Statin Simulation Report

Haodong Li

2022-10-03

Simulation

DGD:

$$
\begin{aligned}
W_1 &\sim Unif(-1,1) \\
W_2 &\sim Unif(-1,1) \\
A &\sim Bernoulli(p) \text{ where } p = \text{expit}(0.1 * W_1 * W_2 - 0.4 * W_1) \\
\tau &= W_1^2 * (W_1 + 7/5) + (5 * W_2/3)^2 \\
\mu_Y &= A * \tau + W_1 * W_2 + 2 * W_2^2 - W_1 \\
Y &\sim N(\mu_Y, 1)
\end{aligned}
$$

Initial estimating models for Q and g:

1) GAM: General Additive Models (Correctly specified based on true DGD of Q and g)
2) Earth: Multivariate Adaptive Regression Splines

Note this is not SL, just individual algorithms. For estimations of other terms (τ, τ_s, γ_s), all use Earth.

CATE estimation

1) DR-learner: regress pseudo outcome estimates φ_n^0 on W.
2) T-learner: just use $\bar{Q}_n^0(1, W)$ - $\bar{Q}_n^0(0, W)$

Truncation:

For TMLE, $\bar{Q}_n^{(0)} \in [0.001, 0.999]$, $g_n \in [0.025, 0.975]$.

For EE, $g_n \in [0.025, 0.975]$

Table 1: Configuration of simulations

Table_ID	Model_Q_g	CATE_learner	CV
Table 2	GAM	DR-learner	No
Table 3	GAM	T-learner	No
Table 4	Earth	DR-learner	No
Table 5	Earth	T-learner	No
Table 6	GAM	DR-learner	Yes
Table 7	GAM	T-learner	Yes
Table 8	Earth	DR-learner	Yes
Table 9	Earth	T-learner	Yes

Table 2: Performance of TMLE and EE for Theta (Correct Q and g, DR-learner, no CV)

n	Method	True_Theta	Variance	Bias	MSE	Coverage	Coverage_or	CI_width
500	TMLE	0.686	0.0265	0.0302	0.0275	0.938	0.948	0.6386
	EE	0.686	0.0414	0.0758	0.0472	0.914	0.928	0.7010
1000	TMLE	0.686	0.0124	0.0100	0.0125	0.950	0.954	0.4369
	EE	0.686	0.0165	0.0430	0.0183	0.930	0.942	0.4652
2000	TMLE	0.686	0.0056	0.0070	0.0057	0.958	0.956	0.3062
	EE	0.686	0.0061	0.0272	0.0068	0.942	0.928	0.3145
3000	TMLE	0.686	0.0040	0.0112	0.0041	0.946	0.940	0.2504
	EE	0.686	0.0041	0.0215	0.0046	0.942	0.938	0.2538
4000	TMLE	0.686	0.0029	0.0054	0.0030	0.952	0.952	0.2152
	EE	0.686	0.0031	0.0106	0.0032	0.954	0.952	0.2169
5000	TMLE	0.686	0.0023	0.0038	0.0023	0.956	0.952	0.1925
	EE	0.686	0.0023	0.0046	0.0023	0.958	0.946	0.1931
10000	TMLE	0.686	0.0013	0.0060	0.0013	0.934	0.938	0.1364
	EE	0.686	0.0013	0.0000	0.0013	0.934	0.946	0.1360
20000	TMLE	0.686	0.0006	0.0079	0.0007	0.938	0.940	0.0966
	EE	0.686	0.0006	-0.0017	0.0006	0.948	0.944	0.0960

Table 3: Performance of TMLE and EE for Theta (Correct Q and g, T-learner, no CV)

n	Method	True_Theta	Variance	Bias	MSE	Coverage	Coverage_or	CI_width
500	TMLE	0.686	0.0234	0.0276	0.0242	0.954	0.954	0.6279
	EE	0.686	0.0238	0.0442	0.0257	0.944	0.946	0.6111
1000	TMLE	0.686	0.0113	0.0179	0.0116	0.954	0.936	0.4373
	EE	0.686	0.0112	0.0285	0.0120	0.950	0.934	0.4287
2000	TMLE	0.686	0.0061	0.0121	0.0062	0.952	0.932	0.3062
	EE	0.686	0.0061	0.0161	0.0064	0.940	0.932	0.3024
3000	TMLE	0.686	0.0039	0.0082	0.0040	0.944	0.948	0.2488
	EE	0.686	0.0039	0.0106	0.0040	0.944	0.952	0.2466
4000	TMLE	0.686	0.0026	0.0057	0.0027	0.964	0.952	0.2154
	EE	0.686	0.0026	0.0073	0.0027	0.964	0.944	0.2137
5000	TMLE	0.686	0.0024	0.0090	0.0025	0.956	0.950	0.1929
	EE	0.686	0.0024	0.0101	0.0025	0.946	0.944	0.1915
10000	TMLE	0.686	0.0012	0.0057	0.0012	0.952	0.944	0.1363
	EE	0.686	0.0012	0.0058	0.0012	0.954	0.942	0.1358
20000	TMLE	0.686	0.0005	0.0044	0.0006	0.960	0.952	0.0962
	EE	0.686	0.0005	0.0041	0.0006	0.958	0.952	0.0960

