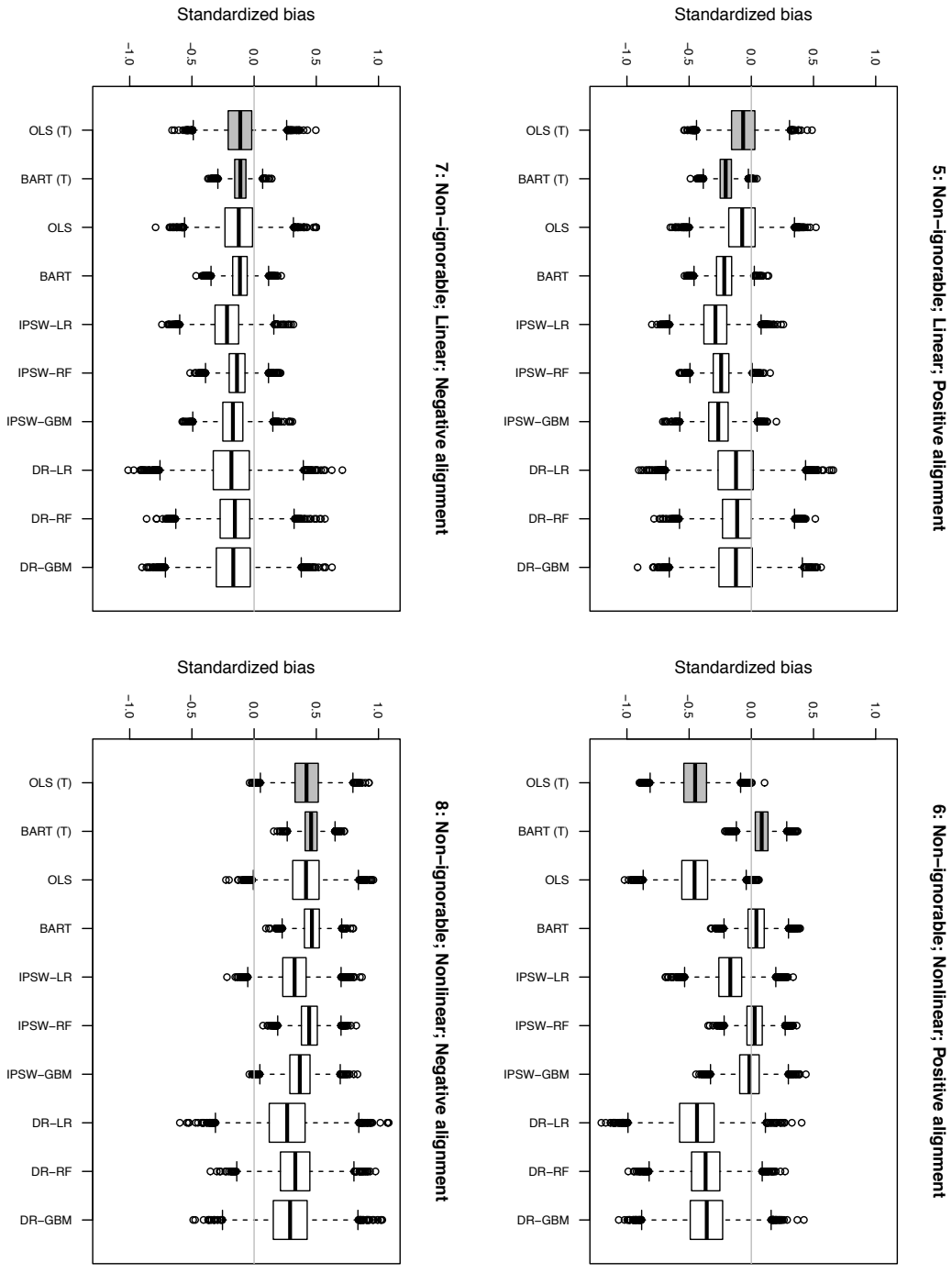

Table S1: Alignment Simulations with reduced sample size: Standardized bias and RSMSE

	Standardized bias		RSMSE	
Alignment	Positive	Negative	Positive	Negative
OLS (T)	−0.036	−0.021	0.064	0.069
BART (T)	−0.047	−0.028	0.061	0.040
OLS	−0.036	−0.021	0.066	0.070
BART	−0.065	−0.042	0.082	0.056
IPSW-LR	−0.031	−0.024	0.073	0.084
IPSW-RF	−0.218	−0.143	0.224	0.156
IPSW-GBM	−0.043	−0.037	0.072	0.075
DR-LR	−0.009	−0.009	0.056	0.067
DR-RF	−0.024	−0.018	0.060	0.064
DR-GBM	−0.015	−0.020	0.054	0.059

Note: The table shows standardized bias and RSMSE averaged over 10,000 simulated datasets for the alignment simulations discussed in the text. Standardized Bias refers to bias divided by the standard deviation of the outcome in the target dataset. RSMSE refers to Root Standardized Mean Square Error. (T) refers to simulations in which control outcomes are available in the target dataset. OLS denotes linear regression; BART denotes Bayesian Additive Regression Trees; IPSW-LR (IPSW-RF/IPSW-GBM) denotes inverse propensity score weighting with propensity scores estimated using logistic regression (random forests/boosting); DR-LR (DR-RF/DR-GBM) refers to double robust weighted linear regression models with propensity scores estimated using logistic regression (random forests/boosting).

