Subset calibration report: marginal risk difference

2024-09-27

Contents

Results **1**
(Base case) MAR: 12% outcome proportion, 40% missingness proportion . . . 1
MAR: 12% outcome proportion, 80% missingness proportion . . . 6
MAR: 5% outcome proportion, 40% missingness proportion . . . 11
MAR: 5% outcome proportion, 80% missingness proportion . . . 16
MNAR: 12% outcome proportion, 40% missingness proportion . . . 21
MNAR: 12% outcome proportion, 80% missingness proportion . . . 26
MNAR: 5% outcome proportion, 40% missingness proportion . . . 31
MNAR: 5% outcome proportion, 80% missingness proportion . . . 37
Other scenarios . . . 42

The tables in this section contain performance for estimating the marginal risk difference (mRD).

Results

(Base case) MAR: 12% outcome proportion, 40% missingness proportion

Table 1: **Synthetic data MAR simulation: oracle marginal risk difference (mRD), 12% outcome proportion, 40% missing proportion**. Comparing estimators under the **simple outcome and simple MAR** scenario. The value of the estimand is 0.04. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Nominal coverage	Oracle coverage	Power	Prop. completed
Benchmark model	0	-0.001	0.074	0.071	0.075	0.071	0.075	0.939	0.953	1.000	100
Complete-case	-0.261	-0.261	0.081	0.08	0.08	0.273	0.273	0.104	0.100	0.428	100
Confounded model	0.238	0.24	0.075	0.072	0.077	0.249	0.252	0.096	0.110	1.000	100
IPW	-0.001	0	0.132	0.135	0.133	0.135	0.133	0.950	0.951	0.868	100
Raking (vanilla)	0	-0.002	0.084	0.08	0.084	0.08	0.084	0.938	0.947	0.998	100
MICE	0	-0.001	0.08	0.077	0.08	0.077	0.08	0.938	0.949	1.000	100
MI-XGB	-0.007	-0.008	0.083	0.078	0.082	0.078	0.082	0.935	0.948	0.999	100
MI-RF	0.007	0.005	0.084	0.076	0.083	0.076	0.083	0.924	0.946	0.999	100
IPCW-TMLE-M	-0.024	-0.027	0.156	0.156	0.155	0.158	0.157	0.932	0.949	0.711	100
IPCW-TMLE-MTO	-0.037	-0.036	0.146	0.141	0.142	0.146	0.147	0.920	0.946	0.759	100

Table 2: **Synthetic data MAR simulation: census marginal risk difference (mRD), 12% outcome proportion, 40% missing proportion**. Comparing estimators under the **simple outcome and simple MAR** scenario. The value of the estimand is 0.041. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Oracle coverage	Nominal coverage	Power	Prop. completed
Benchmark model	-0.003	-0.004	0.074	0.071	0.075	0.071	0.075	0.951	0.938	1.000	100
Complete-case	-0.264	-0.265	0.081	0.08	0.08	0.276	0.276	0.091	0.097	0.428	100
Confounded model	0.235	0.236	0.075	0.072	0.077	0.246	0.249	0.119	0.101	1.000	100
IPW	-0.004	-0.003	0.132	0.135	0.133	0.135	0.133	0.950	0.952	0.868	100
Raking (vanilla)	-0.004	-0.005	0.084	0.08	0.084	0.08	0.084	0.947	0.938	0.998	100
MICE	-0.003	-0.004	0.08	0.077	0.08	0.077	0.08	0.948	0.940	1.000	100
MI-XGB	-0.01	-0.011	0.083	0.078	0.082	0.079	0.083	0.948	0.935	0.999	100
MI-RF	0.004	0.002	0.084	0.076	0.083	0.076	0.083	0.947	0.925	0.999	100
IPCW-TMLE-M	-0.027	-0.031	0.156	0.156	0.155	0.158	0.158	0.948	0.931	0.711	100
IPCW-TMLE-MTO	-0.04	-0.039	0.146	0.141	0.142	0.147	0.148	0.943	0.918	0.759	100

Table 3: **Synthetic data MAR simulation: oracle marginal risk difference (mRD), 12% outcome proportion, 40% missing proportion**. Comparing estimators under the **complex outcome and simple MAR** scenario. The value of the estimand is 0.031. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE. Estimators that are mismatched with the estimand (i.e., are estimating a different parameter) are emphasized using a star.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Nominal coverage	Oracle coverage	Power	Prop. completed
Benchmark model	0	0	0.064	0.064	0.064	0.064	0.064	0.954	0.955	0.999	100
Complete-case*	-0.197	-0.199	0.085	0.084	0.084	0.214	0.216	0.360	0.355	0.279	100
Confounded model*	-0.223	-0.222	0.07	0.069	0.071	0.233	0.233	0.107	0.104	0.260	100
IPW*	0.047	0.044	0.134	0.136	0.135	0.144	0.142	0.948	0.939	0.779	100
Raking (vanilla)*	0.054	0.055	0.081	0.082	0.081	0.098	0.098	0.904	0.896	0.996	100
MICE*	0.122	0.121	0.082	0.082	0.083	0.147	0.147	0.698	0.688	1.000	100
MI-XGB*	0.093	0.093	0.08	0.08	0.08	0.123	0.123	0.797	0.792	0.999	100
MI-RF*	0.061	0.061	0.082	0.079	0.082	0.1	0.103	0.877	0.886	0.998	100
IPCW-TMLE-M	-0.055	-0.075	0.187	0.173	0.182	0.182	0.197	0.892	0.951	0.284	100
IPCW-TMLE-MTO	-0.064	-0.069	0.145	0.136	0.143	0.15	0.158	0.888	0.938	0.453	100
IPCW-a-TMLE-M	-0.058	-0.076	0.187	0.172	0.184	0.182	0.199	0.888	0.951	0.280	100
IPCW-a-TMLE-MTO	-0.08	-0.084	0.137	0.125	0.136	0.148	0.16	0.847	0.914	0.468	100

Table 4: **Synthetic data MAR simulation: census marginal risk difference (mRD), 12% outcome proportion, 40% missing proportion**. Comparing estimators under the **complex outcome and simple MAR** scenario. The value of the estimand is 0.037. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE. Estimators that are mismatched with the estimand (i.e., are estimating a different parameter) are emphasized using a star.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Oracle coverage	Nominal coverage	Power	Prop. completed
Benchmark model	0.002	0.002	0.073	0.072	0.073	0.072	0.073	0.949	0.947	1.000	100
Complete-case	-0.249	-0.251	0.085	0.084	0.084	0.263	0.265	0.164	0.176	0.279	100
Confounded model	-0.275	-0.275	0.07	0.069	0.071	0.284	0.284	0.023	0.023	0.260	100
IPW	-0.005	-0.009	0.134	0.136	0.135	0.136	0.136	0.953	0.953	0.779	100
Raking (vanilla)	0.001	0.002	0.081	0.082	0.081	0.082	0.081	0.950	0.950	0.996	100
MICE	0.069	0.069	0.082	0.082	0.083	0.108	0.107	0.867	0.873	1.000	100
MI-XGB	0.041	0.041	0.08	0.08	0.08	0.09	0.09	0.918	0.923	0.999	100
MI-RF	0.009	0.009	0.082	0.079	0.082	0.079	0.083	0.950	0.935	0.998	100
IPCW-TMLE-M*	-0.108	-0.128	0.187	0.173	0.182	0.204	0.222	0.926	0.823	0.284	100
IPCW-TMLE-MTO*	-0.116	-0.121	0.145	0.136	0.143	0.179	0.187	0.880	0.795	0.453	100
IPCW-a-TMLE-M*	-0.111	-0.129	0.187	0.172	0.184	0.205	0.225	0.925	0.817	0.280	100
IPCW-a-TMLE-MTO*	-0.132	-0.136	0.137	0.125	0.136	0.182	0.192	0.840	0.739	0.468	100

Table 5: **Synthetic data MAR simulation: oracle marginal risk difference (mRD), 12% outcome proportion, 40% missing proportion**. Comparing estimators under the **simple outcome and complex MAR** scenario. The value of the estimand is 0.04. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Nominal cover-age	Oracle cover-age	Power	Prop. com-pleted
Benchmark model	-0.002	-0.002	0.073	0.071	0.072	0.071	0.073	0.941	0.947	1.000	100
Complete-case	0.198	0.198	0.093	0.092	0.095	0.218	0.22	0.426	0.436	1.000	100
Confounded model	0.236	0.236	0.074	0.072	0.073	0.247	0.247	0.101	0.112	1.000	100
IPW	0.123	0.124	0.094	0.098	0.097	0.157	0.157	0.774	0.746	1.000	100
Raking (vanilla)	-0.004	-0.004	0.077	0.073	0.077	0.073	0.077	0.936	0.950	1.000	100
MICE	-0.002	-0.002	0.076	0.074	0.076	0.074	0.076	0.942	0.948	1.000	100
MI-RF	-0.009	-0.008	0.077	0.074	0.078	0.075	0.078	0.940	0.948	1.000	100
IPCW-TMLE-M	0.038	0.036	0.106	0.118	0.109	0.124	0.114	0.975	0.934	0.990	100
IPCW-TMLE-MTO	0.044	0.043	0.103	0.11	0.105	0.118	0.113	0.957	0.930	0.994	100
IPCW-a-TMLE-M	0.035	0.032	0.108	0.119	0.11	0.124	0.114	0.973	0.936	0.984	100
IPCW-a-TMLE-MTO	0.044	0.041	0.104	0.109	0.106	0.118	0.114	0.954	0.927	0.994	100

Table 6: **Synthetic data MAR simulation: census marginal risk difference (mRD), 12% outcome proportion, 40% missing proportion**. Comparing estimators under the **simple outcome and complex MAR** scenario. The value of the estimand is 0.041. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Oracle cover-age	Nominal cover-age	Power	Prop. com-pleted
Benchmark model	-0.005	-0.006	0.073	0.071	0.072	0.071	0.073	0.946	0.942	1.000	100
Complete-case	0.195	0.195	0.093	0.092	0.095	0.215	0.217	0.450	0.440	1.000	100
Confounded model	0.233	0.233	0.074	0.072	0.073	0.244	0.244	0.118	0.108	1.000	100
IPW	0.12	0.12	0.094	0.098	0.097	0.154	0.155	0.754	0.788	1.000	100
Raking (vanilla)	-0.007	-0.007	0.077	0.073	0.077	0.074	0.077	0.949	0.936	1.000	100
MICE	-0.006	-0.005	0.076	0.074	0.076	0.074	0.076	0.950	0.945	1.000	100
MI-RF	-0.013	-0.012	0.077	0.074	0.078	0.075	0.079	0.946	0.936	1.000	100
IPCW-TMLE-M	0.034	0.033	0.106	0.118	0.109	0.123	0.113	0.939	0.977	0.990	100
IPCW-TMLE-MTO	0.041	0.039	0.103	0.11	0.105	0.117	0.112	0.934	0.958	0.994	100
IPCW-a-TMLE-M	0.032	0.029	0.108	0.119	0.11	0.123	0.113	0.939	0.974	0.984	100
IPCW-a-TMLE-MTO	0.04	0.037	0.104	0.109	0.106	0.117	0.113	0.930	0.956	0.994	100

Table 7: **Synthetic data MAR simulation: oracle marginal risk difference (mRD), 12% outcome proportion, 40% missing proportion**. Comparing estimators under the **complex outcome and complex MAR** scenario. The value of the estimand is 0.031. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE. Estimators that are mismatched with the estimand (i.e., are estimating a different parameter) are emphasized using a star.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Nominal coverage	Oracle coverage	Power	Prop. completed
Benchmark model	-0.002	-0.001	0.063	0.064	0.064	0.064	0.064	0.957	0.956	0.998	100
Complete-case*	0.247	0.246	0.098	0.099	0.101	0.267	0.266	0.296	0.294	1.000	100
Confounded model*	-0.225	-0.224	0.068	0.069	0.068	0.235	0.234	0.098	0.088	0.242	100
IPW*	0.132	0.132	0.097	0.098	0.097	0.165	0.164	0.746	0.724	0.997	100
Raking (vanilla)*	0.057	0.057	0.076	0.074	0.079	0.093	0.097	0.882	0.890	0.999	100
MICE*	0.102	0.102	0.076	0.077	0.079	0.127	0.129	0.735	0.729	1.000	100
MI-RF*	0.055	0.054	0.076	0.076	0.08	0.094	0.097	0.896	0.895	0.999	100
IPCW-TMLE-M	0.023	0.018	0.121	0.127	0.121	0.129	0.123	0.961	0.946	0.807	100
IPCW-TMLE-MTO	0.011	0.006	0.103	0.102	0.103	0.102	0.104	0.950	0.951	0.912	100
IPCW-a-TMLE-M	0.031	0.028	0.124	0.127	0.125	0.131	0.128	0.958	0.940	0.820	100
IPCW-a-TMLE-MTO	0.017	0.013	0.104	0.101	0.102	0.102	0.103	0.945	0.945	0.920	100

Table 8: **Synthetic data MAR simulation: census marginal risk difference (mRD), 12% outcome proportion, 40% missing proportion**. Comparing estimators under the **complex outcome and complex MAR** scenario. The value of the estimand is 0.037. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE. Estimators that are mismatched with the estimand (i.e., are estimating a different parameter) are emphasized using a star.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Oracle coverage	Nominal coverage	Power	Prop. completed
Benchmark model	0.001	0	0.071	0.072	0.071	0.072	0.071	0.954	0.960	1.000	100
Complete-case	0.195	0.193	0.098	0.099	0.101	0.219	0.218	0.498	0.502	1.000	100
Confounded model	-0.277	-0.277	0.068	0.069	0.068	0.286	0.285	0.020	0.021	0.242	100
IPW	0.08	0.079	0.097	0.098	0.097	0.127	0.126	0.870	0.887	0.997	100
Raking (vanilla)	0.004	0.004	0.076	0.074	0.079	0.074	0.079	0.951	0.944	0.999	100
MICE	0.049	0.049	0.076	0.077	0.079	0.091	0.093	0.906	0.908	1.000	100
MI-RF	0.003	0.002	0.076	0.076	0.08	0.076	0.08	0.948	0.948	0.999	100
IPCW-TMLE-M*	-0.03	-0.035	0.121	0.127	0.121	0.13	0.126	0.947	0.938	0.807	100
IPCW-TMLE-MTO*	-0.042	-0.047	0.103	0.102	0.103	0.11	0.114	0.929	0.917	0.912	100
IPCW-a-TMLE-M*	-0.021	-0.025	0.124	0.127	0.125	0.129	0.127	0.950	0.940	0.820	100
IPCW-a-TMLE-MTO*	-0.036	-0.04	0.104	0.101	0.102	0.107	0.11	0.936	0.919	0.920	100

MAR: 12% outcome proportion, 80% missingness proportion

Table 9: **Synthetic data MAR simulation: oracle marginal risk difference (mRD), 12% outcome proportion, 80% missing proportion**. Comparing estimators under the **simple outcome and simple MAR** scenario. The value of the estimand is 0.04. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Nominal coverage	Oracle coverage	Power	Prop. completed
Benchmark model	0.003	0.003	0.07	0.071	0.069	0.071	0.069	0.949	0.947	1.000	100
Complete-case	-0.251	-0.251	0.12	0.12	0.123	0.278	0.279	0.442	0.436	0.224	100
Confounded model	0.241	0.241	0.071	0.072	0.071	0.252	0.251	0.069	0.066	1.000	100
IPW	0.002	-0.003	0.28	0.279	0.278	0.279	0.278	0.944	0.951	0.267	100
Raking (vanilla)	0.001	0.002	0.113	0.111	0.11	0.111	0.11	0.939	0.944	0.942	100
MICE	0.004	0.004	0.098	0.097	0.095	0.097	0.095	0.949	0.950	0.981	100
MI-RF	0.072	0.075	0.101	0.083	0.096	0.11	0.122	0.810	0.895	0.998	100
IPCW-TMLE-M	-0.048	-0.078	0.313	0.296	0.294	0.3	0.304	0.913	0.952	0.167	100
IPCW-TMLE-MTO	-0.066	-0.087	0.283	0.258	0.279	0.266	0.293	0.892	0.945	0.230	100

Table 10: **Synthetic data MAR simulation: census marginal risk difference (mRD), 12% outcome proportion, 80% missing proportion**. Comparing estimators under the **simple outcome and simple MAR** scenario. The value of the estimand is 0.041. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Oracle coverage	Nominal coverage	Power	Prop. completed
Benchmark model	0	0	0.07	0.071	0.069	0.071	0.069	0.947	0.948	1.000	100
Complete-case	-0.254	-0.254	0.12	0.12	0.123	0.281	0.282	0.424	0.432	0.224	100
Confounded model	0.238	0.238	0.071	0.072	0.071	0.249	0.248	0.077	0.079	1.000	100
IPW	-0.002	-0.006	0.28	0.279	0.278	0.279	0.278	0.952	0.943	0.267	100
Raking (vanilla)	-0.002	-0.001	0.113	0.111	0.11	0.111	0.11	0.942	0.938	0.942	100
MICE	0	0.001	0.098	0.097	0.095	0.097	0.095	0.951	0.949	0.981	100
MI-RF	0.068	0.072	0.101	0.083	0.096	0.108	0.12	0.897	0.816	0.998	100
IPCW-TMLE-M	-0.052	-0.081	0.313	0.296	0.294	0.301	0.305	0.951	0.910	0.167	100
IPCW-TMLE-MTO	-0.07	-0.09	0.283	0.258	0.279	0.267	0.294	0.946	0.891	0.230	100

Table 11: **Synthetic data MAR simulation: oracle marginal risk difference (mRD), 12% outcome proportion, 80% missing proportion**. Comparing estimators under the **complex outcome and simple MAR** scenario. The value of the estimand is 0.031. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE. Estimators that are mismatched with the estimand (i.e., are estimating a different parameter) are emphasized using a star.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Nominal coverage	Oracle coverage	Power	Prop. completed
Benchmark model	-0.003	-0.002	0.066	0.064	0.065	0.064	0.065	0.940	0.946	0.998	100
Complete-case*	-0.197	-0.199	0.134	0.132	0.135	0.237	0.241	0.652	0.682	0.119	100
Confounded model*	-0.226	-0.225	0.069	0.069	0.07	0.236	0.236	0.103	0.098	0.250	100
IPW*	0.035	0.025	0.297	0.274	0.286	0.276	0.287	0.933	0.958	0.219	100
Raking (vanilla)*	0.048	0.046	0.12	0.117	0.122	0.126	0.13	0.926	0.932	0.872	100
MICE*	0.142	0.141	0.119	0.113	0.122	0.181	0.186	0.760	0.774	0.984	100
MI-RF*	0.062	0.059	0.112	0.086	0.114	0.106	0.128	0.819	0.912	0.972	100
IPCW-TMLE-M	-0.09	-0.132	0.366	0.315	0.339	0.327	0.364	0.842	0.955	0.074	100
IPCW-TMLE-MTO	-0.114	-0.134	0.277	0.243	0.277	0.269	0.308	0.843	0.939	0.119	100