Table 4: Performance of TMLE and EE for Theta (earth est Q and g, DR-learner, no CV)

n	Method	True_Theta	Variance	Bias	MSE	Coverage	Coverage_or	CI_width
500	TMLE	0.686	0.0930	0.0069	0.0930	0.918	0.978	0.7526
	EE	0.686	0.3765	0.1823	0.4097	0.920	0.968	1.1256
1000	TMLE	0.686	0.0274	-0.0071	0.0275	0.932	0.980	0.4766
	EE	0.686	0.0467	0.0648	0.0509	0.932	0.972	0.5946
2000	TMLE	0.686	0.0077	-0.0053	0.0077	0.944	0.964	0.3217
	EE	0.686	0.0245	0.0496	0.0269	0.952	0.964	0.4007
3000	TMLE	0.686	0.0053	-0.0072	0.0053	0.920	0.960	0.2539
	EE	0.686	0.0071	0.0293	0.0080	0.926	0.954	0.2831
4000	TMLE	0.686	0.0032	-0.0148	0.0035	0.934	0.950	0.2171
	EE	0.686	0.0056	0.0157	0.0059	0.952	0.976	0.2379
5000	TMLE	0.686	0.0027	-0.0170	0.0029	0.930	0.944	0.1978
	EE	0.686	0.0033	0.0075	0.0033	0.958	0.964	0.2069
10000	TMLE	0.686	0.0014	-0.0174	0.0017	0.884	0.922	0.1360
	EE	0.686	0.0013	-0.0018	0.0013	0.926	0.944	0.1379
20000	TMLE	0.686	0.0006	-0.0160	0.0009	0.876	0.896	0.0957
	EE	0.686	0.0006	-0.0036	0.0006	0.942	0.950	0.0968

Table 5: Performance of TMLE and EE for Theta (earth est Q and g, T-learner, no CV)

n	Method	True_Theta	Variance	Bias	MSE	Coverage	Coverage_or	CI_width
500	TMLE	0.686	0.0239	0.0121	0.0240	0.958	0.962	0.6459
	EE	0.686	0.0254	-0.0018	0.0255	0.962	0.956	0.6530
1000	TMLE	0.686	0.0114	0.0093	0.0115	0.958	0.940	0.4492
	EE	0.686	0.0122	-0.0005	0.0122	0.952	0.940	0.4557
2000	TMLE	0.686	0.0065	0.0002	0.0065	0.946	0.948	0.3133
	EE	0.686	0.0066	-0.0051	0.0066	0.942	0.948	0.3171
3000	TMLE	0.686	0.0042	-0.0053	0.0042	0.934	0.936	0.2500
	EE	0.686	0.0043	-0.0088	0.0044	0.944	0.940	0.2529
4000	TMLE	0.686	0.0028	-0.0087	0.0029	0.942	0.940	0.2165
	EE	0.686	0.0028	-0.0107	0.0029	0.946	0.942	0.2186
5000	TMLE	0.686	0.0025	-0.0060	0.0026	0.942	0.946	0.1924
	EE	0.686	0.0026	-0.0082	0.0026	0.942	0.946	0.1946
10000	TMLE	0.686	0.0012	-0.0099	0.0013	0.942	0.944	0.1355
	EE	0.686	0.0012	-0.0108	0.0013	0.938	0.946	0.1372
20000	TMLE	0.686	0.0006	-0.0119	0.0007	0.920	0.920	0.0955
	EE	0.686	0.0006	-0.0125	0.0007	0.924	0.916	0.0966

Table 6: Performance of TMLE and EE for Theta (Correct Q and g, DR-learner, CV)

n	Method	True_Theta	Variance	Bias	MSE	Coverage	Coverage_or	CI_width
500	TMLE	0.686	0.0299	0.0302	0.0308	0.964	0.960	0.6897
	EE	0.686	0.0868	-0.1703	0.1158	0.852	0.962	0.8408
1000	TMLE	0.686	0.0129	0.0162	0.0132	0.960	0.950	0.4583
	EE	0.686	0.0168	-0.0946	0.0257	0.870	0.894	0.4990
2000	TMLE	0.686	0.0058	0.0072	0.0059	0.960	0.954	0.3119
	EE	0.686	0.0068	-0.0512	0.0094	0.902	0.914	0.3246
3000	TMLE	0.686	0.0041	0.0094	0.0041	0.948	0.944	0.2536
	EE	0.686	0.0045	-0.0306	0.0054	0.908	0.930	0.2594
4000	TMLE	0.686	0.0030	0.0040	0.0030	0.964	0.956	0.2169
	EE	0.686	0.0032	-0.0231	0.0038	0.918	0.932	0.2202
5000	TMLE	0.686	0.0023	0.0023	0.0023	0.958	0.954	0.1937
	EE	0.686	0.0024	-0.0206	0.0028	0.922	0.922	0.1956
10000	TMLE	0.686	0.0013	0.0036	0.0013	0.930	0.946	0.1366
	EE	0.686	0.0013	-0.0084	0.0014	0.918	0.936	0.1367
20000	TMLE	0.686	0.0006	0.0055	0.0006	0.936	0.936	0.0966
	EE	0.686	0.0006	-0.0053	0.0006	0.942	0.946	0.0962