Table S2: Standardized bias (minimized distance)

Scenario	1	2	3	4	5	6	7	8	avg bias ignorable	avg bias non-ignorable
OLS (T)	0.01	0.09	0.02	0.05	−0.07	−0.45	−0.11	0.42	0.04	−0.05
BART (T)	−0.02	−0.08	−0.03	−0.04	−0.21	0.08	−0.11	0.46	−0.04	0.06
OLS	0.01	0.06	0.01	0.02	−0.08	−0.46	−0.12	0.42	0.02	−0.06
BART	0.00	−0.04	−0.02	−0.02	−0.22	0.04	−0.11	0.46	−0.02	0.04
IPSW-LR	−0.06	−0.12	−0.07	−0.06	−0.29	−0.17	−0.22	0.33	−0.08	−0.09
IPSW-RF	−0.02	−0.07	−0.04	−0.04	−0.24	0.03	−0.14	0.44	−0.04	0.02
IPSW-GBM	−0.05	−0.09	−0.05	−0.04	−0.27	−0.02	−0.17	0.37	−0.06	−0.02
DR-LR	−0.03	−0.06	−0.02	0.01	−0.13	−0.44	−0.18	0.27	−0.02	−0.12
DR-RF	0.00	−0.01	0.00	0.02	−0.11	−0.37	−0.15	0.33	0.00	−0.08
DR-GBM	−0.01	−0.03	0.00	0.02	−0.13	−0.36	−0.17	0.29	−0.01	−0.09

Note: All results reported here average over 10,000 simulated datasets. Covariate distances between experimental and target datasets are minimized (see text). See Figures **??** and **??** for a description of the scenarios. (T) refers to simulations in which control outcomes are available in the target dataset. OLS denotes linear regression; BART denotes Bayesian Additive Regression Trees; IPSW-LR (IPSW-RF/IPSW-GBM) denotes inverse propensity score weighting with propensity scores estimated using logistic regression (random forests/boosting); DR-LR (DR-RF/DR-GBM) refers to double robust weighted linear regression models with propensity scores estimated using logistic regression (random forests/boosting). The last two columns show the average standardized bias for the ignorable (1–4) and non-ignorable (5–8) scenarios.

Table S3: RSMSE (minimized distance)

Scenario	1	2	3	4	5	6	7	8	avg RSMSE ignorable	avg RSMSE non-ignorable
OLS (T)	0.10	0.13	0.10	0.11	0.15	0.47	0.18	0.44	0.11	0.31
BART (T)	0.06	0.10	0.06	0.07	0.22	0.11	0.13	0.46	0.07	0.23
OLS	0.12	0.13	0.12	0.12	0.18	0.48	0.20	0.45	0.12	0.33
BART	0.07	0.08	0.07	0.07	0.24	0.10	0.14	0.47	0.07	0.24
IPSW-LR	0.13	0.16	0.14	0.13	0.32	0.22	0.26	0.35	0.14	0.29
IPSW-RF	0.08	0.10	0.09	0.08	0.26	0.10	0.17	0.45	0.09	0.24
IPSW-GBM	0.11	0.13	0.11	0.10	0.29	0.12	0.21	0.39	0.11	0.25
DR-LR	0.20	0.19	0.19	0.18	0.25	0.48	0.28	0.34	0.19	0.34
DR-RF	0.14	0.14	0.14	0.14	0.21	0.40	0.24	0.37	0.14	0.31
DR-GBM	0.17	0.17	0.17	0.16	0.24	0.41	0.26	0.35	0.17	0.31

Note: All results reported here average over 10,000 simulated datasets. Covariate distances between experimental and target datasets are minimized (see text). See Figures **??** and **??** for a description of the scenarios. (T) refers to simulations in which control outcomes are available in the target dataset. OLS denotes linear regression; BART denotes Bayesian Additive Regression Trees; IPSW-LR (IPSW-RF/IPSW-GBM) denotes inverse propensity score weighting with propensity scores estimated using logistic regression (random forests/boosting); DR-LR (DR-RF/DR-GBM) refers to double robust weighted linear regression models with propensity scores estimated using logistic regression (random forests/boosting). The last two columns show the average RSMSE for the ignorable (1–4) and non-ignorable (5–8) scenarios.