Table 12: **Synthetic data MAR simulation: census marginal risk difference (mRD), 12% outcome proportion, 80% missing proportion**. Comparing estimators under the **complex outcome and simple MAR** scenario. The value of the estimand is 0.037. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE. Estimators that are mismatched with the estimand (i.e., are estimating a different parameter) are emphasized using a star.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Oracle coverage	Nominal coverage	Power	Prop. completed
Benchmark model	-0.001	-0.002	0.074	0.072	0.075	0.073	0.075	0.949	0.945	0.999	100
Complete-case	-0.25	-0.252	0.134	0.132	0.135	0.283	0.286	0.526	0.515	0.119	100
Confounded model	-0.279	-0.278	0.069	0.069	0.07	0.287	0.286	0.020	0.021	0.250	100
IPW	-0.017	-0.027	0.297	0.274	0.286	0.275	0.287	0.959	0.931	0.219	100
Raking (vanilla)	-0.004	-0.006	0.12	0.117	0.122	0.117	0.122	0.950	0.940	0.872	100
MICE	0.089	0.088	0.119	0.113	0.122	0.144	0.151	0.885	0.881	0.984	100
MI-RF	0.009	0.006	0.112	0.086	0.114	0.087	0.114	0.947	0.868	0.972	100
IPCW-TMLE-M*	-0.143	-0.185	0.366	0.315	0.339	0.345	0.386	0.948	0.799	0.074	100
IPCW-TMLE-MTO*	-0.167	-0.186	0.277	0.243	0.277	0.295	0.334	0.918	0.796	0.119	100

Table 13: **Synthetic data MAR simulation: oracle marginal risk difference (mRD), 12% outcome proportion, 80% missing proportion**. Comparing estimators under the **simple outcome and complex MAR** scenario. The value of the estimand is 0.04. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Nominal coverage	Oracle coverage	Power	Prop. completed
Benchmark model	0	0	0.072	0.071	0.072	0.071	0.072	0.944	0.950	1.000	100
Complete-case	0.719	0.717	0.179	0.18	0.176	0.742	0.738	0.015	0.018	1.000	100
Confounded model	0.237	0.238	0.074	0.072	0.072	0.248	0.249	0.101	0.116	1.000	100
IPW	0.26	0.26	0.169	0.17	0.168	0.311	0.31	0.694	0.675	0.986	100
Raking (vanilla)	-0.012	-0.013	0.097	0.098	0.096	0.098	0.097	0.947	0.946	0.978	100
MICE	0.003	0.003	0.087	0.086	0.089	0.086	0.089	0.951	0.954	0.998	100
MI-RF	-0.003	-0.001	0.09	0.081	0.093	0.081	0.093	0.926	0.954	0.999	100
IPCW-TMLE-M	0.049	0.038	0.184	0.202	0.178	0.208	0.182	0.976	0.941	0.646	100
IPCW-TMLE-MTO	0.091	0.083	0.174	0.181	0.171	0.203	0.19	0.946	0.910	0.824	100

Table 14: **Synthetic data MAR simulation: census marginal risk difference (mRD), 12% outcome proportion, 80% missing proportion**. Comparing estimators under the **simple outcome and complex MAR** scenario. The value of the estimand is 0.041. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Oracle coverage	Nominal coverage	Power	Prop. completed
Benchmark model	-0.004	-0.004	0.072	0.071	0.072	0.071	0.072	0.948	0.943	1.000	100
Complete-case	0.716	0.713	0.179	0.18	0.176	0.738	0.735	0.018	0.015	1.000	100
Confounded model	0.234	0.235	0.074	0.072	0.072	0.245	0.246	0.124	0.111	1.000	100
IPW	0.257	0.257	0.169	0.17	0.168	0.308	0.307	0.684	0.701	0.986	100
Raking (vanilla)	-0.015	-0.016	0.097	0.098	0.096	0.099	0.097	0.945	0.947	0.978	100
MICE	0	0	0.087	0.086	0.089	0.086	0.089	0.954	0.952	0.998	100
MI-RF	-0.006	-0.005	0.09	0.081	0.093	0.081	0.093	0.952	0.925	0.999	100
IPCW-TMLE-M	0.045	0.035	0.184	0.202	0.178	0.207	0.181	0.942	0.975	0.646	100
IPCW-TMLE-MTO	0.088	0.08	0.174	0.181	0.171	0.201	0.188	0.912	0.946	0.824	100

Table 15: **Synthetic data MAR simulation: oracle marginal risk difference (mRD), 12% outcome proportion, 80% missing proportion**. Comparing estimators under the **complex outcome and complex MAR** scenario. The value of the estimand is 0.031. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE. Estimators that are mismatched with the estimand (i.e., are estimating a different parameter) are emphasized using a star.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Nominal coverage	Oracle coverage	Power	Prop. completed
Benchmark model	-0.003	-0.004	0.064	0.064	0.065	0.064	0.065	0.948	0.952	1.000	100
Complete-case*	0.707	0.699	0.196	0.199	0.193	0.734	0.726	0.040	0.043	0.999	100
Confounded model*	-0.224	-0.224	0.069	0.069	0.068	0.234	0.234	0.109	0.106	0.252	100
IPW*	0.148	0.144	0.16	0.161	0.16	0.219	0.215	0.881	0.851	0.863	100
Raking (vanilla)*	0.067	0.065	0.104	0.104	0.102	0.124	0.121	0.914	0.903	0.970	100
MICE*	0.208	0.205	0.099	0.092	0.097	0.228	0.227	0.393	0.450	1.000	100
MI-RF*	0.13	0.128	0.096	0.084	0.097	0.155	0.161	0.647	0.734	0.999	100
IPCW-TMLE-M	0.031	0.017	0.199	0.214	0.189	0.216	0.19	0.972	0.939	0.319	100
IPCW-TMLE-MTO	0.025	0.015	0.171	0.169	0.163	0.171	0.164	0.955	0.945	0.518	100

Table 16: **Synthetic data MAR simulation: census marginal risk difference (mRD), 12% outcome proportion, 80% missing proportion**. Comparing estimators under the **complex outcome and complex MAR** scenario. The value of the estimand is 0.037. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE. Estimators that are mismatched with the estimand (i.e., are estimating a different parameter) are emphasized using a star.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Oracle coverage	Nominal coverage	Power	Prop. completed
Benchmark model	0.001	-0.001	0.073	0.072	0.073	0.072	0.073	0.953	0.949	1.000	100
Complete-case	0.654	0.647	0.196	0.199	0.193	0.684	0.675	0.079	0.072	0.999	100
Confounded model	-0.276	-0.276	0.069	0.069	0.068	0.285	0.285	0.020	0.023	0.252	100
IPW	0.095	0.092	0.16	0.161	0.16	0.187	0.184	0.912	0.932	0.863	100
Raking (vanilla)	0.015	0.013	0.104	0.104	0.102	0.105	0.103	0.947	0.952	0.970	100
MICE	0.156	0.153	0.099	0.092	0.097	0.181	0.181	0.662	0.615	1.000	100
MI-RF	0.078	0.076	0.096	0.084	0.097	0.114	0.123	0.880	0.830	0.999	100
IPCW-TMLE-M*	-0.022	-0.035	0.199	0.214	0.189	0.215	0.192	0.951	0.957	0.319	100
IPCW-TMLE-MTO*	-0.027	-0.038	0.171	0.169	0.163	0.171	0.167	0.953	0.942	0.518	100

MAR: 5% outcome proportion, 40% missingness proportion

Table 17: **Synthetic data MAR simulation: oracle marginal risk difference (mRD), 5% outcome proportion, 40% missing proportion**. Comparing estimators under the **simple outcome and simple MAR** scenario. The value of the estimand is 0.019. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Nominal coverage	Oracle coverage	Power	Prop. completed
Benchmark model	0	-0.001	0.048	0.048	0.05	0.048	0.05	0.944	0.947	0.974	100
Complete-case	-0.124	-0.123	0.053	0.052	0.053	0.134	0.134	0.338	0.345	0.233	100
Confounded model	0.111	0.112	0.05	0.049	0.051	0.122	0.123	0.382	0.393	1.000	100
IPW	0	0	0.096	0.095	0.095	0.095	0.095	0.941	0.948	0.515	100
Raking (vanilla)	0	-0.001	0.054	0.054	0.054	0.054	0.054	0.950	0.952	0.940	100
MICE	0	0.001	0.053	0.052	0.052	0.052	0.052	0.947	0.949	0.952	100
MI-RF	0.019	0.019	0.054	0.052	0.054	0.055	0.057	0.927	0.939	0.976	100
IPCW-TMLE-M	-0.014	-0.018	0.107	0.104	0.103	0.105	0.104	0.929	0.947	0.370	100
IPCW-TMLE-MTO	-0.018	-0.021	0.1	0.096	0.097	0.097	0.099	0.920	0.942	0.423	100
r-IPCW-TMLE-MTO	-0.019	-0.022	0.1	0.096	0.098	0.097	0.1	0.919	0.944	0.418	100

Table 18: **Synthetic data MAR simulation: census marginal risk difference (mRD), 5% outcome proportion, 40% missing proportion**. Comparing estimators under the **simple outcome and simple MAR** scenario. The value of the estimand is 0.019. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Oracle coverage	Nominal coverage	Power	Prop. completed
Benchmark model	0.003	0.003	0.048	0.048	0.05	0.048	0.05	0.948	0.947	0.974	100
Complete-case	-0.121	-0.12	0.053	0.052	0.053	0.131	0.131	0.369	0.363	0.233	100
Confounded model	0.114	0.115	0.05	0.049	0.051	0.125	0.125	0.372	0.357	1.000	100
IPW	0.004	0.003	0.096	0.095	0.095	0.095	0.095	0.950	0.942	0.515	100
Raking (vanilla)	0.003	0.002	0.054	0.054	0.054	0.054	0.054	0.955	0.952	0.940	100
MICE	0.004	0.004	0.053	0.052	0.052	0.052	0.053	0.948	0.948	0.952	100
MI-RF	0.022	0.022	0.054	0.052	0.054	0.056	0.058	0.935	0.920	0.976	100
IPCW-TMLE-M	-0.01	-0.014	0.107	0.104	0.103	0.105	0.104	0.948	0.932	0.370	100
IPCW-TMLE-MTO	-0.015	-0.018	0.1	0.096	0.097	0.097	0.099	0.946	0.922	0.423	100
r-IPCW-TMLE-MTO	-0.016	-0.019	0.1	0.096	0.098	0.097	0.1	0.946	0.923	0.418	100

Table 19: **Synthetic data MAR simulation: oracle marginal risk difference (mRD), 5% outcome proportion, 40% missing proportion**. Comparing estimators under the **complex outcome and simple MAR** scenario. The value of the estimand is 0.015. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE. Estimators that are mismatched with the estimand (i.e., are estimating a different parameter) are emphasized using a star.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Nominal coverage	Oracle coverage	Power	Prop. completed
Benchmark model	0.002	0.001	0.044	0.046	0.043	0.046	0.043	0.953	0.945	0.936	100
Complete-case*	-0.092	-0.093	0.058	0.058	0.06	0.109	0.111	0.632	0.635	0.164	100
Confounded model*	-0.139	-0.139	0.048	0.049	0.047	0.147	0.147	0.182	0.165	0.058	100
IPW*	0.036	0.032	0.097	0.098	0.095	0.104	0.1	0.952	0.933	0.477	100
Raking (vanilla)*	0.036	0.035	0.059	0.06	0.059	0.069	0.068	0.920	0.908	0.898	100
MICE*	0.085	0.084	0.06	0.06	0.06	0.104	0.103	0.721	0.707	0.985	100
MI-RF*	0.045	0.045	0.058	0.056	0.057	0.072	0.073	0.880	0.883	0.946	100
IPCW-TMLE-M	-0.028	-0.045	0.144	0.124	0.126	0.127	0.134	0.872	0.958	0.099	100
IPCW-TMLE-MTO	-0.032	-0.04	0.109	0.098	0.104	0.103	0.112	0.883	0.946	0.201	100
r-IPCW-TMLE-MTO	-0.014	-0.022	0.112	0.099	0.109	0.1	0.111	0.900	0.949	0.263	100

Table 20: **Synthetic data MAR simulation: census marginal risk difference (mRD), 5% outcome proportion, 40% missing proportion**. Comparing estimators under the **complex outcome and simple MAR** scenario. The value of the estimand is 0.019. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE. Estimators that are mismatched with the estimand (i.e., are estimating a different parameter) are emphasized using a star.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Oracle coverage	Nominal coverage	Power	Prop. completed
Benchmark model	-0.003	-0.003	0.05	0.052	0.051	0.052	0.051	0.948	0.953	0.971	100
Complete-case	-0.132	-0.132	0.058	0.058	0.06	0.144	0.145	0.370	0.381	0.164	100
Confounded model	-0.178	-0.179	0.048	0.049	0.047	0.185	0.185	0.046	0.051	0.058	100
IPW	-0.003	-0.008	0.097	0.098	0.095	0.098	0.095	0.952	0.945	0.477	100
Raking (vanilla)	-0.004	-0.005	0.059	0.06	0.059	0.06	0.059	0.946	0.950	0.898	100
MICE	0.046	0.045	0.06	0.06	0.06	0.076	0.075	0.884	0.892	0.985	100
MI-RF	0.005	0.006	0.058	0.056	0.057	0.056	0.058	0.948	0.941	0.946	100
IPCW-TMLE-M*	-0.067	-0.085	0.144	0.124	0.126	0.141	0.152	0.949	0.797	0.099	100
IPCW-TMLE-MTO*	-0.072	-0.08	0.109	0.098	0.104	0.121	0.131	0.910	0.784	0.201	100
r-IPCW-TMLE-MTO*	-0.054	-0.061	0.112	0.099	0.109	0.113	0.125	0.931	0.823	0.263	100

Table 21: **Synthetic data MAR simulation: oracle marginal risk difference (mRD), 5% outcome proportion, 40% missing proportion**. Comparing estimators under the **simple outcome and complex MAR** scenario. The value of the estimand is 0.019. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Nominal coverage	Oracle coverage	Power	Prop. completed
Benchmark model	-0.004	-0.003	0.047	0.048	0.046	0.048	0.046	0.948	0.943	0.968	100
Complete-case	0.101	0.102	0.062	0.063	0.062	0.119	0.119	0.654	0.632	0.998	100
Confounded model	0.108	0.108	0.048	0.049	0.047	0.119	0.118	0.402	0.389	1.000	100
IPW	0.063	0.062	0.064	0.067	0.063	0.092	0.088	0.867	0.832	0.976	100
Raking (vanilla)	-0.006	-0.005	0.049	0.049	0.049	0.05	0.05	0.943	0.943	0.958	100
MICE	-0.004	-0.003	0.049	0.05	0.048	0.05	0.048	0.951	0.946	0.963	100
MI-RF	-0.002	-0.001	0.05	0.05	0.05	0.05	0.05	0.952	0.948	0.962	100
IPCW-TMLE-M	0.032	0.031	0.073	0.08	0.073	0.086	0.079	0.970	0.928	0.853	100
IPCW-TMLE-MTO	0.037	0.035	0.07	0.074	0.07	0.083	0.078	0.952	0.912	0.900	100
IPCW-a-TMLE-M	0.035	0.033	0.074	0.08	0.074	0.087	0.081	0.964	0.919	0.854	100
IPCW-a-TMLE-MTO	0.04	0.039	0.072	0.074	0.069	0.084	0.079	0.940	0.913	0.911	100
r-IPCW-TMLE-MTO	0.034	0.033	0.069	0.074	0.07	0.081	0.077	0.952	0.920	0.894	100

Table 22: **Synthetic data MAR simulation: census marginal risk difference (mRD), 5% outcome proportion, 40% missing proportion**. Comparing estimators under the **simple outcome and complex MAR** scenario. The value of the estimand is 0.019. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Oracle coverage	Nominal coverage	Power	Prop. completed
Benchmark model	-0.001	0	0.047	0.048	0.046	0.048	0.046	0.945	0.950	0.968	100
Complete-case	0.104	0.105	0.062	0.063	0.062	0.122	0.122	0.615	0.633	0.998	100
Confounded model	0.111	0.111	0.048	0.049	0.047	0.122	0.12	0.365	0.378	1.000	100
IPW	0.066	0.065	0.064	0.067	0.063	0.094	0.09	0.823	0.857	0.976	100
Raking (vanilla)	-0.003	-0.002	0.049	0.049	0.049	0.049	0.05	0.946	0.946	0.958	100
MICE	-0.001	0	0.049	0.05	0.048	0.05	0.048	0.948	0.955	0.963	100
MI-RF	0.001	0.002	0.05	0.05	0.05	0.05	0.05	0.946	0.948	0.962	100
IPCW-TMLE-M	0.035	0.034	0.073	0.08	0.073	0.087	0.081	0.924	0.970	0.853	100
IPCW-TMLE-MTO	0.041	0.039	0.07	0.074	0.07	0.084	0.08	0.906	0.946	0.900	100
IPCW-a-TMLE-M	0.038	0.036	0.074	0.08	0.074	0.088	0.083	0.917	0.962	0.854	100
IPCW-a-TMLE-MTO	0.044	0.042	0.072	0.074	0.069	0.085	0.081	0.906	0.935	0.911	100
r-IPCW-TMLE-MTO	0.037	0.036	0.069	0.074	0.07	0.083	0.078	0.917	0.948	0.894	100

Table 23: **Synthetic data MAR simulation: oracle marginal risk difference (mRD), 5% outcome proportion, 40% missing proportion**. Comparing estimators under the **complex outcome and complex MAR** scenario. The value of the estimand is 0.015. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE. Estimators that are mismatched with the estimand (i.e., are estimating a different parameter) are emphasized using a star.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Nominal coverage	Oracle coverage	Power	Prop. completed
Benchmark model	0	0.001	0.046	0.046	0.043	0.046	0.043	0.943	0.943	0.927	100
Complete-case*	0.168	0.169	0.073	0.073	0.072	0.183	0.183	0.352	0.367	0.996	100
Confounded model*	-0.141	-0.141	0.048	0.049	0.046	0.15	0.148	0.168	0.160	0.051	100
IPW*	0.084	0.082	0.067	0.069	0.064	0.108	0.104	0.803	0.772	0.952	100
Raking (vanilla)*	0.041	0.04	0.054	0.052	0.052	0.066	0.066	0.874	0.875	0.960	100
MICE*	0.074	0.073	0.054	0.055	0.053	0.092	0.091	0.751	0.730	0.988	100
MI-RF*	0.047	0.046	0.054	0.055	0.052	0.072	0.07	0.868	0.861	0.959	100
IPCW-TMLE-M	0.036	0.027	0.094	0.093	0.089	0.099	0.093	0.962	0.936	0.538	100
IPCW-TMLE-MTO	0.025	0.021	0.076	0.072	0.075	0.077	0.078	0.939	0.935	0.713	100
IPCW-a-TMLE-M	0.043	0.034	0.096	0.093	0.091	0.102	0.097	0.960	0.928	0.567	100
IPCW-a-TMLE-MTO	0.032	0.029	0.077	0.072	0.076	0.079	0.081	0.930	0.928	0.744	100
r-IPCW-TMLE-MTO	0.053	0.05	0.08	0.074	0.08	0.091	0.094	0.893	0.899	0.797	100