Table 7: Performance of TMLE and EE for Theta (Correct Q and g, T-learner, CV)

n	Method	True_Theta	Variance	Bias	MSE	Coverage	Coverage_or	CI_width
500	TMLE	0.686	0.0252	-0.0295	0.0261	0.926	0.950	0.6457
	EE	0.686	0.0277	-0.0506	0.0302	0.918	0.932	0.6667
1000	TMLE	0.686	0.0111	-0.0136	0.0113	0.950	0.950	0.4395
	EE	0.686	0.0118	-0.0207	0.0122	0.946	0.944	0.4456
2000	TMLE	0.686	0.0061	-0.0055	0.0062	0.942	0.944	0.3068
	EE	0.686	0.0061	-0.0086	0.0062	0.938	0.940	0.3085
3000	TMLE	0.686	0.0039	-0.0049	0.0040	0.940	0.938	0.2488
	EE	0.686	0.0040	-0.0067	0.0040	0.938	0.944	0.2496
4000	TMLE	0.686	0.0027	-0.0052	0.0027	0.952	0.938	0.2153
	EE	0.686	0.0026	-0.0059	0.0027	0.954	0.940	0.2159
5000	TMLE	0.686	0.0025	-0.0016	0.0025	0.950	0.964	0.1928
	EE	0.686	0.0025	-0.0008	0.0025	0.952	0.958	0.1931
10000	TMLE	0.686	0.0012	-0.0009	0.0012	0.952	0.954	0.1363
	EE	0.686	0.0012	-0.0001	0.0012	0.954	0.954	0.1364
20000	TMLE	0.686	0.0005	0.0006	0.0005	0.958	0.950	0.0962
	EE	0.686	0.0005	0.0008	0.0005	0.958	0.950	0.0962

Table 8: Performance of TMLE and EE for Theta (earth est Q and g, DR-learner, CV)

n	Method	True_Theta	Variance	Bias	MSE	Coverage	Coverage_or	CI_width
500	TMLE	0.686	0.0913	0.0484	0.0936	0.948	0.968	0.9364
	EE	0.686	0.5247	-0.2821	0.6043	0.844	0.948	1.4933
1000	TMLE	0.686	0.0209	0.0283	0.0217	0.952	0.954	0.5865
	EE	0.686	0.1500	-0.1901	0.1861	0.850	0.944	0.8569
2000	TMLE	0.686	0.0092	0.0108	0.0094	0.946	0.950	0.3653
	EE	0.686	0.0345	-0.0549	0.0375	0.884	0.970	0.4765
3000	TMLE	0.686	0.0053	0.0006	0.0053	0.938	0.954	0.2753
	EE	0.686	0.0151	-0.0443	0.0171	0.898	0.960	0.3244
4000	TMLE	0.686	0.0034	-0.0102	0.0035	0.940	0.956	0.2270
	EE	0.686	0.0085	-0.0356	0.0098	0.912	0.966	0.2580
5000	TMLE	0.686	0.0030	-0.0125	0.0032	0.930	0.954	0.2065
	EE	0.686	0.0047	-0.0235	0.0053	0.936	0.966	0.2282
10000	TMLE	0.686	0.0014	-0.0179	0.0017	0.878	0.920	0.1369
	EE	0.686	0.0016	-0.0108	0.0017	0.918	0.944	0.1412
20000	TMLE	0.686	0.0007	-0.0170	0.0009	0.870	0.900	0.0966
	EE	0.686	0.0007	-0.0075	0.0007	0.932	0.950	0.0969

Table 9: Performance of TMLE and EE for Theta (earth est Q and g, T-learner, CV)

n	Method	True_Theta	Variance	Bias	MSE	Coverage	Coverage_or	CI_width
500	TMLE	0.686	0.0287	-0.0312	0.0297	0.916	0.942	0.6818
	EE	0.686	0.0436	-0.0780	0.0497	0.910	0.950	0.7574
1000	TMLE	0.686	0.0134	-0.0229	0.0139	0.928	0.944	0.4649
	EE	0.686	0.0155	-0.0501	0.0180	0.924	0.930	0.4948
2000	TMLE	0.686	0.0063	-0.0168	0.0066	0.942	0.946	0.3184
	EE	0.686	0.0067	-0.0295	0.0075	0.940	0.942	0.3299
3000	TMLE	0.686	0.0043	-0.0128	0.0045	0.928	0.946	0.2541
	EE	0.686	0.0046	-0.0203	0.0050	0.932	0.950	0.2610
4000	TMLE	0.686	0.0030	-0.0176	0.0034	0.934	0.944	0.2166
	EE	0.686	0.0031	-0.0230	0.0036	0.934	0.936	0.2214
5000	TMLE	0.686	0.0024	-0.0195	0.0028	0.926	0.938	0.1932
	EE	0.686	0.0025	-0.0226	0.0030	0.930	0.928	0.1969
10000	TMLE	0.686	0.0013	-0.0181	0.0017	0.886	0.912	0.1352
	EE	0.686	0.0013	-0.0189	0.0017	0.892	0.904	0.1372
20000	TMLE	0.686	0.0006	-0.0168	0.0009	0.884	0.902	0.0953
	EE	0.686	0.0006	-0.0173	0.0009	0.894	0.896	0.0967

Table 10: Performance of SS and TMLE for Theta (earth est Q and g, DR-learner, no CV)

n	Method	True_Theta	Variance	Bias	MSE
1000	SS	0.686	0.4599	4.3097	19.0335
	TMLE	0.686	0.0274	-0.0071	0.0275

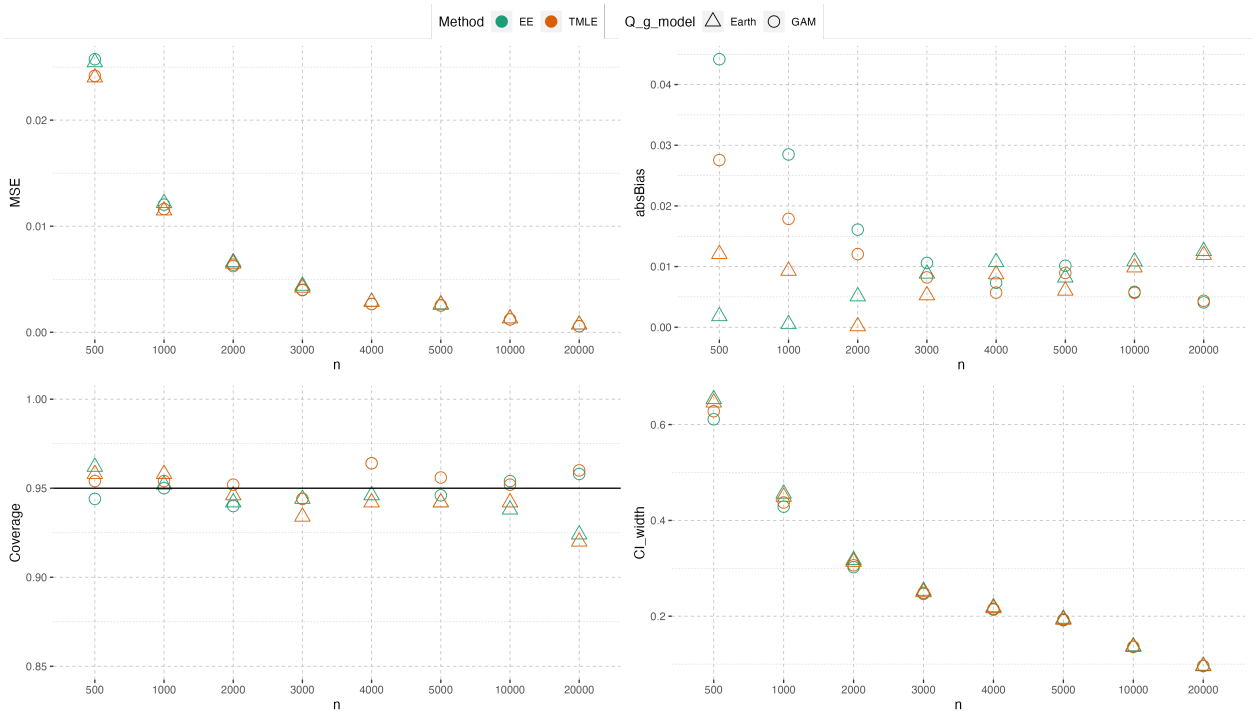

Figure 1: Main performance metrics (T-learner, no CV)

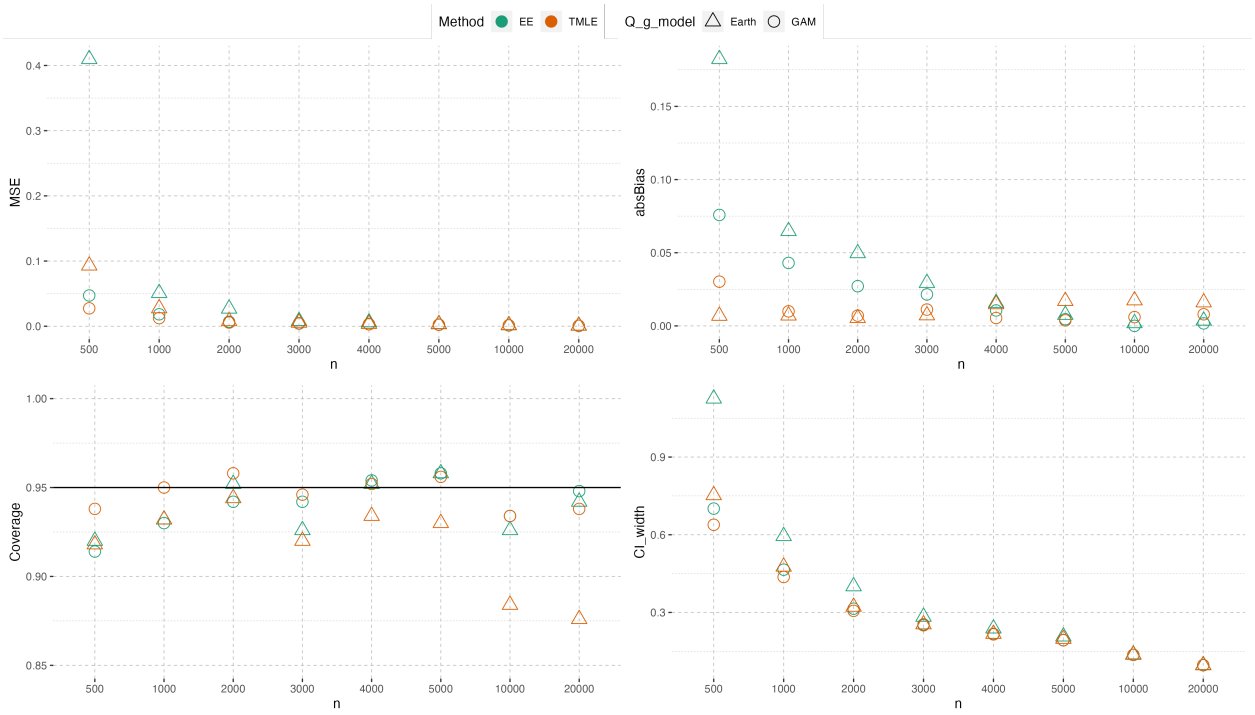

Figure 2: Main performance metrics (DR-learner, no CV)