Table 24: **Synthetic data MAR simulation: census marginal risk difference (mRD), 5% outcome proportion, 40% missing proportion**. Comparing estimators under the **complex outcome and complex MAR** scenario. The value of the estimand is 0.019. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE. Estimators that are mismatched with the estimand (i.e., are estimating a different parameter) are emphasized using a star.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Oracle coverage	Nominal coverage	Power	Prop. completed
Benchmark model	-0.004	-0.005	0.051	0.052	0.05	0.052	0.05	0.950	0.951	0.964	100
Complete-case	0.129	0.13	0.073	0.073	0.072	0.148	0.148	0.584	0.575	0.996	100
Confounded model	-0.181	-0.181	0.048	0.049	0.046	0.187	0.186	0.033	0.040	0.051	100
IPW	0.044	0.043	0.067	0.069	0.064	0.082	0.077	0.898	0.921	0.952	100
Raking (vanilla)	0.002	0.001	0.054	0.052	0.052	0.052	0.052	0.955	0.946	0.960	100
MICE	0.034	0.034	0.054	0.055	0.053	0.065	0.063	0.901	0.913	0.988	100
MI-RF	0.007	0.007	0.054	0.055	0.052	0.055	0.053	0.948	0.948	0.959	100
IPCW-TMLE-M*	-0.003	-0.013	0.094	0.093	0.089	0.093	0.09	0.951	0.938	0.538	100
IPCW-TMLE-MTO*	-0.014	-0.018	0.076	0.072	0.075	0.074	0.077	0.948	0.919	0.713	100
IPCW-a-TMLE-M*	0.004	-0.006	0.096	0.093	0.091	0.093	0.091	0.948	0.943	0.567	100
IPCW-a-TMLE-MTO*	-0.008	-0.011	0.077	0.072	0.076	0.073	0.077	0.952	0.929	0.744	100
r-IPCW-TMLE-MTO*	0.014	0.011	0.08	0.074	0.08	0.076	0.081	0.946	0.936	0.797	100

MAR: 5% outcome proportion, 80% missingness proportion

Table 25: **Synthetic data MAR simulation: oracle marginal risk difference (mRD), 5% outcome proportion, 80% missing proportion**. Comparing estimators under the **simple outcome and simple MAR** scenario. The value of the estimand is 0.019. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Nominal coverage	Oracle coverage	Power	Prop. completed
Benchmark model	-0.001	-0.001	0.048	0.048	0.048	0.048	0.048	0.949	0.946	0.968	100
Complete-case	-0.124	-0.126	0.076	0.076	0.075	0.146	0.146	0.610	0.628	0.101	100
Confounded model	0.111	0.11	0.05	0.049	0.047	0.122	0.12	0.383	0.393	1.000	100
IPW	-0.115	-0.122	0.098	0.094	0.093	0.149	0.153	0.680	0.779	0.072	100
Raking (vanilla)	-0.005	-0.005	0.077	0.089	0.075	0.089	0.075	0.972	0.946	0.550	100
MICE	0	-0.002	0.067	0.066	0.067	0.066	0.067	0.938	0.950	0.810	100
MI-RF	0.061	0.06	0.059	0.055	0.06	0.082	0.085	0.790	0.827	0.988	100
IPCW-TMLE-M	-0.113	-0.129	0.12	0.102	0.102	0.153	0.164	0.633	0.868	0.068	100
IPCW-TMLE-MTO	-0.112	-0.123	0.106	0.09	0.098	0.144	0.158	0.625	0.814	0.108	100
r-IPCW-TMLE-MTO	-0.113	-0.125	0.109	0.091	0.098	0.145	0.159	0.622	0.828	0.108	100

Table 26: **Synthetic data MAR simulation: census marginal risk difference (mRD), 5% outcome proportion, 80% missing proportion**. Comparing estimators under the **simple outcome and simple MAR** scenario. The value of the estimand is 0.019. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Oracle coverage	Nominal coverage	Power	Prop. completed
Benchmark model	0.002	0.002	0.048	0.048	0.048	0.048	0.048	0.948	0.948	0.968	100
Complete-case	-0.121	-0.122	0.076	0.076	0.075	0.143	0.144	0.644	0.626	0.101	100
Confounded model	0.114	0.113	0.05	0.049	0.047	0.124	0.123	0.363	0.353	1.000	100
IPW	-0.112	-0.119	0.098	0.094	0.093	0.146	0.151	0.786	0.689	0.072	100
Raking (vanilla)	-0.002	-0.002	0.077	0.089	0.075	0.089	0.075	0.946	0.974	0.550	100
MICE	0.003	0.001	0.067	0.066	0.067	0.066	0.067	0.950	0.937	0.810	100
MI-RF	0.064	0.064	0.059	0.055	0.06	0.085	0.087	0.812	0.778	0.988	100
IPCW-TMLE-M	-0.11	-0.126	0.12	0.102	0.102	0.15	0.162	0.873	0.648	0.068	100
IPCW-TMLE-MTO	-0.109	-0.12	0.106	0.09	0.098	0.142	0.155	0.825	0.636	0.108	100
r-IPCW-TMLE-MTO	-0.11	-0.122	0.109	0.091	0.098	0.143	0.156	0.838	0.630	0.108	100

Table 27: **Synthetic data MAR simulation: oracle marginal risk difference (mRD), 5% outcome proportion, 80% missing proportion**. Comparing estimators under the **complex outcome and simple MAR** scenario. The value of the estimand is 0.015. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE. Estimators that are mismatched with the estimand (i.e., are estimating a different parameter) are emphasized using a star.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Nominal coverage	Oracle coverage	Power	Prop. completed
Benchmark model	0.002	0.002	0.046	0.046	0.047	0.046	0.047	0.955	0.955	0.936	100
Complete-case*	-0.09	-0.094	0.09	0.091	0.09	0.128	0.13	0.790	0.827	0.080	100
Confounded model*	-0.139	-0.139	0.049	0.049	0.047	0.147	0.147	0.192	0.181	0.054	100
IPW*	-0.083	-0.091	0.098	0.094	0.096	0.126	0.132	0.749	0.869	0.064	100
Raking (vanilla)*	0.034	0.035	0.092	0.1	0.092	0.106	0.099	0.959	0.932	0.466	100
MICE*	0.105	0.104	0.088	0.081	0.088	0.132	0.136	0.744	0.772	0.890	100
MI-RF*	0.019	0.016	0.074	0.061	0.074	0.064	0.076	0.889	0.944	0.771	100
IPCW-TMLE-M	-0.104	-0.133	0.157	0.114	0.108	0.154	0.172	0.613	0.960	0.058	100
IPCW-TMLE-MTO	-0.1	-0.115	0.115	0.094	0.1	0.138	0.152	0.647	0.878	0.074	100
r-IPCW-TMLE-MTO	-0.091	-0.109	0.123	0.095	0.109	0.131	0.155	0.665	0.910	0.112	100

Table 28: **Synthetic data MAR simulation: census marginal risk difference (mRD), 5% outcome proportion, 80% missing proportion**. Comparing estimators under the **complex outcome and simple MAR** scenario. The value of the estimand is 0.019. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE. Estimators that are mismatched with the estimand (i.e., are estimating a different parameter) are emphasized using a star.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Oracle coverage	Nominal coverage	Power	Prop. completed
Benchmark model	-0.003	-0.002	0.052	0.052	0.05	0.052	0.05	0.953	0.949	0.962	100
Complete-case	-0.129	-0.134	0.09	0.091	0.09	0.158	0.161	0.689	0.661	0.080	100
Confounded model	-0.178	-0.179	0.049	0.049	0.047	0.185	0.185	0.044	0.049	0.054	100
IPW	-0.123	-0.13	0.098	0.094	0.096	0.154	0.162	0.740	0.636	0.064	100
Raking (vanilla)	-0.006	-0.004	0.092	0.1	0.092	0.101	0.092	0.948	0.963	0.466	100
MICE	0.065	0.065	0.088	0.081	0.088	0.104	0.109	0.881	0.856	0.890	100
MI-RF	-0.021	-0.023	0.074	0.061	0.074	0.065	0.078	0.941	0.872	0.771	100
IPCW-TMLE-M*	-0.143	-0.173	0.157	0.114	0.108	0.183	0.204	0.924	0.506	0.058	100
IPCW-TMLE-MTO*	-0.14	-0.154	0.115	0.094	0.1	0.169	0.183	0.767	0.530	0.074	100
r-IPCW-TMLE-MTO*	-0.13	-0.149	0.123	0.095	0.109	0.161	0.184	0.822	0.560	0.112	100

Table 29: **Synthetic data MAR simulation: oracle marginal risk difference (mRD), 5% outcome proportion, 80% missing proportion**. Comparing estimators under the **simple outcome and complex MAR** scenario. The value of the estimand is 0.019. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Nominal coverage	Oracle coverage	Power	Prop. completed
Benchmark model	-0.002	-0.003	0.048	0.048	0.049	0.048	0.049	0.950	0.954	0.979	100
Complete-case	0.413	0.409	0.14	0.137	0.142	0.435	0.433	0.123	0.159	0.996	100
Confounded model	0.11	0.109	0.049	0.049	0.05	0.121	0.12	0.397	0.394	1.000	100
IPW	0.16	0.157	0.117	0.117	0.115	0.198	0.195	0.756	0.733	0.890	100
Raking (vanilla)	-0.008	-0.009	0.064	0.066	0.064	0.066	0.064	0.955	0.950	0.796	100
MICE	0	-0.001	0.057	0.057	0.057	0.057	0.057	0.954	0.951	0.918	100
MI-RF	0.016	0.015	0.058	0.054	0.059	0.056	0.061	0.917	0.939	0.964	100
IPCW-TMLE-M	0.043	0.036	0.137	0.136	0.127	0.143	0.131	0.966	0.935	0.369	100
IPCW-TMLE-MTO	0.071	0.066	0.131	0.121	0.127	0.14	0.143	0.923	0.916	0.581	100
IPCW-a-TMLE-M	0.043	0.032	0.143	0.138	0.131	0.145	0.135	0.965	0.936	0.370	100
IPCW-a-TMLE-MTO	0.076	0.069	0.135	0.122	0.132	0.144	0.149	0.914	0.911	0.592	100
r-IPCW-TMLE-MTO	0.053	0.048	0.128	0.122	0.125	0.133	0.134	0.940	0.929	0.514	100

Table 30: **Synthetic data MAR simulation: census marginal risk difference (mRD), 5% outcome proportion, 80% missing proportion**. Comparing estimators under the **simple outcome and complex MAR** scenario. The value of the estimand is 0.019. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Oracle coverage	Nominal coverage	Power	Prop. completed
Benchmark model	0.001	0	0.048	0.048	0.049	0.048	0.049	0.952	0.952	0.979	100
Complete-case	0.416	0.412	0.14	0.137	0.142	0.438	0.436	0.152	0.118	0.996	100
Confounded model	0.113	0.113	0.049	0.049	0.05	0.124	0.123	0.369	0.371	1.000	100
IPW	0.163	0.16	0.117	0.117	0.115	0.2	0.197	0.726	0.748	0.890	100
Raking (vanilla)	-0.005	-0.006	0.064	0.066	0.064	0.066	0.064	0.951	0.956	0.796	100
MICE	0.003	0.002	0.057	0.057	0.057	0.057	0.057	0.951	0.950	0.918	100
MI-RF	0.019	0.019	0.058	0.054	0.059	0.057	0.062	0.935	0.914	0.964	100
IPCW-TMLE-M	0.046	0.039	0.137	0.136	0.127	0.144	0.132	0.934	0.966	0.369	100
IPCW-TMLE-MTO	0.074	0.069	0.131	0.121	0.127	0.142	0.144	0.914	0.920	0.581	100
IPCW-a-TMLE-M	0.046	0.035	0.143	0.138	0.131	0.145	0.136	0.935	0.965	0.370	100
IPCW-a-TMLE-MTO	0.079	0.072	0.135	0.122	0.132	0.145	0.151	0.909	0.912	0.592	100
r-IPCW-TMLE-MTO	0.057	0.051	0.128	0.122	0.125	0.134	0.135	0.927	0.938	0.514	100

Table 31: **Synthetic data MAR simulation: oracle marginal risk difference (mRD), 5% outcome proportion, 80% missing proportion**. Comparing estimators under the **complex outcome and complex MAR** scenario. The value of the estimand is 0.015. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE. Estimators that are mismatched with the estimand (i.e., are estimating a different parameter) are emphasized using a star.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Nominal coverage	Oracle coverage	Power	Prop. completed
Benchmark model	0.001	0	0.046	0.046	0.046	0.046	0.046	0.950	0.951	0.930	100
Complete-case*	0.517	0.513	0.167	0.162	0.164	0.541	0.539	0.089	0.125	0.991	100
Confounded model*	-0.141	-0.142	0.049	0.049	0.05	0.149	0.151	0.178	0.175	0.058	100
IPW*	0.101	0.098	0.105	0.105	0.1	0.146	0.14	0.870	0.841	0.702	100
Raking (vanilla)*	0.048	0.045	0.075	0.076	0.075	0.09	0.087	0.915	0.901	0.778	100
MICE*	0.151	0.15	0.071	0.066	0.07	0.165	0.165	0.378	0.434	0.998	100
MI-RF*	0.098	0.096	0.068	0.06	0.068	0.115	0.117	0.632	0.704	0.982	100
IPCW-TMLE-M	0.055	0.039	0.16	0.15	0.144	0.16	0.149	0.960	0.944	0.227	100
IPCW-TMLE-MTO	0.056	0.047	0.127	0.115	0.123	0.128	0.132	0.926	0.926	0.436	100
IPCW-a-TMLE-M	0.06	0.047	0.164	0.151	0.146	0.162	0.153	0.957	0.944	0.246	100
IPCW-a-TMLE-MTO	0.072	0.063	0.129	0.117	0.125	0.137	0.14	0.910	0.914	0.489	100
r-IPCW-TMLE-MTO	0.11	0.104	0.137	0.122	0.133	0.165	0.169	0.850	0.878	0.580	100

Table 32: **Synthetic data MAR simulation: census marginal risk difference (mRD), 5% outcome proportion, 80% missing proportion**. Comparing estimators under the **complex outcome and complex MAR** scenario. The value of the estimand is 0.019. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE. Estimators that are mismatched with the estimand (i.e., are estimating a different parameter) are emphasized using a star.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Oracle coverage	Nominal coverage	Power	Prop. completed
Benchmark model	-0.004	-0.004	0.052	0.052	0.051	0.052	0.052	0.950	0.944	0.955	100
Complete-case	0.477	0.474	0.167	0.162	0.164	0.504	0.502	0.180	0.148	0.991	100
Confounded model	-0.18	-0.182	0.049	0.049	0.05	0.187	0.188	0.048	0.050	0.058	100
IPW	0.062	0.059	0.105	0.105	0.1	0.122	0.116	0.903	0.928	0.702	100
Raking (vanilla)	0.009	0.006	0.075	0.076	0.075	0.076	0.075	0.950	0.956	0.778	100
MICE	0.112	0.11	0.071	0.066	0.07	0.13	0.131	0.661	0.619	0.998	100
MI-RF	0.058	0.056	0.068	0.06	0.068	0.084	0.089	0.865	0.820	0.982	100
IPCW-TMLE-M*	0.015	0	0.16	0.15	0.144	0.151	0.144	0.955	0.952	0.227	100
IPCW-TMLE-MTO*	0.017	0.008	0.127	0.115	0.123	0.117	0.123	0.948	0.933	0.436	100
IPCW-a-TMLE-M*	0.021	0.008	0.164	0.151	0.146	0.152	0.146	0.955	0.950	0.246	100
IPCW-a-TMLE-MTO*	0.033	0.024	0.129	0.117	0.125	0.121	0.128	0.938	0.932	0.489	100
r-IPCW-TMLE-MTO*	0.071	0.065	0.137	0.122	0.133	0.141	0.148	0.918	0.898	0.580	100

MNAR: 12% outcome proportion, 40% missingness proportion

Table 33: **Synthetic data MNAR simulation: oracle marginal risk difference (mRD), 12% outcome proportion, 40% missing proportion**. Comparing estimators under the **simple outcome and MNAR-value** scenario. The value of the estimand is 0.04. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Nominal coverage	Oracle coverage	Power	Prop. completed
Benchmark model	-0.003	-0.001	0.073	0.071	0.073	0.071	0.073	0.943	0.950	1.000	100
Complete-case	-0.199	-0.199	0.084	0.085	0.084	0.216	0.216	0.344	0.336	0.678	100
Confounded model	0.235	0.237	0.075	0.072	0.075	0.246	0.248	0.109	0.126	1.000	100
IPW	-0.204	-0.206	0.09	0.091	0.089	0.223	0.224	0.390	0.379	0.593	100
Raking (vanilla)	-0.12	-0.119	0.08	0.086	0.08	0.148	0.144	0.724	0.683	0.922	100
MICE	-0.118	-0.116	0.08	0.077	0.08	0.141	0.141	0.663	0.685	0.953	100
MI-XGB	-0.119	-0.117	0.08	0.097	0.083	0.154	0.144	0.804	0.688	0.860	100
MI-RF	-0.116	-0.115	0.078	0.074	0.079	0.138	0.14	0.654	0.690	0.966	100
IPCW-TMLE-M	-0.201	-0.205	0.11	0.107	0.105	0.227	0.23	0.481	0.539	0.479	100
IPCW-TMLE-MTO	-0.202	-0.204	0.103	0.098	0.098	0.224	0.226	0.433	0.492	0.548	100

Table 34: **Synthetic data MNAR simulation: census marginal risk difference (mRD), 12% outcome proportion, 40% missing proportion**. Comparing estimators under the **simple outcome and MNAR-value** scenario. The value of the estimand is 0.041. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Oracle coverage	Nominal coverage	Power	Prop. completed
Benchmark model	-0.006	-0.004	0.073	0.071	0.073	0.071	0.073	0.948	0.940	1.000	100
Complete-case	-0.202	-0.202	0.084	0.085	0.084	0.219	0.219	0.321	0.330	0.678	100
Confounded model	0.232	0.233	0.075	0.072	0.075	0.243	0.245	0.130	0.119	1.000	100
IPW	-0.207	-0.209	0.09	0.091	0.089	0.226	0.227	0.363	0.380	0.593	100
Raking (vanilla)	-0.124	-0.122	0.08	0.086	0.08	0.15	0.146	0.672	0.710	0.922	100
MICE	-0.122	-0.119	0.08	0.077	0.08	0.144	0.143	0.672	0.650	0.953	100
MI-XGB	-0.123	-0.121	0.08	0.097	0.083	0.157	0.146	0.679	0.794	0.860	100
MI-RF	-0.119	-0.118	0.078	0.074	0.079	0.141	0.142	0.674	0.640	0.966	100
IPCW-TMLE-M	-0.204	-0.208	0.11	0.107	0.105	0.23	0.233	0.527	0.470	0.479	100
IPCW-TMLE-MTO	-0.205	-0.207	0.103	0.098	0.098	0.227	0.229	0.479	0.423	0.548	100

Table 35: **Synthetic data MNAR simulation: oracle marginal risk difference (mRD), 12% outcome proportion, 40% missing proportion**. Comparing estimators under the **complex outcome and MNAR-value** scenario. The value of the estimand is 0.031. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE. Estimators that are mismatched with the estimand (i.e., are estimating a different parameter) are emphasized using a star.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Nominal coverage	Oracle coverage	Power	Prop. completed
Benchmark model	-0.002	-0.003	0.064	0.064	0.064	0.064	0.064	0.944	0.949	0.998	100
Complete-case*	-0.156	-0.157	0.081	0.081	0.079	0.176	0.176	0.502	0.505	0.494	100
Confounded model*	-0.224	-0.226	0.07	0.069	0.069	0.235	0.237	0.108	0.106	0.254	100
IPW*	-0.157	-0.159	0.082	0.082	0.081	0.177	0.178	0.509	0.514	0.484	100
Raking (vanilla)*	-0.406	-0.405	0.078	0.087	0.076	0.415	0.413	0.002	0.001	0.152	100
MICE*	-0.405	-0.407	0.073	0.073	0.073	0.412	0.413	0.000	0.000	0.242	100
MI-XGB*	-0.299	-0.299	0.08	0.084	0.077	0.31	0.309	0.057	0.040	0.046	100
MI-RF*	-0.425	-0.427	0.075	0.072	0.072	0.431	0.434	0.000	0.000	0.340	100
IPCW-TMLE-M	-0.183	-0.189	0.104	0.099	0.098	0.209	0.213	0.486	0.553	0.230	100
IPCW-TMLE-MTO	-0.169	-0.173	0.09	0.085	0.087	0.189	0.193	0.464	0.517	0.383	100
IPCW-a-TMLE-M	-0.183	-0.191	0.104	0.099	0.099	0.209	0.215	0.488	0.551	0.226	100
IPCW-a-TMLE-MTO	-0.168	-0.17	0.085	0.08	0.085	0.186	0.19	0.431	0.486	0.448	100

Table 36: **Synthetic data MNAR simulation: census marginal risk difference (mRD), 12% outcome proportion, 40% missing proportion**. Comparing estimators under the **complex outcome and MNAR-value** scenario. The value of the estimand is 0.037. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE. Estimators that are mismatched with the estimand (i.e., are estimating a different parameter) are emphasized using a star.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Oracle coverage	Nominal coverage	Power	Prop. completed
Benchmark model	0.001	0	0.073	0.072	0.071	0.073	0.071	0.950	0.948	0.999	100
Complete-case	-0.209	-0.21	0.081	0.081	0.079	0.224	0.224	0.265	0.267	0.494	100
Confounded model	-0.277	-0.279	0.07	0.069	0.069	0.286	0.287	0.022	0.022	0.254	100
IPW	-0.209	-0.211	0.082	0.082	0.081	0.225	0.226	0.269	0.276	0.484	100
Raking (vanilla)	-0.458	-0.458	0.078	0.087	0.076	0.466	0.464	0.000	0.000	0.152	100
MICE	-0.458	-0.46	0.073	0.073	0.073	0.464	0.465	0.000	0.000	0.242	100
MI-XGB	-0.351	-0.352	0.08	0.084	0.077	0.361	0.36	0.010	0.015	0.046	100
MI-RF	-0.477	-0.48	0.075	0.072	0.072	0.483	0.485	0.000	0.000	0.340	100
IPCW-TMLE-M*	-0.236	-0.242	0.104	0.099	0.098	0.256	0.261	0.360	0.321	0.230	100
IPCW-TMLE-MTO*	-0.221	-0.226	0.09	0.085	0.087	0.237	0.242	0.297	0.273	0.383	100
IPCW-a-TMLE-M*	-0.236	-0.243	0.104	0.099	0.099	0.256	0.262	0.360	0.323	0.226	100
IPCW-a-TMLE-MTO*	-0.22	-0.223	0.085	0.08	0.085	0.234	0.238	0.262	0.232	0.448	100

Table 37: **Synthetic data MNAR simulation: oracle marginal risk difference (mRD), 12% outcome proportion, 40% missing proportion**. Comparing estimators under the **simple outcome and MNAR-unobserved** scenario. The value of the estimand is 0.04. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Nominal coverage	Oracle coverage	Power	Prop. completed
Benchmark model	-0.003	-0.003	0.07	0.071	0.069	0.071	0.069	0.952	0.949	1.000	100
Complete-case	-0.161	-0.162	0.077	0.079	0.077	0.179	0.179	0.457	0.445	0.882	100
Confounded model	0.236	0.237	0.072	0.072	0.072	0.247	0.248	0.090	0.090	1.000	100
IPW	-0.16	-0.161	0.082	0.083	0.08	0.18	0.18	0.509	0.498	0.846	100
Raking (vanilla)	-0.01	-0.009	0.076	0.086	0.077	0.086	0.077	0.971	0.945	1.000	100
MICE	-0.009	-0.008	0.074	0.075	0.075	0.076	0.075	0.950	0.946	1.000	100
MI-XGB	-0.015	-0.014	0.077	0.078	0.077	0.08	0.078	0.948	0.944	1.000	100
MI-RF	-0.007	-0.007	0.076	0.074	0.077	0.075	0.077	0.942	0.948	1.000	100
IPCW-TMLE-M	-0.163	-0.166	0.098	0.096	0.094	0.189	0.191	0.559	0.596	0.726	100
IPCW-TMLE-MTO	-0.162	-0.165	0.095	0.09	0.09	0.186	0.188	0.530	0.585	0.776	100
IPCW-a-TMLE-M	-0.163	-0.165	0.098	0.096	0.093	0.189	0.189	0.562	0.599	0.728	100
IPCW-a-TMLE-MTO	-0.162	-0.165	0.093	0.09	0.091	0.186	0.188	0.531	0.577	0.777	100

Table 38: **Synthetic data MNAR simulation: census marginal risk difference (mRD), 12% outcome proportion, 40% missing proportion**. Comparing estimators under the **simple outcome and MNAR-unobserved** scenario. The value of the estimand is 0.041. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Oracle coverage	Nominal coverage	Power	Prop. completed
Benchmark model	-0.006	-0.006	0.07	0.071	0.069	0.071	0.07	0.950	0.950	1.000	100
Complete-case	-0.164	-0.166	0.077	0.079	0.077	0.182	0.182	0.426	0.440	0.882	100
Confounded model	0.233	0.234	0.072	0.072	0.072	0.244	0.245	0.096	0.098	1.000	100
IPW	-0.163	-0.164	0.082	0.083	0.08	0.183	0.183	0.479	0.497	0.846	100
Raking (vanilla)	-0.013	-0.013	0.076	0.086	0.077	0.087	0.078	0.942	0.968	1.000	100
MICE	-0.012	-0.012	0.074	0.075	0.075	0.076	0.076	0.946	0.948	1.000	100
MI-XGB	-0.018	-0.018	0.077	0.078	0.077	0.08	0.079	0.941	0.947	1.000	100
MI-RF	-0.011	-0.01	0.076	0.074	0.077	0.075	0.077	0.945	0.939	1.000	100
IPCW-TMLE-M	-0.166	-0.169	0.098	0.096	0.094	0.192	0.194	0.588	0.550	0.726	100
IPCW-TMLE-MTO	-0.166	-0.169	0.095	0.09	0.09	0.188	0.191	0.569	0.518	0.776	100
IPCW-a-TMLE-M	-0.166	-0.168	0.098	0.096	0.093	0.192	0.192	0.589	0.548	0.728	100
IPCW-a-TMLE-MTO	-0.166	-0.168	0.093	0.09	0.091	0.188	0.191	0.564	0.521	0.777	100

Table 39: **Synthetic data MNAR simulation: oracle marginal risk difference (mRD), 12% outcome proportion, 40% missing proportion**. Comparing estimators under the **simple outcome (unobserved covariate) and MNAR-unobserved** scenario. The value of the estimand is 0.038. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE. Estimators that are mismatched with the estimand (i.e., are estimating a different parameter) are emphasized using a star.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Nominal coverage	Oracle coverage	Power	Prop. completed
Benchmark model	-0.001	0.002	0.072	0.07	0.072	0.07	0.072	0.947	0.951	0.999	100
Complete-case*	-0.149	-0.147	0.076	0.076	0.075	0.167	0.165	0.506	0.507	0.855	100
Confounded model*	0.221	0.225	0.072	0.07	0.072	0.232	0.236	0.123	0.140	1.000	100
IPW*	-0.15	-0.149	0.079	0.08	0.078	0.17	0.168	0.540	0.543	0.820	100
Raking (vanilla)*	-0.008	-0.007	0.076	0.083	0.076	0.084	0.077	0.965	0.944	0.996	100
MICE*	-0.006	-0.005	0.075	0.073	0.075	0.073	0.075	0.944	0.949	0.998	100
MI-RF*	-0.004	-0.003	0.076	0.072	0.075	0.072	0.075	0.934	0.948	0.998	100
IPCW-TMLE-M	-0.153	-0.156	0.093	0.092	0.087	0.179	0.179	0.576	0.618	0.707	100
IPCW-TMLE-MTO	-0.153	-0.154	0.089	0.086	0.084	0.175	0.175	0.540	0.593	0.759	100
IPCW-a-TMLE-M	-0.153	-0.156	0.093	0.092	0.087	0.178	0.178	0.577	0.618	0.707	100
IPCW-a-TMLE-MTO	-0.153	-0.155	0.089	0.086	0.085	0.175	0.176	0.545	0.597	0.760	100

Table 40: **Synthetic data MNAR simulation: census marginal risk difference (mRD), 12% outcome proportion, 40% missing proportion**. Comparing estimators under the **simple outcome (unobserved covariate) and MNAR-unobserved** scenario. The value of the estimand is 0.038. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE. Estimators that are mismatched with the estimand (i.e., are estimating a different parameter) are emphasized using a star.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Oracle coverage	Nominal coverage	Power	Prop. completed
Benchmark model	-0.001	0	0.071	0.068	0.069	0.068	0.069	0.950	0.942	0.999	100
Complete-case	-0.149	-0.147	0.076	0.076	0.075	0.167	0.166	0.506	0.505	0.855	100
Confounded model	0.221	0.225	0.072	0.07	0.072	0.232	0.236	0.141	0.124	1.000	100
IPW	-0.15	-0.149	0.079	0.08	0.078	0.17	0.168	0.542	0.538	0.820	100
Raking (vanilla)	-0.008	-0.007	0.076	0.083	0.076	0.084	0.077	0.944	0.965	0.996	100
MICE	-0.006	-0.005	0.075	0.073	0.075	0.073	0.075	0.948	0.944	0.998	100
MI-RF	-0.004	-0.003	0.076	0.072	0.075	0.072	0.075	0.948	0.934	0.998	100
IPCW-TMLE-M*	-0.153	-0.156	0.093	0.092	0.087	0.179	0.179	0.618	0.575	0.707	100
IPCW-TMLE-MTO*	-0.153	-0.154	0.089	0.086	0.084	0.175	0.176	0.592	0.540	0.759	100
IPCW-a-TMLE-M*	-0.153	-0.156	0.093	0.092	0.087	0.179	0.179	0.617	0.575	0.707	100
IPCW-a-TMLE-MTO*	-0.153	-0.155	0.089	0.086	0.085	0.176	0.177	0.596	0.544	0.760	100

MNAR: 12% outcome proportion, 80% missingness proportion

Table 41: **Synthetic data MNAR simulation: oracle marginal risk difference (mRD), 12% outcome proportion, 80% missing proportion**. Comparing estimators under the **simple outcome and MNAR-value** scenario. The value of the estimand is 0.04. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Nominal coverage	Oracle coverage	Power	Prop. completed
Benchmark model	0.001	0.002	0.072	0.071	0.075	0.071	0.075	0.951	0.956	1.000	100
Complete-case	-0.24	-0.242	0.131	0.129	0.133	0.272	0.276	0.535	0.546	0.253	100
Confounded model	0.238	0.239	0.073	0.072	0.075	0.249	0.25	0.090	0.096	1.000	100
IPW	-0.23	-0.236	0.164	0.158	0.165	0.279	0.288	0.646	0.702	0.176	100
Raking (vanilla)	-0.111	-0.108	0.117	0.124	0.113	0.166	0.156	0.864	0.843	0.674	100
MICE	-0.115	-0.115	0.113	0.113	0.113	0.162	0.162	0.793	0.829	0.724	100
MI-XGB	-0.124	-0.122	0.113	0.122	0.109	0.173	0.163	0.843	0.812	0.638	100
MI-RF	-0.096	-0.096	0.104	0.081	0.101	0.125	0.14	0.728	0.849	0.923	100
IPCW-TMLE-M	-0.224	-0.251	0.217	0.187	0.193	0.292	0.317	0.632	0.832	0.109	100
IPCW-TMLE-MTO	-0.225	-0.24	0.187	0.163	0.181	0.277	0.3	0.608	0.766	0.180	100

Table 42: **Synthetic data MNAR simulation: census marginal risk difference (mRD), 12% outcome proportion, 80% missing proportion**. Comparing estimators under the **simple outcome and MNAR-value** scenario. The value of the estimand is 0.041. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Oracle coverage	Nominal coverage	Power	Prop. completed
Benchmark model	-0.003	-0.002	0.072	0.071	0.075	0.071	0.075	0.954	0.950	1.000	100
Complete-case	-0.243	-0.245	0.131	0.129	0.133	0.275	0.279	0.536	0.527	0.253	100
Confounded model	0.235	0.235	0.073	0.072	0.075	0.246	0.247	0.105	0.096	1.000	100
IPW	-0.234	-0.239	0.164	0.158	0.165	0.282	0.291	0.696	0.641	0.176	100
Raking (vanilla)	-0.114	-0.111	0.117	0.124	0.113	0.169	0.158	0.835	0.856	0.674	100
MICE	-0.119	-0.119	0.113	0.113	0.113	0.164	0.164	0.821	0.787	0.724	100
MI-XGB	-0.127	-0.125	0.113	0.122	0.109	0.176	0.166	0.803	0.834	0.638	100
MI-RF	-0.099	-0.1	0.104	0.081	0.101	0.128	0.142	0.844	0.721	0.923	100
IPCW-TMLE-M	-0.228	-0.254	0.217	0.187	0.193	0.295	0.319	0.829	0.624	0.109	100
IPCW-TMLE-MTO	-0.228	-0.243	0.187	0.163	0.181	0.28	0.303	0.758	0.602	0.180	100

Table 43: **Synthetic data MNAR simulation: oracle marginal risk difference (mRD), 12% outcome proportion, 80% missing proportion**. Comparing estimators under the **complex outcome and MNAR-value** scenario. The value of the estimand is 0.031. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE. Estimators that are mismatched with the estimand (i.e., are estimating a different parameter) are emphasized using a star.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Nominal coverage	Oracle coverage	Power	Prop. completed
Benchmark model	0	0	0.064	0.064	0.065	0.064	0.065	0.953	0.954	0.999	100
Complete-case*	-0.165	-0.168	0.133	0.133	0.133	0.212	0.214	0.757	0.762	0.201	100
Confounded model*	-0.223	-0.223	0.068	0.069	0.07	0.234	0.233	0.113	0.102	0.260	100
IPW*	-0.154	-0.157	0.156	0.152	0.157	0.216	0.222	0.784	0.833	0.162	100
Raking (vanilla)*	-0.534	-0.535	0.108	0.115	0.109	0.547	0.546	0.004	0.001	0.484	100
MICE*	-0.532	-0.533	0.101	0.096	0.101	0.54	0.542	0.012	0.001	0.630	100
MI-RF*	-0.552	-0.552	0.09	0.079	0.09	0.557	0.559	0.000	0.000	0.816	100
IPCW-TMLE-M	-0.172	-0.197	0.229	0.18	0.181	0.249	0.268	0.705	0.933	0.092	100
IPCW-TMLE-MTO	-0.176	-0.191	0.174	0.151	0.163	0.232	0.251	0.674	0.832	0.136	100
IPCW-a-TMLE-M	-0.172	-0.197	0.227	0.18	0.18	0.249	0.267	0.700	0.931	0.091	100
IPCW-a-TMLE-MTO	-0.178	-0.191	0.163	0.143	0.157	0.228	0.247	0.662	0.806	0.158	100

Table 44: **Synthetic data MNAR simulation: census marginal risk difference (mRD), 12% outcome proportion, 80% missing proportion**. Comparing estimators under the **complex outcome and MNAR-value** scenario. The value of the estimand is 0.037. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE. Estimators that are mismatched with the estimand (i.e., are estimating a different parameter) are emphasized using a star.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Oracle coverage	Nominal coverage	Power	Prop. completed
Benchmark model	0.002	0.001	0.072	0.072	0.071	0.073	0.071	0.952	0.956	1.000	100
Complete-case	-0.217	-0.221	0.133	0.133	0.133	0.255	0.258	0.615	0.614	0.201	100
Confounded model	-0.276	-0.275	0.068	0.069	0.07	0.284	0.284	0.017	0.020	0.260	100
IPW	-0.206	-0.209	0.156	0.152	0.157	0.256	0.262	0.730	0.678	0.162	100
Raking (vanilla)	-0.587	-0.588	0.108	0.115	0.109	0.598	0.598	0.000	0.002	0.484	100
MICE	-0.584	-0.585	0.101	0.096	0.101	0.592	0.594	0.000	0.008	0.630	100
MI-RF	-0.604	-0.605	0.09	0.079	0.09	0.609	0.611	0.000	0.000	0.816	100
IPCW-TMLE-M*	-0.224	-0.25	0.229	0.18	0.181	0.288	0.308	0.880	0.604	0.092	100
IPCW-TMLE-MTO*	-0.229	-0.244	0.174	0.151	0.163	0.274	0.293	0.736	0.565	0.136	100
IPCW-a-TMLE-M*	-0.225	-0.25	0.227	0.18	0.18	0.288	0.308	0.876	0.602	0.091	100
IPCW-a-TMLE-MTO*	-0.231	-0.243	0.163	0.143	0.157	0.271	0.29	0.696	0.554	0.158	100

Table 45: **Synthetic data MNAR simulation: oracle marginal risk difference (mRD), 12% outcome proportion, 80% missing proportion**. Comparing estimators under the **simple outcome and MNAR-unobserved** scenario. The value of the estimand is 0.04. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Nominal coverage	Oracle coverage	Power	Prop. completed
Benchmark model	-0.003	-0.003	0.07	0.071	0.071	0.071	0.071	0.952	0.949	1.000	100
Complete-case	-0.215	-0.214	0.117	0.118	0.121	0.245	0.246	0.555	0.552	0.347	100
Confounded model	0.235	0.235	0.071	0.072	0.072	0.246	0.245	0.093	0.091	1.000	100
IPW	-0.209	-0.216	0.142	0.142	0.139	0.253	0.257	0.652	0.686	0.245	100
Raking (vanilla)	-0.013	-0.011	0.11	0.122	0.107	0.122	0.107	0.965	0.946	0.916	100
MICE	-0.012	-0.013	0.095	0.095	0.095	0.096	0.096	0.945	0.946	0.980	100
MI-XGB	0.006	0.007	0.112	0.103	0.109	0.103	0.109	0.927	0.951	0.947	100
MI-RF	0.043	0.045	0.1	0.083	0.098	0.093	0.108	0.871	0.928	0.998	100
IPCW-TMLE-M	-0.213	-0.226	0.172	0.16	0.159	0.267	0.276	0.636	0.764	0.168	100
IPCW-TMLE-MTO	-0.211	-0.222	0.16	0.144	0.15	0.255	0.267	0.608	0.736	0.240	100
IPCW-a-TMLE-M	-0.213	-0.228	0.171	0.16	0.158	0.267	0.277	0.634	0.762	0.169	100
IPCW-a-TMLE-MTO	-0.211	-0.221	0.158	0.143	0.15	0.254	0.267	0.606	0.729	0.245	100