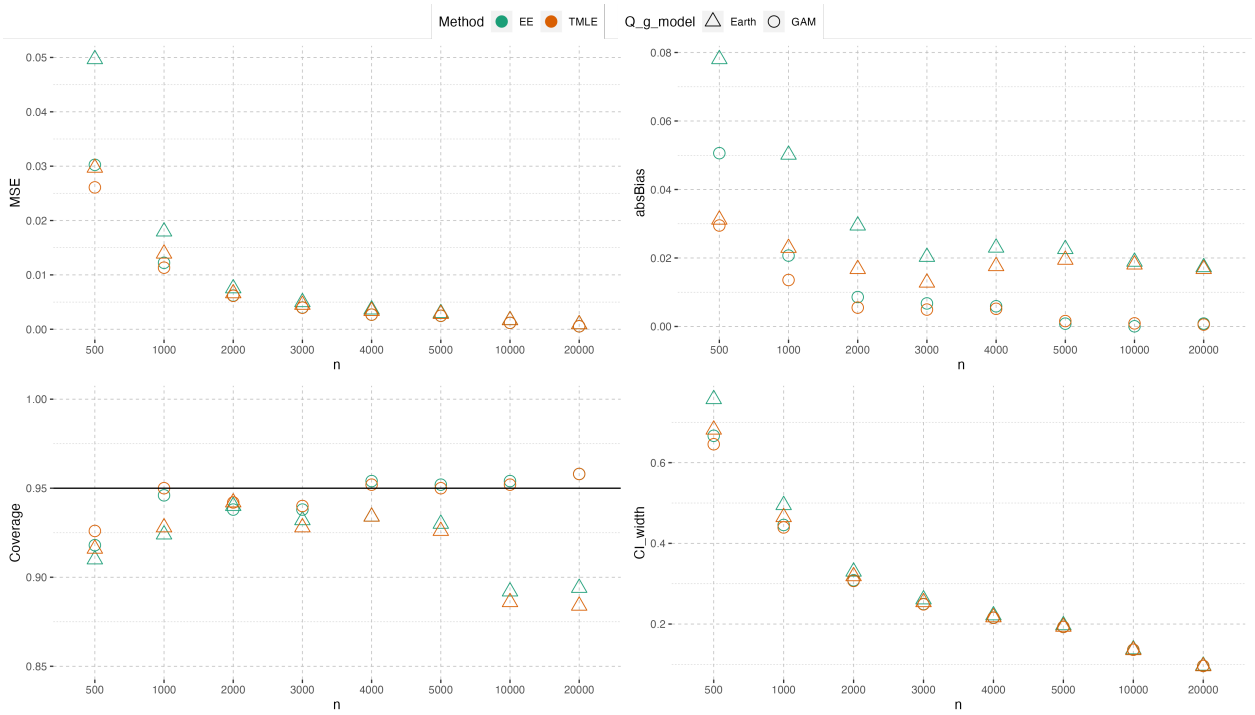

Figure 3: Main performance metrics (T-learner, CV)

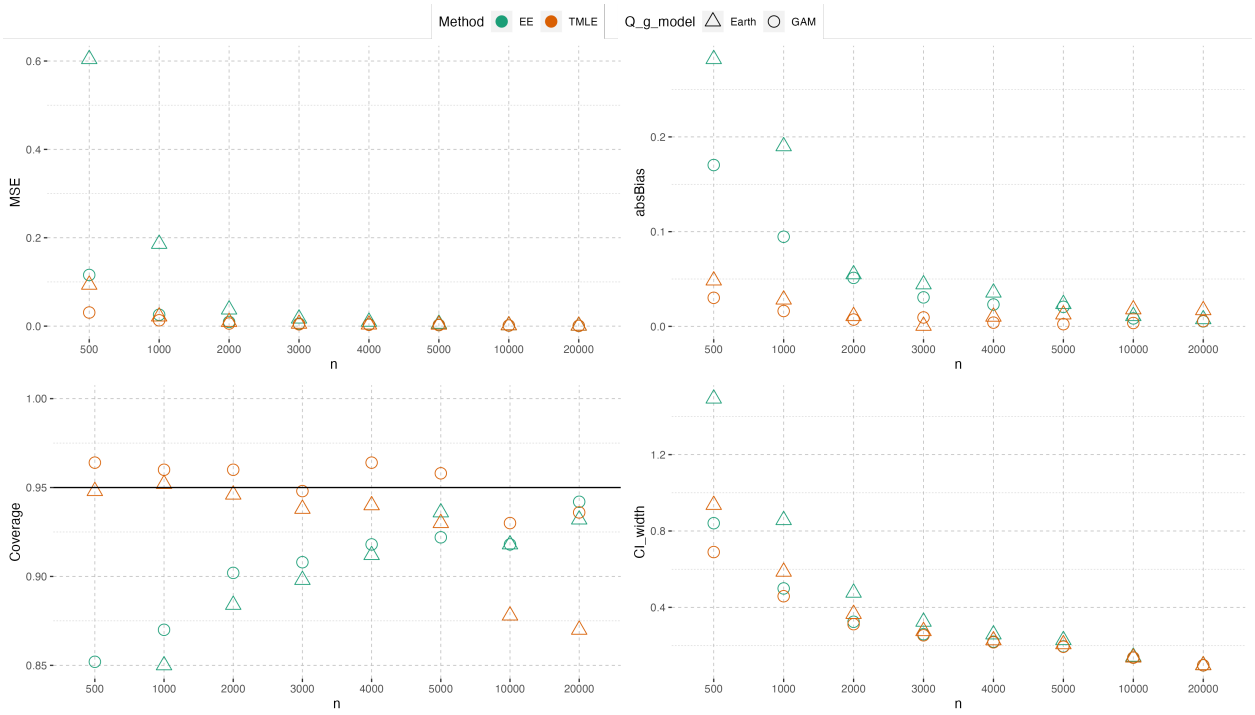

Figure 4: Main performance metrics (DR-learner, CV)