Table 46: **Synthetic data MNAR simulation: census marginal risk difference (mRD), 12% outcome proportion, 80% missing proportion**. Comparing estimators under the **simple outcome and MNAR-unobserved** scenario. The value of the estimand is 0.041. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Oracle coverage	Nominal coverage	Power	Prop. completed
Benchmark model	-0.006	-0.007	0.07	0.071	0.071	0.071	0.071	0.945	0.949	1.000	100
Complete-case	-0.218	-0.217	0.117	0.118	0.121	0.248	0.249	0.541	0.546	0.347	100
Confounded model	0.232	0.231	0.071	0.072	0.072	0.243	0.242	0.097	0.101	1.000	100
IPW	-0.212	-0.219	0.142	0.142	0.139	0.256	0.26	0.677	0.642	0.245	100
Raking (vanilla)	-0.016	-0.014	0.11	0.122	0.107	0.123	0.108	0.946	0.966	0.916	100
MICE	-0.015	-0.016	0.095	0.095	0.095	0.096	0.096	0.946	0.944	0.980	100
MI-XGB	0.002	0.004	0.112	0.103	0.109	0.103	0.109	0.952	0.927	0.947	100
MI-RF	0.04	0.042	0.1	0.083	0.098	0.092	0.107	0.934	0.874	0.998	100
IPCW-TMLE-M	-0.216	-0.229	0.172	0.16	0.159	0.269	0.279	0.755	0.626	0.168	100
IPCW-TMLE-MTO	-0.214	-0.225	0.16	0.144	0.15	0.258	0.27	0.728	0.599	0.240	100
IPCW-a-TMLE-M	-0.217	-0.231	0.171	0.16	0.158	0.269	0.28	0.755	0.627	0.169	100
IPCW-a-TMLE-MTO	-0.214	-0.224	0.158	0.143	0.15	0.257	0.27	0.721	0.598	0.245	100

Table 47: **Synthetic data MNAR simulation: oracle marginal risk difference (mRD), 12% outcome proportion, 80% missing proportion**. Comparing estimators under the **simple outcome (unobserved covariate) and MNAR-unobserved** scenario. The value of the estimand is 0.038. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE. Estimators that are mismatched with the estimand (i.e., are estimating a different parameter) are emphasized using a star.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Nominal coverage	Oracle coverage	Power	Prop. completed
Benchmark model	0.002	0.002	0.072	0.07	0.073	0.07	0.073	0.950	0.958	1.000	100
Complete-case*	-0.204	-0.204	0.111	0.113	0.109	0.233	0.232	0.562	0.559	0.323	100
Confounded model*	0.225	0.225	0.071	0.07	0.072	0.235	0.236	0.109	0.117	1.000	100
IPW*	-0.199	-0.2	0.135	0.136	0.134	0.241	0.241	0.654	0.685	0.224	100
Raking (vanilla)*	-0.009	-0.01	0.105	0.118	0.107	0.118	0.107	0.974	0.952	0.897	100
MICE*	-0.009	-0.01	0.094	0.092	0.094	0.092	0.095	0.941	0.948	0.975	100
MI-XGB*	0.003	0.003	0.1	0.094	0.101	0.094	0.101	0.932	0.950	0.969	100
MI-RF*	0.049	0.047	0.097	0.08	0.099	0.094	0.11	0.856	0.923	0.997	100
IPCW-TMLE-M	-0.201	-0.212	0.168	0.154	0.155	0.253	0.263	0.638	0.777	0.144	100
IPCW-TMLE-MTO	-0.199	-0.207	0.153	0.137	0.144	0.242	0.252	0.617	0.746	0.226	100
IPCW-a-TMLE-M	-0.201	-0.213	0.168	0.153	0.154	0.253	0.262	0.639	0.777	0.147	100
IPCW-a-TMLE-MTO	-0.2	-0.207	0.151	0.136	0.141	0.242	0.251	0.612	0.742	0.225	100

Table 48: **Synthetic data MNAR simulation: census marginal risk difference (mRD), 12% outcome proportion, 80% missing proportion**. Comparing estimators under the **simple outcome (unobserved covariate) and MNAR-unobserved** scenario. The value of the estimand is 0.038. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE. Estimators that are mismatched with the estimand (i.e., are estimating a different parameter) are emphasized using a star.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Oracle coverage	Nominal coverage	Power	Prop. completed
Benchmark model	0.002	0.002	0.07	0.068	0.071	0.068	0.071	0.954	0.948	0.999	100
Complete-case	-0.204	-0.204	0.111	0.113	0.109	0.233	0.232	0.559	0.562	0.323	100
Confounded model	0.224	0.225	0.071	0.07	0.072	0.235	0.236	0.118	0.109	1.000	100
IPW	-0.199	-0.2	0.135	0.136	0.134	0.241	0.241	0.685	0.654	0.224	100
Raking (vanilla)	-0.01	-0.01	0.105	0.118	0.107	0.118	0.107	0.952	0.974	0.897	100
MICE	-0.009	-0.011	0.094	0.092	0.094	0.092	0.095	0.948	0.941	0.975	100
MI-XGB	0.003	0.003	0.1	0.094	0.101	0.094	0.101	0.950	0.933	0.969	100
MI-RF	0.049	0.047	0.097	0.08	0.099	0.093	0.11	0.924	0.857	0.997	100
IPCW-TMLE-M*	-0.201	-0.212	0.168	0.154	0.155	0.253	0.263	0.776	0.638	0.144	100
IPCW-TMLE-MTO*	-0.199	-0.207	0.153	0.137	0.144	0.242	0.252	0.746	0.616	0.226	100
IPCW-a-TMLE-M*	-0.201	-0.213	0.168	0.153	0.154	0.253	0.263	0.776	0.638	0.147	100
IPCW-a-TMLE-MTO*	-0.2	-0.208	0.151	0.136	0.141	0.242	0.251	0.740	0.611	0.225	100

MNAR: 5% outcome proportion, 40% missingness proportion

Table 49: **Synthetic data MNAR simulation: oracle marginal risk difference (mRD), 5% outcome proportion, 40% missing proportion**. Comparing estimators under the **simple outcome and MNAR-value** scenario. The value of the estimand is 0.019. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Nominal coverage	Oracle coverage	Power	Prop. completed
Benchmark model	-0.001	-0.002	0.046	0.048	0.045	0.048	0.045	0.958	0.950	0.984	100
Complete-case	-0.097	-0.097	0.055	0.056	0.056	0.112	0.112	0.591	0.580	0.375	100
Confounded model	0.111	0.109	0.047	0.049	0.047	0.121	0.119	0.390	0.362	1.000	100
IPW	-0.099	-0.098	0.06	0.06	0.059	0.115	0.115	0.610	0.622	0.321	100
Raking (vanilla)	-0.055	-0.056	0.052	0.059	0.051	0.081	0.076	0.881	0.809	0.622	100
MICE	-0.056	-0.056	0.051	0.052	0.051	0.076	0.076	0.808	0.802	0.726	100
MI-XGB	-0.055	-0.056	0.052	0.058	0.052	0.08	0.076	0.872	0.813	0.644	100
MI-RF	-0.051	-0.051	0.05	0.05	0.051	0.072	0.072	0.820	0.825	0.776	100
IPCW-TMLE-M	-0.099	-0.102	0.07	0.068	0.066	0.12	0.122	0.637	0.702	0.228	100
IPCW-TMLE-MTO	-0.098	-0.1	0.067	0.063	0.065	0.117	0.119	0.614	0.688	0.280	100

Table 50: **Synthetic data MNAR simulation: census marginal risk difference (mRD), 5% outcome proportion, 40% missing proportion**. Comparing estimators under the **simple outcome and MNAR-value** scenario. The value of the estimand is 0.019. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Oracle coverage	Nominal coverage	Power	Prop. completed
Benchmark model	0.002	0.001	0.046	0.048	0.045	0.048	0.045	0.950	0.958	0.984	100
Complete-case	-0.094	-0.094	0.055	0.056	0.056	0.109	0.109	0.599	0.610	0.375	100
Confounded model	0.114	0.112	0.047	0.049	0.047	0.124	0.121	0.336	0.357	1.000	100
IPW	-0.095	-0.095	0.06	0.06	0.059	0.113	0.112	0.640	0.631	0.321	100
Raking (vanilla)	-0.052	-0.053	0.052	0.059	0.051	0.079	0.074	0.828	0.896	0.622	100
MICE	-0.053	-0.053	0.051	0.052	0.051	0.074	0.074	0.821	0.823	0.726	100
MI-XGB	-0.052	-0.053	0.052	0.058	0.052	0.078	0.074	0.830	0.882	0.644	100
MI-RF	-0.048	-0.048	0.05	0.05	0.051	0.069	0.07	0.844	0.834	0.776	100
IPCW-TMLE-M	-0.096	-0.099	0.07	0.068	0.066	0.118	0.119	0.718	0.652	0.228	100
IPCW-TMLE-MTO	-0.095	-0.097	0.067	0.063	0.065	0.114	0.117	0.700	0.629	0.280	100

Table 51: **Synthetic data MNAR simulation: oracle marginal risk difference (mRD), 5% outcome proportion, 40% missing proportion**. Comparing estimators under the **complex outcome and MNAR-value** scenario. The value of the estimand is 0.015. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE. Estimators that are mismatched with the estimand (i.e., are estimating a different parameter) are emphasized using a star.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Nominal coverage	Oracle coverage	Power	Prop. completed
Benchmark model	0.002	0.002	0.046	0.046	0.045	0.046	0.045	0.943	0.944	0.928	100
Complete-case*	-0.076	-0.077	0.054	0.054	0.052	0.093	0.093	0.716	0.716	0.286	100
Confounded model*	-0.139	-0.138	0.048	0.049	0.048	0.148	0.146	0.184	0.173	0.055	100
IPW*	-0.078	-0.078	0.053	0.054	0.052	0.094	0.094	0.683	0.693	0.273	100
Raking (vanilla)*	-0.243	-0.243	0.056	0.067	0.057	0.252	0.25	0.032	0.012	0.237	100
MICE*	-0.256	-0.256	0.051	0.052	0.051	0.261	0.261	0.003	0.001	0.513	100
MI-RF*	-0.247	-0.246	0.053	0.051	0.053	0.252	0.252	0.001	0.001	0.457	100
IPCW-TMLE-M	-0.092	-0.096	0.069	0.065	0.062	0.113	0.114	0.609	0.736	0.120	100
IPCW-TMLE-MTO	-0.083	-0.085	0.059	0.057	0.057	0.101	0.103	0.633	0.705	0.207	100
IPCW-a-TMLE-M	-0.092	-0.097	0.069	0.065	0.061	0.113	0.114	0.606	0.735	0.119	100
IPCW-a-TMLE-MTO	-0.082	-0.085	0.055	0.052	0.052	0.098	0.1	0.607	0.666	0.254	100
r-IPCW-TMLE-MTO	-0.055	-0.057	0.067	0.059	0.065	0.081	0.087	0.768	0.866	0.379	100

Table 52: **Synthetic data MNAR simulation: census marginal risk difference (mRD), 5% outcome proportion, 40% missing proportion**. Comparing estimators under the **complex outcome and MNAR-value** scenario. The value of the estimand is 0.019. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE. Estimators that are mismatched with the estimand (i.e., are estimating a different parameter) are emphasized using a star.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Oracle coverage	Nominal coverage	Power	Prop. completed
Benchmark model	-0.003	-0.003	0.051	0.052	0.051	0.052	0.051	0.948	0.948	0.959	100
Complete-case	-0.115	-0.116	0.054	0.054	0.052	0.128	0.128	0.422	0.433	0.286	100
Confounded model	-0.179	-0.178	0.048	0.049	0.048	0.185	0.184	0.039	0.043	0.055	100
IPW	-0.117	-0.118	0.053	0.054	0.052	0.129	0.129	0.399	0.411	0.273	100
Raking (vanilla)	-0.282	-0.283	0.056	0.067	0.057	0.29	0.288	0.002	0.007	0.237	100
MICE	-0.295	-0.295	0.051	0.052	0.051	0.3	0.299	0.000	0.000	0.513	100
MI-RF	-0.286	-0.285	0.053	0.051	0.053	0.291	0.29	0.000	0.000	0.457	100
IPCW-TMLE-M*	-0.132	-0.136	0.069	0.065	0.062	0.147	0.149	0.495	0.393	0.120	100
IPCW-TMLE-MTO*	-0.122	-0.125	0.059	0.057	0.057	0.135	0.137	0.436	0.395	0.207	100
IPCW-a-TMLE-M*	-0.132	-0.136	0.069	0.065	0.061	0.147	0.149	0.493	0.391	0.119	100
IPCW-a-TMLE-MTO*	-0.122	-0.124	0.055	0.052	0.052	0.133	0.135	0.374	0.348	0.254	100
r-IPCW-TMLE-MTO*	-0.095	-0.096	0.067	0.059	0.065	0.111	0.117	0.697	0.574	0.379	100

Table 53: **Synthetic data MNAR simulation: oracle marginal risk difference (mRD), 5% outcome proportion, 40% missing proportion**. Comparing estimators under the **simple outcome and MNAR-unobserved** scenario. The value of the estimand is 0.019. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Nominal coverage	Oracle coverage	Power	Prop. completed
Benchmark model	-0.002	-0.002	0.048	0.048	0.048	0.048	0.048	0.947	0.950	0.974	100
Complete-case	-0.082	-0.082	0.052	0.051	0.051	0.096	0.097	0.639	0.648	0.540	100
Confounded model	0.11	0.108	0.049	0.049	0.05	0.12	0.119	0.391	0.397	1.000	100
IPW	-0.082	-0.082	0.054	0.054	0.053	0.098	0.098	0.659	0.673	0.495	100
Raking (vanilla)	-0.005	-0.006	0.052	0.06	0.052	0.06	0.052	0.975	0.947	0.901	100
MICE	-0.005	-0.005	0.051	0.051	0.051	0.051	0.051	0.948	0.949	0.951	100
MI-RF	0.007	0.006	0.052	0.051	0.052	0.051	0.052	0.944	0.952	0.968	100
IPCW-TMLE-M	-0.083	-0.086	0.063	0.062	0.061	0.104	0.106	0.674	0.735	0.385	100
IPCW-TMLE-MTO	-0.083	-0.085	0.061	0.058	0.058	0.101	0.103	0.652	0.722	0.440	100
IPCW-a-TMLE-M	-0.083	-0.086	0.063	0.062	0.061	0.104	0.105	0.672	0.734	0.383	100
IPCW-a-TMLE-MTO	-0.083	-0.085	0.06	0.058	0.059	0.101	0.104	0.654	0.718	0.438	100
r-IPCW-TMLE-MTO	-0.084	-0.086	0.061	0.058	0.059	0.102	0.104	0.646	0.714	0.434	100

Table 54: **Synthetic data MNAR simulation: census marginal risk difference (mRD), 5% outcome proportion, 40% missing proportion**. Comparing estimators under the **simple outcome and MNAR-unobserved** scenario. The value of the estimand is 0.019. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Oracle coverage	Nominal coverage	Power	Prop. completed
Benchmark model	0.001	0.001	0.048	0.048	0.048	0.048	0.048	0.952	0.948	0.974	100
Complete-case	-0.078	-0.079	0.052	0.051	0.051	0.094	0.094	0.674	0.662	0.540	100
Confounded model	0.113	0.111	0.049	0.049	0.05	0.123	0.122	0.372	0.367	1.000	100
IPW	-0.079	-0.079	0.054	0.054	0.053	0.096	0.095	0.695	0.682	0.495	100
Raking (vanilla)	-0.002	-0.003	0.052	0.06	0.052	0.06	0.052	0.950	0.978	0.901	100
MICE	-0.002	-0.002	0.051	0.051	0.051	0.051	0.051	0.948	0.948	0.951	100
MI-RF	0.01	0.009	0.052	0.051	0.052	0.052	0.053	0.948	0.943	0.968	100
IPCW-TMLE-M	-0.08	-0.083	0.063	0.062	0.061	0.101	0.103	0.753	0.690	0.385	100
IPCW-TMLE-MTO	-0.08	-0.082	0.061	0.058	0.058	0.098	0.1	0.740	0.671	0.440	100
IPCW-a-TMLE-M	-0.08	-0.083	0.063	0.062	0.061	0.101	0.103	0.750	0.689	0.383	100
IPCW-a-TMLE-MTO	-0.08	-0.082	0.06	0.058	0.059	0.098	0.101	0.740	0.667	0.438	100
r-IPCW-TMLE-MTO	-0.081	-0.083	0.061	0.058	0.059	0.099	0.102	0.734	0.668	0.434	100

Table 55: **Synthetic data MNAR simulation: oracle marginal risk difference (mRD), 5% outcome proportion, 40% missing proportion**. Comparing estimators under the **simple outcome (unobserved covariate) and MNAR-unobserved** scenario. The value of the estimand is 0.016. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE. Estimators that are mismatched with the estimand (i.e., are estimating a different parameter) are emphasized using a star.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Nominal coverage	Oracle coverage	Power	Prop. completed
Benchmark model	0	0	0.045	0.046	0.045	0.046	0.045	0.956	0.949	0.955	100
Complete-case*	-0.069	-0.068	0.047	0.048	0.048	0.084	0.083	0.697	0.689	0.508	100
Confounded model*	0.098	0.096	0.045	0.046	0.044	0.109	0.106	0.441	0.416	1.000	100
IPW*	-0.068	-0.069	0.051	0.051	0.051	0.085	0.086	0.715	0.734	0.460	100
Raking (vanilla)*	-0.002	-0.003	0.047	0.056	0.048	0.056	0.048	0.980	0.952	0.865	100
MICE*	-0.002	-0.003	0.046	0.048	0.046	0.048	0.046	0.956	0.951	0.939	100
MI-RF*	0.011	0.01	0.047	0.047	0.046	0.049	0.047	0.947	0.942	0.965	100
IPCW-TMLE-M	-0.07	-0.072	0.059	0.057	0.058	0.09	0.093	0.708	0.768	0.360	100
IPCW-TMLE-MTO	-0.07	-0.071	0.056	0.054	0.057	0.088	0.09	0.691	0.752	0.415	100
IPCW-a-TMLE-M	-0.07	-0.072	0.058	0.057	0.058	0.09	0.093	0.710	0.767	0.362	100
IPCW-a-TMLE-MTO	-0.07	-0.071	0.057	0.053	0.057	0.088	0.091	0.694	0.755	0.411	100
r-IPCW-TMLE-MTO	-0.07	-0.072	0.056	0.054	0.057	0.089	0.091	0.688	0.750	0.407	100

Table 56: **Synthetic data MNAR simulation: census marginal risk difference (mRD), 5% outcome proportion, 40% missing proportion**. Comparing estimators under the **simple outcome (unobserved covariate) and MNAR-unobserved** scenario. The value of the estimand is 0.016. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE. Estimators that are mismatched with the estimand (i.e., are estimating a different parameter) are emphasized using a star.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Oracle coverage	Nominal coverage	Power	Prop. completed
Benchmark model	0.001	0	0.044	0.045	0.043	0.045	0.043	0.951	0.958	0.966	100
Complete-case	-0.068	-0.068	0.047	0.048	0.048	0.084	0.083	0.692	0.703	0.508	100
Confounded model	0.099	0.097	0.045	0.046	0.044	0.109	0.107	0.412	0.436	1.000	100
IPW	-0.068	-0.069	0.051	0.051	0.051	0.085	0.086	0.735	0.719	0.460	100
Raking (vanilla)	-0.002	-0.002	0.047	0.056	0.048	0.056	0.048	0.952	0.980	0.865	100
MICE	-0.001	-0.002	0.046	0.048	0.046	0.048	0.046	0.951	0.956	0.939	100
MI-RF	0.012	0.011	0.047	0.047	0.046	0.049	0.047	0.941	0.947	0.965	100
IPCW-TMLE-M*	-0.07	-0.072	0.059	0.057	0.058	0.09	0.093	0.772	0.710	0.360	100
IPCW-TMLE-MTO*	-0.069	-0.07	0.056	0.054	0.057	0.088	0.09	0.754	0.695	0.415	100
IPCW-a-TMLE-M*	-0.07	-0.072	0.058	0.057	0.058	0.09	0.092	0.769	0.712	0.362	100
IPCW-a-TMLE-MTO*	-0.069	-0.07	0.057	0.053	0.057	0.088	0.091	0.759	0.698	0.411	100
r-IPCW-TMLE-MTO*	-0.07	-0.071	0.056	0.054	0.057	0.088	0.091	0.752	0.691	0.407	100

MNAR: 5% outcome proportion, 80% missingness proportion

Table 57: **Synthetic data MNAR simulation: oracle marginal risk difference (mRD), 5% outcome proportion, 80% missing proportion**. Comparing estimators under the **simple outcome and MNAR-value** scenario. The value of the estimand is 0.019. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Nominal coverage	Oracle coverage	Power	Prop. completed
Benchmark model	-0.002	-0.001	0.048	0.048	0.05	0.048	0.051	0.954	0.956	0.977	100
Complete-case	-0.116	-0.115	0.086	0.084	0.086	0.143	0.144	0.713	0.732	0.147	100
Confounded model	0.11	0.11	0.049	0.049	0.052	0.121	0.122	0.396	0.393	1.000	100
IPW	-0.113	-0.116	0.106	0.101	0.105	0.151	0.157	0.730	0.808	0.093	100
Raking (vanilla)	-0.055	-0.055	0.079	0.085	0.081	0.102	0.098	0.914	0.898	0.346	100
MICE	-0.057	-0.056	0.074	0.074	0.073	0.093	0.092	0.862	0.876	0.454	100
MI-XGB	-0.048	-0.047	0.075	0.071	0.075	0.085	0.088	0.885	0.900	0.530	100
MI-RF	-0.015	-0.014	0.067	0.055	0.07	0.057	0.072	0.881	0.944	0.830	100
IPCW-TMLE-M	-0.109	-0.126	0.149	0.116	0.114	0.159	0.17	0.674	0.937	0.071	100
IPCW-TMLE-MTO	-0.109	-0.119	0.12	0.101	0.106	0.149	0.16	0.676	0.864	0.104	100

Table 58: **Synthetic data MNAR simulation: census marginal risk difference (mRD), 5% outcome proportion, 80% missing proportion**. Comparing estimators under the **simple outcome and MNAR-value** scenario. The value of the estimand is 0.019. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Oracle coverage	Nominal coverage	Power	Prop. completed
Benchmark model	0.001	0.002	0.048	0.048	0.05	0.048	0.051	0.957	0.954	0.977	100
Complete-case	-0.113	-0.112	0.086	0.084	0.086	0.14	0.141	0.748	0.723	0.147	100
Confounded model	0.113	0.113	0.049	0.049	0.052	0.124	0.125	0.376	0.372	1.000	100
IPW	-0.11	-0.113	0.106	0.101	0.105	0.149	0.154	0.818	0.740	0.093	100
Raking (vanilla)	-0.052	-0.052	0.079	0.085	0.081	0.1	0.096	0.903	0.918	0.346	100
MICE	-0.054	-0.052	0.074	0.074	0.073	0.092	0.09	0.884	0.870	0.454	100
MI-XGB	-0.044	-0.044	0.075	0.071	0.075	0.083	0.087	0.905	0.892	0.530	100
MI-RF	-0.012	-0.011	0.067	0.055	0.07	0.057	0.071	0.944	0.888	0.830	100
IPCW-TMLE-M	-0.106	-0.123	0.149	0.116	0.114	0.157	0.167	0.942	0.685	0.071	100
IPCW-TMLE-MTO	-0.106	-0.116	0.12	0.101	0.106	0.147	0.158	0.871	0.684	0.104	100

Table 59: **Synthetic data MNAR simulation: oracle marginal risk difference (mRD), 5% outcome proportion, 80% missing proportion**. Comparing estimators under the **complex outcome and MNAR-value** scenario. The value of the estimand is 0.015. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE. Estimators that are mismatched with the estimand (i.e., are estimating a different parameter) are emphasized using a star.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Nominal coverage	Oracle coverage	Power	Prop. completed
Benchmark model	0.002	0.001	0.047	0.046	0.047	0.046	0.047	0.935	0.944	0.926	100
Complete-case*	-0.069	-0.071	0.092	0.092	0.092	0.115	0.117	0.882	0.885	0.147	100
Confounded model*	-0.139	-0.139	0.049	0.049	0.049	0.148	0.148	0.187	0.184	0.062	100
IPW*	-0.068	-0.072	0.105	0.102	0.103	0.123	0.125	0.860	0.899	0.111	100
Raking (vanilla)*	-0.335	-0.336	0.08	0.086	0.076	0.345	0.345	0.035	0.017	0.586	100
MICE*	-0.351	-0.354	0.069	0.065	0.07	0.357	0.36	0.006	0.001	0.832	100
MI-RF*	-0.331	-0.331	0.062	0.054	0.062	0.335	0.337	0.000	0.000	0.868	100
IPCW-TMLE-M	-0.082	-0.098	0.134	0.113	0.106	0.14	0.144	0.797	0.934	0.066	100
IPCW-TMLE-MTO	-0.083	-0.089	0.112	0.1	0.103	0.13	0.136	0.780	0.898	0.098	100
IPCW-a-TMLE-M	-0.082	-0.097	0.131	0.112	0.105	0.139	0.143	0.797	0.927	0.069	100
IPCW-a-TMLE-MTO	-0.082	-0.088	0.106	0.095	0.098	0.126	0.132	0.780	0.888	0.111	100
r-IPCW-TMLE-MTO	-0.075	-0.085	0.119	0.102	0.106	0.127	0.136	0.794	0.913	0.117	100

Table 60: **Synthetic data MNAR simulation: census marginal risk difference (mRD), 5% outcome proportion, 80% missing proportion**. Comparing estimators under the **complex outcome and MNAR-value** scenario. The value of the estimand is 0.019. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE. Estimators that are mismatched with the estimand (i.e., are estimating a different parameter) are emphasized using a star.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Oracle coverage	Nominal coverage	Power	Prop. completed
Benchmark model	-0.002	-0.004	0.053	0.052	0.052	0.052	0.053	0.949	0.946	0.966	100
Complete-case	-0.108	-0.111	0.092	0.092	0.092	0.142	0.144	0.771	0.773	0.147	100
Confounded model	-0.179	-0.179	0.049	0.049	0.049	0.185	0.185	0.048	0.051	0.062	100
IPW	-0.107	-0.111	0.105	0.102	0.103	0.148	0.151	0.817	0.769	0.111	100
Raking (vanilla)	-0.374	-0.375	0.08	0.086	0.076	0.384	0.383	0.006	0.019	0.586	100
MICE	-0.391	-0.393	0.069	0.065	0.07	0.396	0.399	0.000	0.004	0.832	100
MI-RF	-0.37	-0.37	0.062	0.054	0.062	0.374	0.376	0.000	0.000	0.868	100
IPCW-TMLE-M*	-0.122	-0.138	0.134	0.113	0.106	0.166	0.174	0.885	0.680	0.066	100
IPCW-TMLE-MTO*	-0.122	-0.129	0.112	0.1	0.103	0.158	0.164	0.820	0.671	0.098	100
IPCW-a-TMLE-M*	-0.122	-0.137	0.131	0.112	0.105	0.166	0.172	0.877	0.680	0.069	100
IPCW-a-TMLE-MTO*	-0.122	-0.127	0.106	0.095	0.098	0.154	0.161	0.796	0.669	0.111	100
r-IPCW-TMLE-MTO*	-0.115	-0.124	0.119	0.102	0.106	0.153	0.163	0.844	0.696	0.117	100

Table 61: **Synthetic data MNAR simulation: oracle marginal risk difference (mRD), 5% outcome proportion, 80% missing proportion**. Comparing estimators under the **simple outcome and MNAR-unobserved** scenario. The value of the estimand is 0.019. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Nominal coverage	Oracle coverage	Power	Prop. completed
Benchmark model	0	0	0.049	0.048	0.049	0.048	0.049	0.950	0.952	0.976	100
Complete-case	-0.108	-0.109	0.077	0.076	0.077	0.132	0.133	0.677	0.700	0.177	100
Confounded model	0.112	0.112	0.05	0.049	0.051	0.122	0.123	0.380	0.390	1.000	100
IPW	-0.103	-0.108	0.095	0.09	0.093	0.137	0.143	0.706	0.804	0.120	100
Raking (vanilla)	-0.006	-0.006	0.074	0.086	0.071	0.086	0.072	0.974	0.940	0.588	100
MICE	-0.005	-0.004	0.065	0.064	0.064	0.064	0.064	0.948	0.950	0.811	100
MI-RF	0.053	0.055	0.059	0.055	0.057	0.076	0.08	0.830	0.854	0.986	100
IPCW-TMLE-M	-0.101	-0.116	0.116	0.101	0.099	0.143	0.152	0.673	0.890	0.088	100
IPCW-TMLE-MTO	-0.101	-0.113	0.103	0.09	0.094	0.135	0.147	0.668	0.846	0.140	100
IPCW-a-TMLE-M	-0.101	-0.116	0.116	0.101	0.099	0.143	0.153	0.676	0.889	0.093	100
IPCW-a-TMLE-MTO	-0.101	-0.112	0.1	0.088	0.094	0.134	0.146	0.673	0.830	0.143	100
r-IPCW-TMLE-MTO	-0.102	-0.114	0.103	0.09	0.095	0.136	0.148	0.667	0.843	0.136	100

Table 62: **Synthetic data MNAR simulation: census marginal risk difference (mRD), 5% outcome proportion, 80% missing proportion**. Comparing estimators under the **simple outcome and MNAR-unobserved** scenario. The value of the estimand is 0.019. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Oracle coverage	Nominal coverage	Power	Prop. completed
Benchmark model	0.003	0.004	0.049	0.048	0.049	0.048	0.049	0.952	0.949	0.976	100
Complete-case	-0.105	-0.106	0.077	0.076	0.077	0.129	0.131	0.714	0.690	0.177	100
Confounded model	0.115	0.115	0.05	0.049	0.051	0.125	0.126	0.371	0.360	1.000	100
IPW	-0.099	-0.105	0.095	0.09	0.093	0.134	0.14	0.817	0.720	0.120	100
Raking (vanilla)	-0.003	-0.003	0.074	0.086	0.071	0.086	0.072	0.941	0.974	0.588	100
MICE	-0.002	-0.001	0.065	0.064	0.064	0.064	0.064	0.950	0.947	0.811	100
MI-RF	0.056	0.059	0.059	0.055	0.057	0.079	0.082	0.844	0.816	0.986	100
IPCW-TMLE-M	-0.098	-0.113	0.116	0.101	0.099	0.14	0.15	0.896	0.686	0.088	100
IPCW-TMLE-MTO	-0.098	-0.109	0.103	0.09	0.094	0.133	0.144	0.854	0.680	0.140	100
IPCW-a-TMLE-M	-0.098	-0.113	0.116	0.101	0.099	0.14	0.15	0.895	0.684	0.093	100
IPCW-a-TMLE-MTO	-0.098	-0.109	0.1	0.088	0.094	0.132	0.144	0.840	0.683	0.143	100
r-IPCW-TMLE-MTO	-0.098	-0.111	0.103	0.09	0.095	0.133	0.146	0.850	0.676	0.136	100

Table 63: **Synthetic data MNAR simulation: oracle marginal risk difference (mRD), 5% outcome proportion, 80% missing proportion**. Comparing estimators under the **simple outcome (unobserved covariate) and MNAR-unobserved** scenario. The value of the estimand is 0.016. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE. Estimators that are mismatched with the estimand (i.e., are estimating a different parameter) are emphasized using a star.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Nominal coverage	Oracle coverage	Power	Prop. completed
Benchmark model	0	0.001	0.046	0.046	0.048	0.046	0.048	0.951	0.949	0.948	100
Complete-case*	-0.092	-0.094	0.07	0.07	0.07	0.116	0.117	0.724	0.742	0.164	100
Confounded model*	0.098	0.097	0.046	0.046	0.047	0.108	0.108	0.439	0.438	1.000	100
IPW*	-0.089	-0.092	0.086	0.083	0.085	0.122	0.125	0.734	0.823	0.104	100
Raking (vanilla)*	-0.003	-0.003	0.07	0.081	0.07	0.081	0.07	0.976	0.948	0.515	100
MICE*	-0.002	-0.003	0.061	0.06	0.062	0.06	0.062	0.947	0.950	0.763	100
MI-XGB*	0.025	0.025	0.062	0.056	0.064	0.061	0.069	0.898	0.932	0.889	100
MI-RF*	0.053	0.053	0.054	0.051	0.054	0.073	0.076	0.813	0.833	0.984	100
IPCW-TMLE-M	-0.089	-0.101	0.105	0.091	0.091	0.128	0.136	0.681	0.889	0.076	100
IPCW-TMLE-MTO	-0.088	-0.096	0.095	0.082	0.086	0.12	0.129	0.683	0.858	0.118	100
IPCW-a-TMLE-M	-0.089	-0.102	0.106	0.091	0.091	0.128	0.136	0.682	0.893	0.074	100
IPCW-a-TMLE-MTO	-0.088	-0.095	0.093	0.081	0.086	0.119	0.128	0.684	0.854	0.130	100

Table 64: **Synthetic data MNAR simulation: census marginal risk difference (mRD), 5% outcome proportion, 80% missing proportion**. Comparing estimators under the **simple outcome (unobserved covariate) and MNAR-unobserved** scenario. The value of the estimand is 0.016. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE. Estimators that are mismatched with the estimand (i.e., are estimating a different parameter) are emphasized using a star.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Oracle coverage	Nominal coverage	Power	Prop. completed
Benchmark model	0.001	0	0.045	0.045	0.045	0.045	0.045	0.950	0.949	0.956	100
Complete-case	-0.092	-0.093	0.07	0.07	0.07	0.115	0.117	0.746	0.725	0.164	100
Confounded model	0.098	0.097	0.046	0.046	0.047	0.109	0.108	0.432	0.436	1.000	100
IPW	-0.088	-0.092	0.086	0.083	0.085	0.122	0.125	0.826	0.738	0.104	100
Raking (vanilla)	-0.003	-0.002	0.07	0.081	0.07	0.081	0.07	0.948	0.976	0.515	100
MICE	-0.001	-0.002	0.061	0.06	0.062	0.06	0.062	0.950	0.948	0.763	100
MI-XGB	0.025	0.026	0.062	0.056	0.064	0.061	0.069	0.932	0.896	0.889	100
MI-RF	0.053	0.053	0.054	0.051	0.054	0.074	0.076	0.829	0.811	0.984	100
IPCW-TMLE-M*	-0.089	-0.1	0.105	0.091	0.091	0.127	0.135	0.891	0.683	0.076	100
IPCW-TMLE-MTO*	-0.087	-0.096	0.095	0.082	0.086	0.12	0.129	0.860	0.685	0.118	100
IPCW-a-TMLE-M*	-0.089	-0.101	0.106	0.091	0.091	0.127	0.136	0.894	0.684	0.074	100
IPCW-a-TMLE-MTO*	-0.087	-0.094	0.093	0.081	0.086	0.119	0.128	0.856	0.686	0.130	100

Other scenarios

Table 65: **Synthetic data MAR simulation: oracle marginal risk difference (mRD), 12% outcome proportion, 40% missing proportion**. Comparing estimators under the **simple outcome (no treatment effect) and simple MAR (no dependence on Y)** scenario. The value of the estimand is 0. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Nominal coverage	Oracle coverage	Power	Prop. completed
Benchmark model	-0.002	-0.003	0.065	0.065	0.066	0.065	0.066	0.951	0.950	0.049	100
Complete-case	-0.003	-0.003	0.082	0.082	0.082	0.082	0.082	0.944	0.944	0.056	100
Confounded model	0.198	0.199	0.066	0.066	0.068	0.209	0.21	0.146	0.155	0.856	100
IPW	-0.003	-0.004	0.099	0.099	0.098	0.099	0.098	0.945	0.949	0.055	100
Raking (vanilla)	-0.002	-0.004	0.07	0.07	0.07	0.07	0.07	0.949	0.950	0.051	100
MICE	-0.002	-0.003	0.068	0.067	0.068	0.067	0.068	0.948	0.951	0.053	100
MI-XGB	-0.005	-0.006	0.07	0.07	0.071	0.07	0.071	0.948	0.947	0.052	100
MI-RF	0.007	0.007	0.069	0.068	0.069	0.068	0.07	0.946	0.949	0.054	100
IPCW-TMLE-M	-0.004	-0.006	0.124	0.117	0.123	0.117	0.123	0.937	0.952	0.063	100
IPCW-TMLE-MTO	-0.002	-0.004	0.114	0.106	0.113	0.106	0.113	0.933	0.949	0.066	100
IPCW-a-TMLE-M	-0.004	-0.007	0.123	0.117	0.123	0.117	0.123	0.936	0.950	0.064	100
IPCW-a-TMLE-MTO	-0.002	-0.005	0.112	0.105	0.11	0.105	0.11	0.932	0.949	0.068	100

Table 66: **Synthetic data MAR simulation: census marginal risk difference (mRD), 12% outcome proportion, 40% missing proportion**. Comparing estimators under the **simple outcome (no treatment effect) and simple MAR (no dependence on Y)** scenario. The value of the estimand is 0. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Oracle coverage	Nominal coverage	Power	Prop. completed
Benchmark model	-0.006	-0.007	0.065	0.065	0.066	0.065	0.066	0.950	0.950	0.049	100
Complete-case	-0.007	-0.007	0.082	0.082	0.082	0.082	0.083	0.946	0.943	0.056	100
Confounded model	0.194	0.195	0.066	0.066	0.068	0.205	0.207	0.167	0.162	0.856	100
IPW	-0.007	-0.008	0.099	0.099	0.098	0.099	0.098	0.950	0.944	0.055	100
Raking (vanilla)	-0.006	-0.008	0.07	0.07	0.07	0.07	0.071	0.950	0.946	0.051	100
MICE	-0.006	-0.007	0.068	0.067	0.068	0.068	0.068	0.949	0.946	0.053	100
MI-XGB	-0.009	-0.01	0.07	0.07	0.071	0.071	0.072	0.944	0.945	0.052	100
MI-RF	0.003	0.003	0.069	0.068	0.069	0.068	0.069	0.951	0.946	0.054	100
IPCW-TMLE-M	-0.008	-0.01	0.124	0.117	0.123	0.118	0.124	0.954	0.936	0.063	100
IPCW-TMLE-MTO	-0.006	-0.008	0.114	0.106	0.113	0.106	0.113	0.949	0.929	0.066	100
IPCW-a-TMLE-M	-0.008	-0.011	0.123	0.117	0.123	0.117	0.123	0.952	0.934	0.064	100
IPCW-a-TMLE-MTO	-0.006	-0.009	0.112	0.105	0.11	0.105	0.11	0.948	0.928	0.068	100