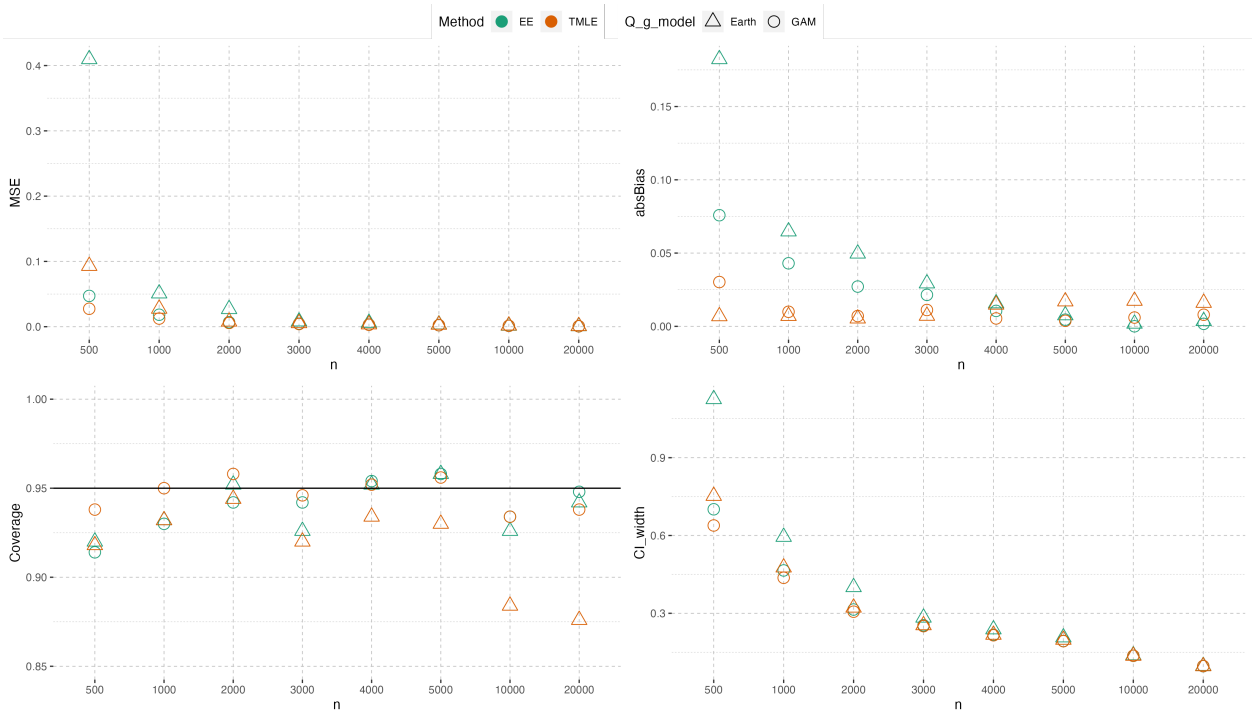

Figure 5: Main performance metrics (DR-learner, no CV)

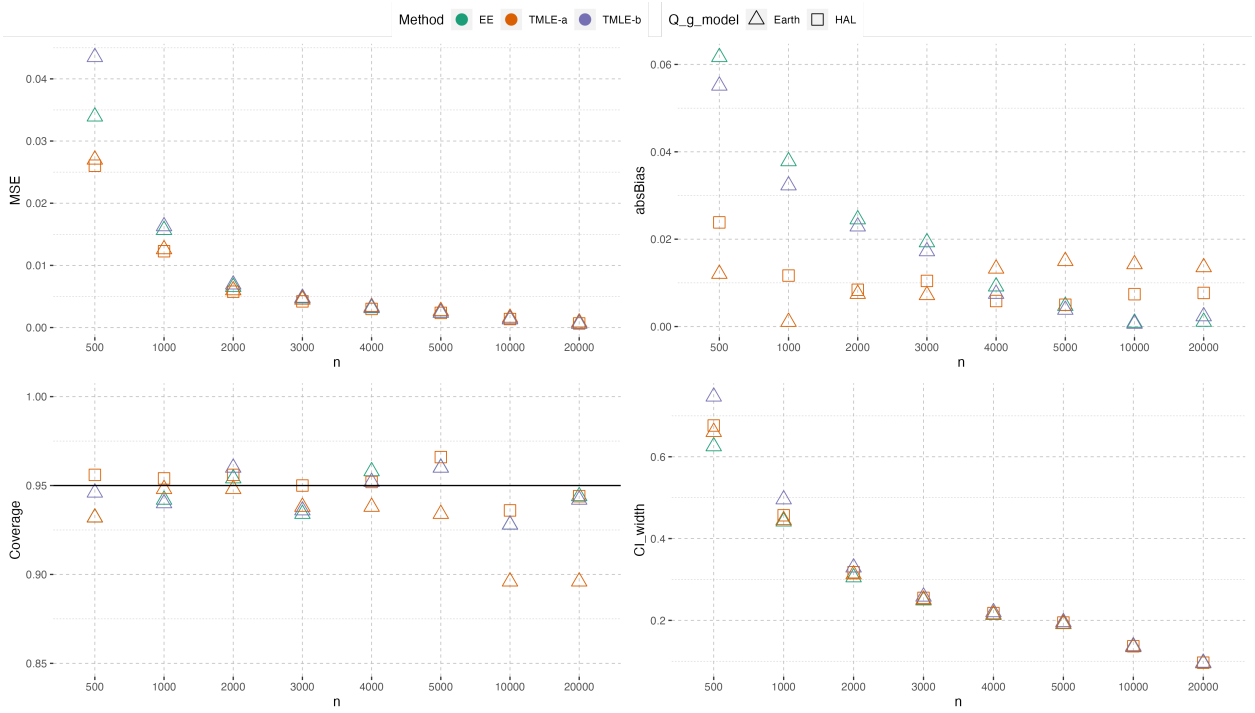

Figure 6: Main performance metrics (DR-learner, no CV)