Table 67: **Synthetic data MAR simulation: oracle marginal risk difference (mRD), 12% outcome proportion, 40% missing proportion**. Comparing estimators under the **simple outcome (no treatment effect) and simple MAR** scenario. The value of the estimand is 0. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Nominal coverage	Oracle coverage	Power	Prop. completed
Benchmark model	-0.001	0	0.065	0.065	0.065	0.065	0.065	0.944	0.945	0.056	100
Complete-case	-0.11	-0.11	0.073	0.072	0.074	0.131	0.133	0.649	0.666	0.348	100
Confounded model	0.199	0.2	0.066	0.066	0.065	0.21	0.21	0.139	0.142	0.862	100
IPW	-0.005	-0.006	0.123	0.122	0.123	0.122	0.124	0.945	0.953	0.055	100
Raking (vanilla)	-0.001	0	0.071	0.072	0.071	0.072	0.071	0.951	0.948	0.048	100
MICE	0	0.001	0.069	0.07	0.068	0.07	0.068	0.950	0.947	0.050	100
MI-XGB	-0.002	-0.001	0.071	0.071	0.07	0.071	0.07	0.951	0.950	0.048	100
MI-RF	0.01	0.012	0.071	0.069	0.071	0.07	0.072	0.943	0.947	0.057	100
IPCW-TMLE-M	-0.021	-0.025	0.144	0.139	0.141	0.141	0.143	0.935	0.957	0.064	100
IPCW-TMLE-MTO	-0.027	-0.029	0.132	0.127	0.135	0.13	0.138	0.924	0.950	0.074	100

Table 68: **Synthetic data MAR simulation: census marginal risk difference (mRD), 12% outcome proportion, 40% missing proportion**. Comparing estimators under the **simple outcome (no treatment effect) and simple MAR** scenario. The value of the estimand is 0. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Oracle coverage	Nominal coverage	Power	Prop. completed
Benchmark model	-0.005	-0.004	0.065	0.065	0.065	0.065	0.065	0.944	0.946	0.056	100
Complete-case	-0.114	-0.114	0.073	0.072	0.074	0.134	0.136	0.648	0.629	0.348	100
Confounded model	0.195	0.196	0.066	0.066	0.065	0.206	0.206	0.154	0.152	0.862	100
IPW	-0.008	-0.01	0.123	0.122	0.123	0.122	0.124	0.952	0.943	0.055	100
Raking (vanilla)	-0.005	-0.004	0.071	0.072	0.071	0.072	0.071	0.947	0.952	0.048	100
MICE	-0.003	-0.003	0.069	0.07	0.068	0.07	0.068	0.946	0.952	0.050	100
MI-XGB	-0.006	-0.005	0.071	0.071	0.07	0.071	0.07	0.946	0.948	0.048	100
MI-RF	0.006	0.009	0.071	0.069	0.071	0.069	0.072	0.948	0.943	0.057	100
IPCW-TMLE-M	-0.025	-0.029	0.144	0.139	0.141	0.141	0.144	0.956	0.932	0.064	100
IPCW-TMLE-MTO	-0.031	-0.033	0.132	0.127	0.135	0.131	0.139	0.948	0.922	0.074	100

Table 69: **Synthetic data MAR simulation: oracle marginal risk difference (mRD), 12% outcome proportion, 40% missing proportion**. Comparing estimators under the **simple outcome and simple MAR (no dependence on Y)** scenario. The value of the estimand is 0.04. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Nominal coverage	Oracle coverage	Power	Prop. completed
Benchmark model	0.001	0	0.071	0.071	0.072	0.071	0.072	0.957	0.956	1.000	100
Complete-case	-0.016	-0.016	0.09	0.089	0.09	0.09	0.092	0.944	0.948	0.993	100
Confounded model	0.239	0.238	0.073	0.072	0.074	0.249	0.249	0.082	0.088	1.000	100
IPW	-0.003	-0.003	0.113	0.109	0.113	0.109	0.113	0.940	0.948	0.968	100
Raking (vanilla)	0.001	0	0.077	0.077	0.079	0.077	0.079	0.944	0.949	1.000	100
MICE	0	0.001	0.074	0.074	0.075	0.074	0.075	0.951	0.954	1.000	100
MI-XGB	-0.005	-0.004	0.076	0.077	0.078	0.077	0.078	0.953	0.952	1.000	100
MI-RF	0.009	0.009	0.075	0.075	0.076	0.075	0.076	0.950	0.951	1.000	100
IPCW-TMLE-M	-0.002	-0.011	0.141	0.132	0.133	0.132	0.134	0.931	0.949	0.892	100
IPCW-TMLE-MTO	-0.002	-0.007	0.13	0.118	0.123	0.118	0.123	0.922	0.949	0.935	100

Table 70: **Synthetic data MAR simulation: census marginal risk difference (mRD), 12% outcome proportion, 40% missing proportion**. Comparing estimators under the **simple outcome and simple MAR (no dependence on Y)** scenario. The value of the estimand is 0.041. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Oracle coverage	Nominal coverage	Power	Prop. completed
Benchmark model	-0.002	-0.003	0.071	0.071	0.072	0.071	0.072	0.957	0.957	1.000	100
Complete-case	-0.019	-0.02	0.09	0.089	0.09	0.091	0.092	0.946	0.941	0.993	100
Confounded model	0.236	0.235	0.073	0.072	0.074	0.246	0.246	0.097	0.093	1.000	100
IPW	-0.006	-0.006	0.113	0.109	0.113	0.109	0.114	0.949	0.938	0.968	100
Raking (vanilla)	-0.003	-0.003	0.077	0.077	0.079	0.077	0.079	0.947	0.944	1.000	100
MICE	-0.003	-0.002	0.074	0.074	0.075	0.074	0.076	0.954	0.955	1.000	100
MI-XGB	-0.009	-0.008	0.076	0.077	0.078	0.077	0.078	0.950	0.951	1.000	100
MI-RF	0.005	0.006	0.075	0.075	0.076	0.075	0.076	0.954	0.951	1.000	100
IPCW-TMLE-M	-0.005	-0.014	0.141	0.132	0.133	0.132	0.134	0.949	0.929	0.892	100
IPCW-TMLE-MTO	-0.005	-0.01	0.13	0.118	0.123	0.118	0.123	0.949	0.919	0.935	100

Table 71: **Synthetic data MAR simulation: oracle marginal risk difference (mRD), 12% outcome proportion, 40% missing proportion**. Comparing estimators under the **simple outcome and simple MAR (no dependence on Y)** scenario. The value of the estimand is 0.04. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Nominal coverage	Oracle coverage	Power	Prop. completed
Benchmark model	0	0	0.072	0.071	0.072	0.071	0.072	0.948	0.954	1.000	100
Complete-case	-0.033	-0.037	0.154	0.152	0.154	0.155	0.158	0.935	0.942	0.685	100
Confounded model	0.237	0.238	0.074	0.072	0.073	0.247	0.249	0.099	0.109	1.000	100
IPW	-0.002	-0.013	0.286	0.251	0.258	0.251	0.258	0.936	0.966	0.342	100
Raking (vanilla)	-0.003	-0.004	0.112	0.11	0.111	0.11	0.112	0.941	0.952	0.948	100
MICE	-0.002	-0.002	0.089	0.087	0.091	0.087	0.091	0.948	0.954	0.994	100
MI-RF	0.03	0.031	0.096	0.084	0.095	0.089	0.1	0.906	0.941	0.997	100
IPCW-TMLE-M	-0.008	-0.026	0.311	0.28	0.305	0.28	0.306	0.922	0.951	0.261	100
IPCW-TMLE-MTO	-0.003	-0.011	0.272	0.235	0.274	0.235	0.274	0.906	0.952	0.396	100

Table 72: **Synthetic data MAR simulation: census marginal risk difference (mRD), 12% outcome proportion, 40% missing proportion**. Comparing estimators under the **simple outcome and simple MAR (no dependence on Y)** scenario. The value of the estimand is 0.041. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Oracle coverage	Nominal coverage	Power	Prop. completed
Benchmark model	-0.003	-0.003	0.072	0.071	0.072	0.071	0.072	0.952	0.947	1.000	100
Complete-case	-0.037	-0.041	0.154	0.152	0.154	0.156	0.159	0.939	0.933	0.685	100
Confounded model	0.233	0.234	0.074	0.072	0.073	0.244	0.246	0.116	0.107	1.000	100
IPW	-0.005	-0.017	0.286	0.251	0.258	0.251	0.259	0.967	0.936	0.342	100
Raking (vanilla)	-0.007	-0.007	0.112	0.11	0.111	0.11	0.112	0.952	0.941	0.948	100
MICE	-0.005	-0.006	0.089	0.087	0.091	0.087	0.092	0.954	0.944	0.994	100
MI-RF	0.026	0.028	0.096	0.084	0.095	0.088	0.099	0.941	0.908	0.997	100
IPCW-TMLE-M	-0.011	-0.029	0.311	0.28	0.305	0.28	0.307	0.951	0.918	0.261	100
IPCW-TMLE-MTO	-0.007	-0.014	0.272	0.235	0.274	0.235	0.274	0.952	0.904	0.396	100

Table 73: **Synthetic data MAR simulation: oracle marginal risk difference (mRD), 12% outcome proportion, 40% missing proportion**. Comparing estimators under the **semi-complex outcome and simple MAR** scenario. The value of the estimand is 0. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE. Estimators that are mismatched with the estimand (i.e., are estimating a different parameter) are emphasized using a star. The semi-complex outcome is a function of exponentiated and squared terms on Z.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Nominal coverage	Oracle coverage	Power	Prop. completed
Benchmark model	-0.001	-0.001	0.037	0.036	0.037	0.036	0.037	0.945	0.952	0.055	100
Complete-case*	0.038	0.038	0.054	0.053	0.052	0.065	0.065	0.892	0.894	0.108	100
Confounded model*	0.128	0.127	0.049	0.049	0.048	0.137	0.135	0.253	0.252	0.748	100
IPW*	0.069	0.068	0.105	0.103	0.105	0.124	0.125	0.901	0.897	0.100	100
Raking (vanilla)*	0.07	0.07	0.056	0.056	0.056	0.09	0.09	0.760	0.758	0.241	100
MICE*	0.076	0.076	0.054	0.055	0.055	0.094	0.094	0.724	0.712	0.276	100
MI-XGB*	0.093	0.093	0.056	0.055	0.055	0.108	0.108	0.608	0.621	0.392	100
MI-RF*	0.099	0.098	0.055	0.053	0.054	0.112	0.112	0.553	0.569	0.447	100
IPCW-TMLE-M	-0.055	-0.065	0.17	0.152	0.166	0.161	0.178	0.845	0.942	0.155	100
IPCW-TMLE-MTO	-0.022	-0.025	0.094	0.083	0.088	0.086	0.091	0.892	0.940	0.108	100

Table 74: **Synthetic data MAR simulation: census marginal risk difference (mRD), 12% outcome proportion, 40% missing proportion**. Comparing estimators under the **semi-complex outcome and simple MAR** scenario. The value of the estimand is 0.008. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE. Estimators that are mismatched with the estimand (i.e., are estimating a different parameter) are emphasized using a star. The semi-complex outcome is a function of exponentiated and squared terms on Z.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Oracle coverage	Nominal coverage	Power	Prop. completed
Benchmark model	-0.004	-0.004	0.05	0.05	0.048	0.05	0.049	0.943	0.946	0.286	100
Complete-case	-0.037	-0.037	0.054	0.053	0.052	0.064	0.064	0.897	0.883	0.108	100
Confounded model	0.053	0.052	0.049	0.049	0.048	0.072	0.07	0.812	0.815	0.748	100
IPW	-0.006	-0.007	0.105	0.103	0.105	0.103	0.105	0.954	0.939	0.100	100
Raking (vanilla)	-0.005	-0.005	0.056	0.056	0.056	0.056	0.057	0.952	0.950	0.241	100
MICE	0.001	0.001	0.054	0.055	0.055	0.055	0.055	0.948	0.949	0.276	100
MI-XGB	0.018	0.018	0.056	0.055	0.055	0.058	0.058	0.933	0.931	0.392	100
MI-RF	0.024	0.023	0.055	0.053	0.054	0.059	0.059	0.924	0.920	0.447	100
IPCW-TMLE-M*	-0.13	-0.14	0.17	0.152	0.166	0.2	0.217	0.885	0.741	0.155	100
IPCW-TMLE-MTO*	-0.097	-0.1	0.094	0.083	0.088	0.128	0.133	0.826	0.712	0.108	100

Table 75: **Synthetic data MAR simulation: oracle marginal risk difference (mRD), 12% outcome proportion, 40% missing proportion**. Comparing estimators under the **semi-complex outcome and simple MAR** scenario. The value of the estimand is -0.016. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE. Estimators that are mismatched with the estimand (i.e., are estimating a different parameter) are emphasized using a star. The semi-complex outcome is a function of exponentiated and squared terms on Z.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Nominal coverage	Oracle coverage	Power	Prop. completed
Benchmark model	-0.001	-0.001	0.034	0.034	0.033	0.034	0.033	0.942	0.945	0.997	100
Complete-case*	0.117	0.117	0.05	0.05	0.05	0.127	0.127	0.351	0.352	0.144	100
Confounded model*	0.132	0.132	0.048	0.047	0.048	0.14	0.14	0.194	0.207	0.094	100
IPW*	0.08	0.078	0.098	0.097	0.097	0.126	0.125	0.890	0.877	0.158	100
Raking (vanilla)*	0.082	0.082	0.055	0.053	0.056	0.098	0.099	0.660	0.679	0.323	100
MICE*	0.088	0.087	0.053	0.052	0.053	0.103	0.102	0.612	0.624	0.292	100
MI-XGB*	0.104	0.104	0.054	0.053	0.055	0.117	0.118	0.490	0.512	0.191	100
MI-RF*	0.11	0.109	0.053	0.051	0.053	0.121	0.121	0.428	0.463	0.183	100
IPCW-TMLE-M	-0.048	-0.056	0.16	0.142	0.155	0.15	0.165	0.850	0.944	0.435	100
IPCW-TMLE-MTO	-0.017	-0.016	0.093	0.08	0.092	0.082	0.094	0.880	0.948	0.596	100

Table 76: **Synthetic data MAR simulation: census marginal risk difference (mRD), 12% outcome proportion, 40% missing proportion**. Comparing estimators under the **semi-complex outcome and simple MAR** scenario. The value of the estimand is -0.008. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE. Estimators that are mismatched with the estimand (i.e., are estimating a different parameter) are emphasized using a star. The semi-complex outcome is a function of exponentiated and squared terms on Z.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Oracle coverage	Nominal coverage	Power	Prop. completed
Benchmark model	-0.002	-0.003	0.049	0.048	0.049	0.048	0.049	0.949	0.944	0.388	100
Complete-case	0.033	0.032	0.05	0.05	0.05	0.06	0.06	0.896	0.898	0.144	100
Confounded model	0.048	0.048	0.048	0.047	0.048	0.067	0.068	0.830	0.820	0.094	100
IPW	-0.004	-0.006	0.098	0.097	0.097	0.098	0.098	0.947	0.943	0.158	100
Raking (vanilla)	-0.002	-0.002	0.055	0.053	0.056	0.053	0.056	0.949	0.940	0.323	100
MICE	0.004	0.003	0.053	0.052	0.053	0.052	0.053	0.950	0.947	0.292	100
MI-XGB	0.02	0.02	0.054	0.053	0.055	0.056	0.058	0.934	0.922	0.191	100
MI-RF	0.026	0.025	0.053	0.051	0.053	0.057	0.059	0.920	0.906	0.183	100
IPCW-TMLE-M*	-0.132	-0.14	0.16	0.142	0.155	0.194	0.209	0.876	0.718	0.435	100
IPCW-TMLE-MTO*	-0.101	-0.1	0.093	0.08	0.092	0.129	0.136	0.804	0.681	0.596	100

Table 77: **Synthetic data MAR simulation: oracle marginal risk difference (mRD), 12% outcome proportion, 40% missing proportion**. Comparing estimators under the **semi-complex outcome and simple MAR** scenario. The value of the estimand is 0. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE. Estimators that are mismatched with the estimand (i.e., are estimating a different parameter) are emphasized using a star. The semi-complex outcome is a function of exponentiated and squared terms on W.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Nominal coverage	Oracle coverage	Power	Prop. completed
Benchmark model	0.001	0.001	0.036	0.036	0.037	0.036	0.037	0.950	0.949	0.050	100
Complete-case*	0.017	0.017	0.06	0.059	0.06	0.062	0.063	0.936	0.938	0.064	100
Confounded model*	-0.236	-0.236	0.066	0.065	0.066	0.244	0.245	0.055	0.055	0.946	100
IPW*	0.072	0.071	0.091	0.092	0.089	0.117	0.114	0.896	0.875	0.104	100
Raking (vanilla)*	0.077	0.075	0.073	0.07	0.073	0.104	0.105	0.809	0.822	0.189	100
MICE*	0.057	0.056	0.058	0.06	0.058	0.083	0.081	0.863	0.831	0.136	100
MI-XGB*	0.068	0.069	0.085	0.077	0.082	0.103	0.107	0.828	0.879	0.170	100
MI-RF*	0.126	0.126	0.067	0.064	0.068	0.141	0.143	0.501	0.536	0.497	100
IPCW-TMLE-M	-0.043	-0.048	0.149	0.135	0.137	0.142	0.146	0.881	0.948	0.120	100
IPCW-TMLE-MTO	-0.016	-0.018	0.083	0.076	0.083	0.078	0.085	0.916	0.947	0.084	100

Table 78: **Synthetic data MAR simulation: census marginal risk difference (mRD), 12% outcome proportion, 40% missing proportion**. Comparing estimators under the **semi-complex outcome and simple MAR** scenario. The value of the estimand is 0.008. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE. Estimators that are mismatched with the estimand (i.e., are estimating a different parameter) are emphasized using a star. The semi-complex outcome is a function of exponentiated and squared terms on W.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Oracle coverage	Nominal coverage	Power	Prop. completed
Benchmark model	0	0	0.052	0.05	0.05	0.05	0.05	0.950	0.940	0.321	100
Complete-case	-0.058	-0.058	0.06	0.059	0.06	0.083	0.084	0.837	0.825	0.064	100
Confounded model	-0.311	-0.312	0.066	0.065	0.066	0.318	0.319	0.005	0.005	0.946	100
IPW	-0.004	-0.004	0.091	0.092	0.089	0.092	0.089	0.950	0.950	0.104	100
Raking (vanilla)	0.001	0	0.073	0.07	0.073	0.07	0.073	0.948	0.942	0.189	100
MICE	-0.019	-0.019	0.058	0.06	0.058	0.063	0.061	0.940	0.944	0.136	100
MI-XGB	-0.008	-0.007	0.085	0.077	0.082	0.077	0.083	0.953	0.914	0.170	100
MI-RF	0.05	0.05	0.067	0.064	0.068	0.081	0.085	0.879	0.862	0.497	100
IPCW-TMLE-M*	-0.119	-0.124	0.149	0.135	0.137	0.18	0.185	0.875	0.746	0.120	100
IPCW-TMLE-MTO*	-0.092	-0.094	0.083	0.076	0.083	0.119	0.125	0.801	0.713	0.084	100