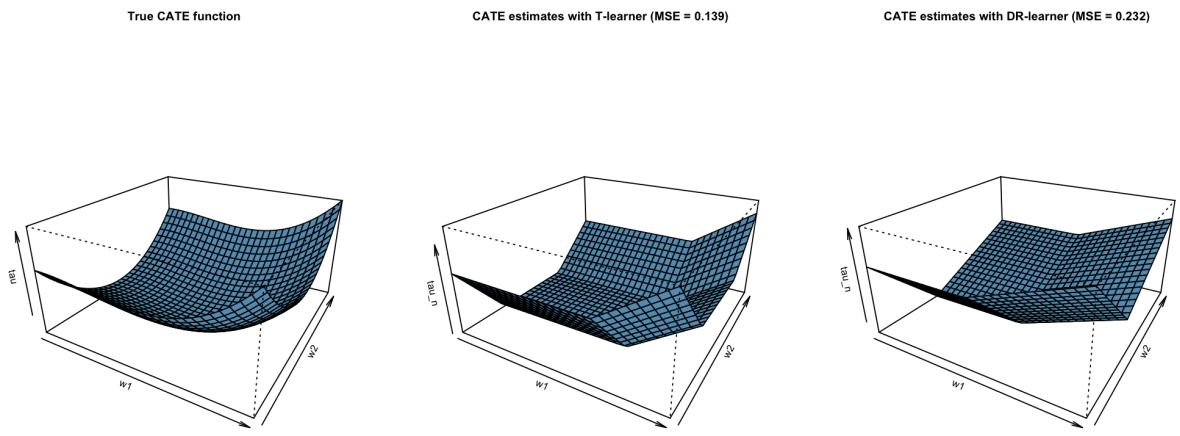

Figure 7: CATE estimation with Earth (n = 500)

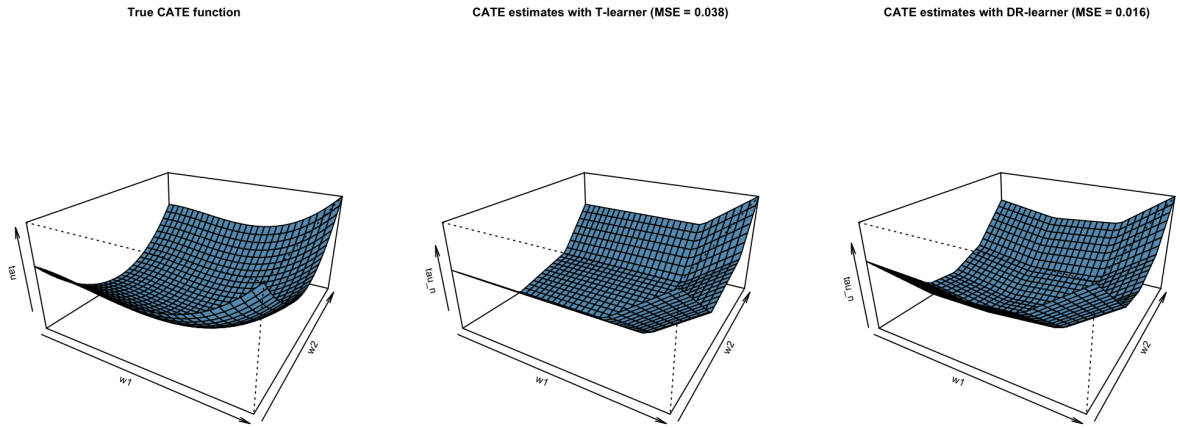

Figure 8: CATE estimation with Earth (n = 5000)

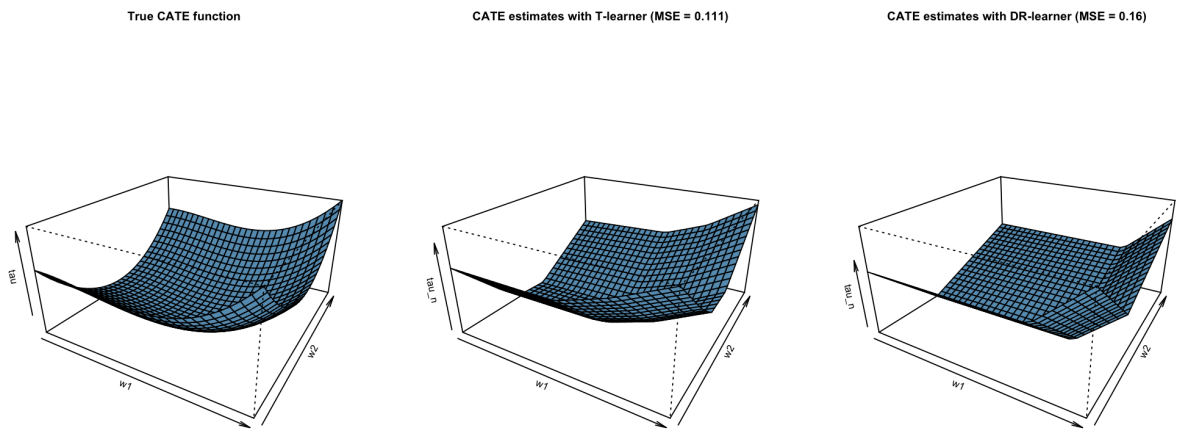

Figure 9: CATE estimation with HAL (n = 500)