Table 79: **Synthetic data MAR simulation: oracle marginal risk difference (mRD), 12% outcome proportion, 40% missing proportion**. Comparing estimators under the **semi-complex outcome and simple MAR** scenario. The value of the estimand is -0.016. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE. Estimators that are mismatched with the estimand (i.e., are estimating a different parameter) are emphasized using a star. The semi-complex outcome is a function of exponentiated and squared terms on W.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Nominal coverage	Oracle coverage	Power	Prop. completed
Benchmark model	0	0	0.034	0.034	0.035	0.034	0.035	0.950	0.951	0.998	100
Complete-case*	0.076	0.076	0.056	0.056	0.056	0.095	0.095	0.727	0.721	0.328	100
Confounded model*	-0.225	-0.225	0.063	0.062	0.062	0.234	0.234	0.049	0.049	1.000	100
IPW*	0.084	0.08	0.088	0.088	0.088	0.121	0.119	0.852	0.837	0.162	100
Raking (vanilla)*	0.086	0.085	0.067	0.067	0.068	0.109	0.109	0.753	0.745	0.212	100
MICE*	0.096	0.096	0.051	0.057	0.05	0.112	0.108	0.626	0.540	0.184	100
MI-XGB*	0.093	0.091	0.076	0.074	0.077	0.119	0.119	0.744	0.770	0.173	100
MI-RF*	0.135	0.135	0.06	0.06	0.062	0.148	0.149	0.389	0.390	0.068	100
IPCW-TMLE-M	-0.035	-0.046	0.141	0.126	0.135	0.131	0.143	0.890	0.954	0.455	100
IPCW-TMLE-MTO	-0.008	-0.011	0.083	0.073	0.083	0.074	0.084	0.906	0.951	0.626	100

Table 80: **Synthetic data MAR simulation: census marginal risk difference (mRD), 12% outcome proportion, 40% missing proportion**. Comparing estimators under the **semi-complex outcome and simple MAR** scenario. The value of the estimand is -0.008. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE. Estimators that are mismatched with the estimand (i.e., are estimating a different parameter) are emphasized using a star. The semi-complex outcome is a function of exponentiated and squared terms on W.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Oracle coverage	Nominal coverage	Power	Prop. completed
Benchmark model	0.002	0.002	0.048	0.048	0.047	0.048	0.047	0.953	0.952	0.337	100
Complete-case	-0.007	-0.007	0.056	0.056	0.056	0.056	0.057	0.950	0.950	0.328	100
Confounded model	-0.309	-0.309	0.063	0.062	0.062	0.315	0.315	0.003	0.003	1.000	100
IPW	0.001	-0.003	0.088	0.088	0.088	0.088	0.088	0.952	0.950	0.162	100
Raking (vanilla)	0.003	0.002	0.067	0.067	0.068	0.067	0.068	0.953	0.949	0.212	100
MICE	0.013	0.012	0.051	0.057	0.05	0.058	0.052	0.941	0.964	0.184	100
MI-XGB	0.01	0.008	0.076	0.074	0.077	0.074	0.077	0.949	0.933	0.173	100
MI-RF	0.052	0.052	0.06	0.06	0.062	0.08	0.081	0.855	0.864	0.068	100
IPCW-TMLE-M*	-0.118	-0.129	0.141	0.126	0.135	0.173	0.187	0.878	0.724	0.455	100
IPCW-TMLE-MTO*	-0.091	-0.094	0.083	0.073	0.083	0.117	0.126	0.806	0.688	0.626	100

Table 81: **Synthetic data MAR simulation: oracle marginal risk difference (mRD), 12% outcome proportion, 40% missing proportion**. Comparing estimators under the **semi-complex outcome and simple MAR (no dependence on Y)** scenario. The value of the estimand is -0.016. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE. Estimators that are mismatched with the estimand (i.e., are estimating a different parameter) are emphasized using a star. The semi-complex outcome is a function of exponentiated and squared terms on W.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Nominal coverage	Oracle coverage	Power	Prop. completed
Benchmark model	-0.002	-0.002	0.034	0.034	0.034	0.034	0.034	0.950	0.952	0.997	100
Complete-case*	0.086	0.084	0.096	0.095	0.092	0.128	0.125	0.847	0.848	0.121	100
Confounded model*	-0.228	-0.229	0.061	0.062	0.063	0.237	0.237	0.046	0.041	1.000	100
IPW*	0.086	0.08	0.179	0.163	0.173	0.184	0.191	0.915	0.922	0.117	100
Raking (vanilla)*	0.091	0.087	0.136	0.122	0.133	0.152	0.159	0.856	0.898	0.131	100
MICE*	-0.006	-0.006	0.057	0.066	0.056	0.066	0.056	0.975	0.945	0.739	100
MI-RF*	0.144	0.144	0.088	0.074	0.087	0.162	0.168	0.509	0.617	0.103	100
IPCW-TMLE-M	-0.031	-0.055	0.265	0.216	0.236	0.218	0.242	0.887	0.960	0.280	100
IPCW-TMLE-MTO	0.004	0.006	0.16	0.134	0.15	0.134	0.15	0.905	0.952	0.281	100

Table 82: **Synthetic data MAR simulation: census marginal risk difference (mRD), 12% outcome proportion, 40% missing proportion**. Comparing estimators under the **semi-complex outcome and simple MAR (no dependence on Y)** scenario. The value of the estimand is -0.008. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE. Estimators that are mismatched with the estimand (i.e., are estimating a different parameter) are emphasized using a star. The semi-complex outcome is a function of exponentiated and squared terms on W.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Oracle coverage	Nominal coverage	Power	Prop. completed
Benchmark model	0.001	0	0.047	0.048	0.049	0.048	0.049	0.952	0.952	0.354	100
Complete-case	0.002	0.001	0.096	0.095	0.092	0.095	0.092	0.947	0.947	0.121	100
Confounded model	-0.312	-0.312	0.061	0.062	0.063	0.318	0.318	0.001	0.001	1.000	100
IPW	0.003	-0.004	0.179	0.163	0.173	0.163	0.173	0.948	0.928	0.117	100
Raking (vanilla)	0.008	0.003	0.136	0.122	0.133	0.122	0.134	0.944	0.922	0.131	100
MICE	-0.09	-0.089	0.057	0.066	0.056	0.111	0.105	0.654	0.746	0.739	100
MI-RF	0.061	0.061	0.088	0.074	0.087	0.096	0.106	0.900	0.830	0.103	100
IPCW-TMLE-M*	-0.114	-0.139	0.265	0.216	0.236	0.244	0.274	0.950	0.810	0.280	100
IPCW-TMLE-MTO*	-0.079	-0.077	0.16	0.134	0.15	0.156	0.169	0.923	0.842	0.281	100

Table 83: **Synthetic data MAR simulation: oracle marginal risk difference (mRD), 12% outcome proportion, 80% missing proportion**. Comparing estimators under the **semi-complex outcome and simple MAR** scenario. The value of the estimand is -0.016. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE. Estimators that are mismatched with the estimand (i.e., are estimating a different parameter) are emphasized using a star. The semi-complex outcome is a function of exponentiated and squared terms on W.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Nominal coverage	Oracle coverage	Power	Prop. completed
Benchmark model	-0.001	0	0.035	0.034	0.035	0.034	0.035	0.947	0.953	0.996	100
Complete-case*	0.13	0.129	0.089	0.086	0.085	0.156	0.155	0.690	0.708	0.078	100
Confounded model*	-0.228	-0.228	0.061	0.062	0.059	0.236	0.235	0.045	0.042	1.000	100
IPW*	0.074	0.069	0.187	0.172	0.185	0.188	0.197	0.919	0.933	0.131	100
Raking (vanilla)*	0.088	0.085	0.126	0.115	0.122	0.145	0.149	0.876	0.896	0.140	100
MICE*	0.104	0.1	0.078	0.077	0.077	0.13	0.126	0.745	0.750	0.138	100
MI-RF*	0.173	0.17	0.094	0.073	0.096	0.188	0.195	0.380	0.558	0.128	100
IPCW-TMLE-M	-0.068	-0.103	0.287	0.218	0.247	0.229	0.268	0.833	0.958	0.358	100
IPCW-TMLE-MTO	-0.031	-0.029	0.171	0.14	0.166	0.143	0.169	0.874	0.944	0.369	100

Table 84: **Synthetic data MAR simulation: census marginal risk difference (mRD), 12% outcome proportion, 80% missing proportion**. Comparing estimators under the **semi-complex outcome and simple MAR** scenario. The value of the estimand is -0.008. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE. Estimators that are mismatched with the estimand (i.e., are estimating a different parameter) are emphasized using a star. The semi-complex outcome is a function of exponentiated and squared terms on W.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Oracle coverage	Nominal coverage	Power	Prop. completed
Benchmark model	0.001	0.001	0.048	0.048	0.048	0.048	0.048	0.954	0.954	0.357	100
Complete-case	0.047	0.046	0.089	0.086	0.085	0.098	0.097	0.915	0.917	0.078	100
Confounded model	-0.311	-0.311	0.061	0.062	0.059	0.317	0.316	0.002	0.002	1.000	100
IPW	-0.01	-0.014	0.187	0.172	0.185	0.173	0.185	0.948	0.930	0.131	100
Raking (vanilla)	0.005	0.002	0.126	0.115	0.122	0.115	0.122	0.951	0.924	0.140	100
MICE	0.021	0.017	0.078	0.077	0.077	0.08	0.079	0.939	0.945	0.138	100
MI-RF	0.09	0.087	0.094	0.073	0.096	0.116	0.129	0.846	0.726	0.128	100
IPCW-TMLE-M*	-0.151	-0.186	0.287	0.218	0.247	0.266	0.309	0.942	0.747	0.358	100
IPCW-TMLE-MTO*	-0.114	-0.112	0.171	0.14	0.166	0.18	0.201	0.892	0.772	0.369	100

Table 85: **Synthetic data MAR simulation: oracle marginal risk difference (mRD), 12% outcome proportion, 40% missing proportion**. Comparing estimators under the **complex outcome and simple MAR** scenario. The value of the estimand is 0. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE. Estimators that are mismatched with the estimand (i.e., are estimating a different parameter) are emphasized using a star.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Nominal coverage	Oracle coverage	Power	Prop. completed
Benchmark model	-0.001	-0.002	0.058	0.059	0.058	0.059	0.058	0.953	0.950	0.046	100
Complete-case*	-0.081	-0.083	0.074	0.077	0.075	0.112	0.111	0.811	0.802	0.183	100
Confounded model*	-0.199	-0.201	0.063	0.064	0.063	0.21	0.21	0.132	0.115	0.866	100
IPW*	0.053	0.054	0.12	0.124	0.12	0.135	0.131	0.947	0.930	0.052	100
Raking (vanilla)*	0.053	0.053	0.074	0.075	0.071	0.092	0.088	0.900	0.890	0.102	100
MICE*	0.114	0.115	0.075	0.077	0.075	0.138	0.137	0.700	0.673	0.305	100
MI-XGB*	0.085	0.084	0.072	0.074	0.069	0.113	0.108	0.806	0.786	0.197	100
MI-RF*	0.064	0.064	0.075	0.073	0.072	0.097	0.096	0.851	0.856	0.151	100
IPCW-TMLE-M	-0.039	-0.049	0.165	0.157	0.161	0.162	0.168	0.892	0.954	0.106	100
IPCW-TMLE-MTO	-0.039	-0.04	0.129	0.124	0.127	0.13	0.134	0.907	0.945	0.092	100

Table 86: **Synthetic data MAR simulation: census marginal risk difference (mRD), 12% outcome proportion, 40% missing proportion**. Comparing estimators under the **complex outcome and simple MAR** scenario. The value of the estimand is 0.006. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE. Estimators that are mismatched with the estimand (i.e., are estimating a different parameter) are emphasized using a star.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Oracle coverage	Nominal coverage	Power	Prop. completed
Benchmark model	-0.002	-0.004	0.065	0.067	0.065	0.067	0.065	0.953	0.956	0.120	100
Complete-case	-0.136	-0.138	0.074	0.077	0.075	0.156	0.157	0.550	0.568	0.183	100
Confounded model	-0.254	-0.256	0.063	0.064	0.063	0.262	0.263	0.019	0.023	0.866	100
IPW	-0.001	-0.001	0.12	0.124	0.12	0.124	0.12	0.951	0.954	0.052	100
Raking (vanilla)	-0.002	-0.002	0.074	0.075	0.071	0.075	0.071	0.951	0.954	0.102	100
MICE	0.06	0.06	0.075	0.077	0.075	0.097	0.096	0.878	0.889	0.305	100
MI-XGB	0.03	0.029	0.072	0.074	0.069	0.08	0.075	0.926	0.936	0.197	100
MI-RF	0.009	0.009	0.075	0.073	0.072	0.073	0.072	0.946	0.939	0.151	100
IPCW-TMLE-M*	-0.094	-0.104	0.165	0.157	0.161	0.183	0.191	0.921	0.824	0.106	100
IPCW-TMLE-MTO*	-0.094	-0.095	0.129	0.124	0.127	0.155	0.159	0.890	0.828	0.092	100

Table 87: **Synthetic data MAR simulation: oracle marginal risk difference (mRD), 12% outcome proportion, 80% missing proportion**. Comparing estimators under the **complex outcome and complex MAR (no dependence on Y)** scenario. The value of the estimand is 0.031. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE. Estimators that are mismatched with the estimand (i.e., are estimating a different parameter) are emphasized using a star.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Nominal coverage	Oracle coverage	Power	Prop. completed
Benchmark model	-0.003	-0.003	0.065	0.064	0.064	0.064	0.064	0.946	0.949	0.999	100.00
Complete-case*	0.021	0.014	0.159	0.158	0.163	0.16	0.164	0.948	0.951	0.554	100.00
Confounded model*	-0.225	-0.228	0.07	0.069	0.069	0.236	0.238	0.104	0.105	0.246	100.00
IPW*	0.231	-0.279	2.06	abs > ln(10)	0.218	abs > ln(10)	0.354	0.404	0.932	0.025	100.00
Raking (vanilla)*	0.562	0.256	1.225	0.726	0.715	0.918	0.759	0.922	0.931	0.178	94.24
MICE*	-0.011	-0.011	0.122	0.103	0.122	0.103	0.123	0.882	0.949	0.821	100.00
MI-RF*	-0.073	-0.071	0.1	0.092	0.102	0.118	0.124	0.835	0.889	0.732	100.00
IPCW-TMLE-M	-0.078	-0.152	0.533	0.43	0.47	0.437	0.494	0.830	0.950	0.079	100.00

Table 88: **Synthetic data MAR simulation: census marginal risk difference (mRD), 12% outcome proportion, 80% missing proportion**. Comparing estimators under the **complex outcome and complex MAR (no dependence on Y)** scenario. The value of the estimand is 0.037. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE. Estimators that are mismatched with the estimand (i.e., are estimating a different parameter) are emphasized using a star.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Oracle coverage	Nominal coverage	Power	Prop. completed
Benchmark model	0	-0.002	0.074	0.072	0.073	0.072	0.073	0.952	0.944	0.999	100.00
Complete-case	-0.032	-0.038	0.159	0.158	0.163	0.161	0.168	0.944	0.940	0.554	100.00
Confounded model	-0.278	-0.28	0.07	0.069	0.069	0.286	0.289	0.022	0.022	0.246	100.00
IPW	0.179	-0.331	2.06	abs > ln(10)	0.218	abs > ln(10)	0.397	0.931	0.359	0.025	100.00
Raking (vanilla)	0.509	0.204	1.225	0.726	0.715	0.887	0.743	0.935	0.922	0.178	94.24
MICE	-0.064	-0.064	0.122	0.103	0.122	0.121	0.138	0.913	0.826	0.821	100.00
MI-RF	-0.126	-0.124	0.1	0.092	0.102	0.156	0.16	0.753	0.692	0.732	100.00
IPCW-TMLE-M*	-0.13	-0.204	0.533	0.43	0.47	0.449	0.513	0.952	0.808	0.079	100.00

Table 89: **Synthetic data MAR simulation: oracle marginal risk difference (mRD), 12% outcome proportion, 40% missing proportion**. Comparing estimators under the **complex outcome and simple MAR (no dependence on Y)** scenario. The value of the estimand is 0.031. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE. Estimators that are mismatched with the estimand (i.e., are estimating a different parameter) are emphasized using a star.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Nominal coverage	Oracle coverage	Power	Prop. completed
Benchmark model	0.001	0	0.064	0.064	0.064	0.064	0.064	0.952	0.951	0.999	100
Complete-case*	0.015	0.014	0.163	0.161	0.163	0.161	0.164	0.946	0.951	0.533	100
Confounded model*	-0.223	-0.222	0.068	0.069	0.067	0.233	0.232	0.103	0.096	0.254	100
IPW*	0.046	0.031	0.343	0.237	0.261	0.241	0.262	0.940	0.980	0.298	100
Raking (vanilla)*	0.051	0.05	0.125	0.118	0.122	0.129	0.132	0.918	0.928	0.872	100
MICE*	0.162	0.16	0.103	0.101	0.106	0.191	0.192	0.650	0.656	0.997	100
MI-RF*	0.064	0.063	0.1	0.089	0.1	0.11	0.118	0.855	0.896	0.981	100
IPCW-TMLE-M	-0.031	-0.076	0.345	0.292	0.305	0.294	0.314	0.899	0.956	0.123	100
IPCW-TMLE-MTO	-0.034	-0.051	0.264	0.222	0.252	0.224	0.257	0.889	0.950	0.248	100

Table 90: **Synthetic data MAR simulation: census marginal risk difference (mRD), 12% outcome proportion, 40% missing proportion**. Comparing estimators under the **complex outcome and simple MAR (no dependence on Y)** scenario. The value of the estimand is 0.037. The sample size is n = 10000. Maximum observed Monte-Carlo error over the 2500 simulation replications was 0.010 for all summaries besides coverage and 0.012 for coverage. ESE = empirical standard error, ASE = asymptotic standard error, MAD = mean absolute deviation, RMSE = root mean squared error, rRMSE = robust RMSE (using median bias and MAD), Oracle coverage = coverage of a confidence interval based on the ESE, Nominal coverage = coverage of a confidence interval based on the ASE. Estimators that are mismatched with the estimand (i.e., are estimating a different parameter) are emphasized using a star.

Estimator	Mean bias	Median bias	ESE	ASE	MAD	RMSE	rRMSE	Oracle coverage	Nominal coverage	Power	Prop. completed
Benchmark model	0.003	0.003	0.072	0.072	0.071	0.073	0.071	0.948	0.949	0.999	100
Complete-case	-0.038	-0.039	0.163	0.161	0.163	0.165	0.168	0.946	0.940	0.533	100
Confounded model	-0.275	-0.275	0.068	0.069	0.067	0.284	0.283	0.020	0.022	0.254	100
IPW	-0.007	-0.022	0.343	0.237	0.261	0.237	0.261	0.983	0.924	0.298	100
Raking (vanilla)	-0.001	-0.003	0.125	0.118	0.122	0.118	0.122	0.945	0.933	0.872	100
MICE	0.109	0.107	0.103	0.101	0.106	0.149	0.151	0.811	0.807	0.997	100
MI-RF	0.012	0.011	0.1	0.089	0.1	0.089	0.1	0.949	0.912	0.981	100
IPCW-TMLE-M*	-0.084	-0.129	0.345	0.292	0.305	0.304	0.331	0.952	0.870	0.123	100
IPCW-TMLE-MTO*	-0.086	-0.104	0.264	0.222	0.252	0.238	0.273	0.945	0.853	0.248	100