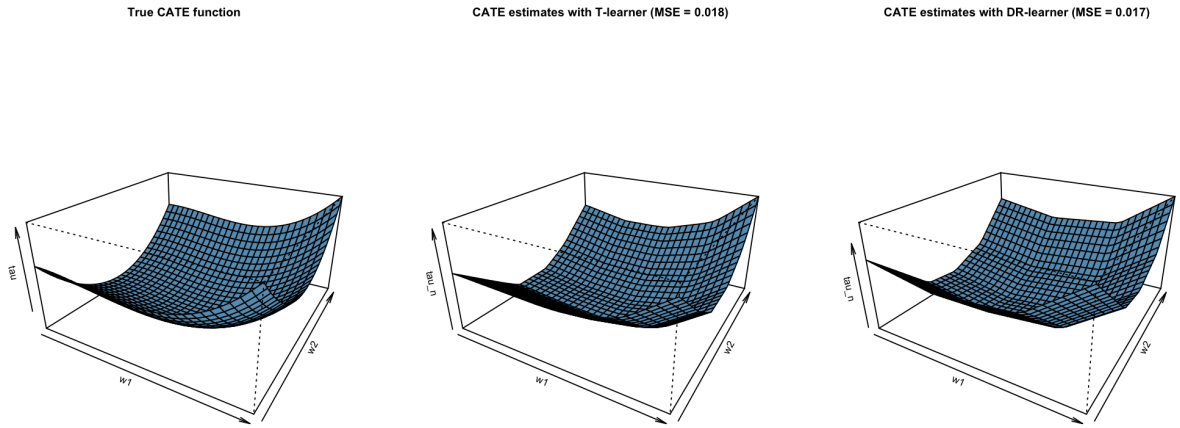

Figure 10: CATE estimation with HAL (n = 5000)

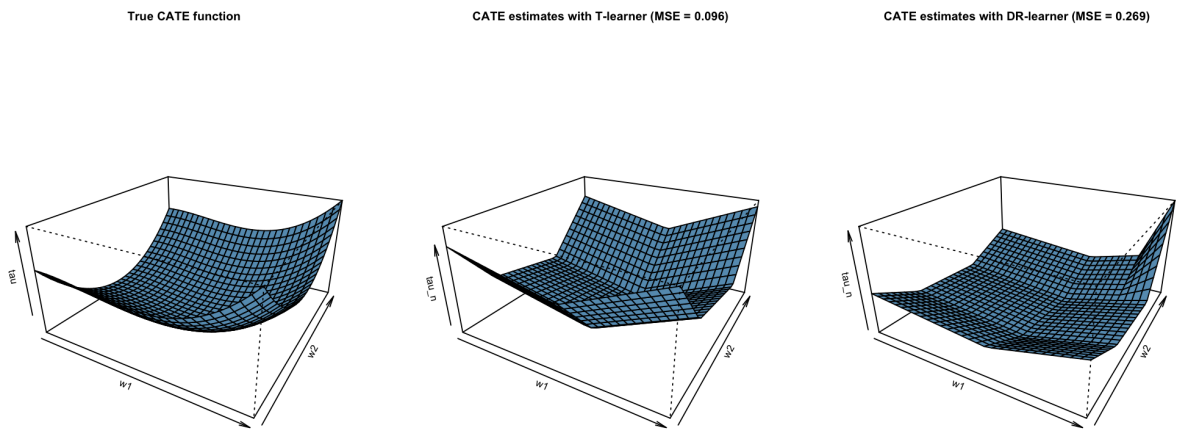

Figure 11: CATE estimation with Earth (n = 500, seed 2)

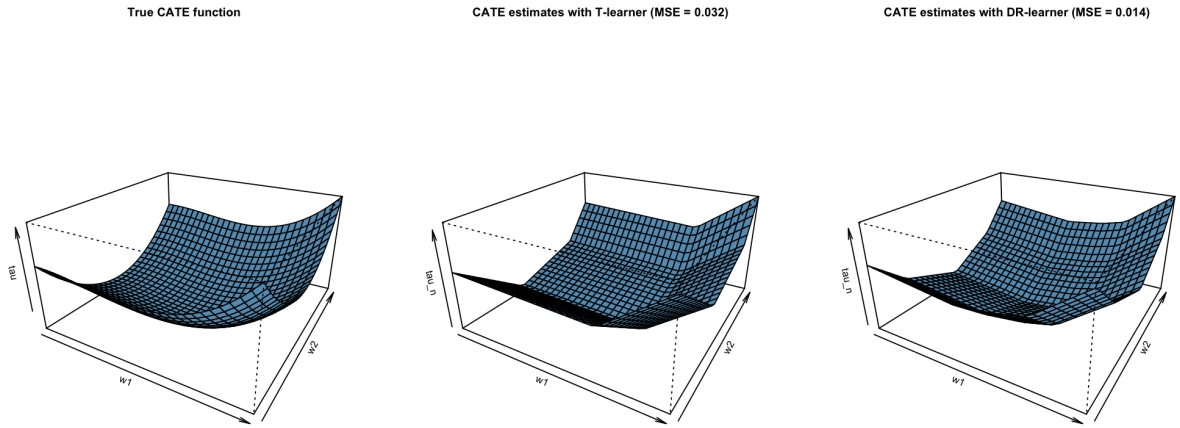

Figure 12: CATE estimation with Earth (n = 5000, seed 2)

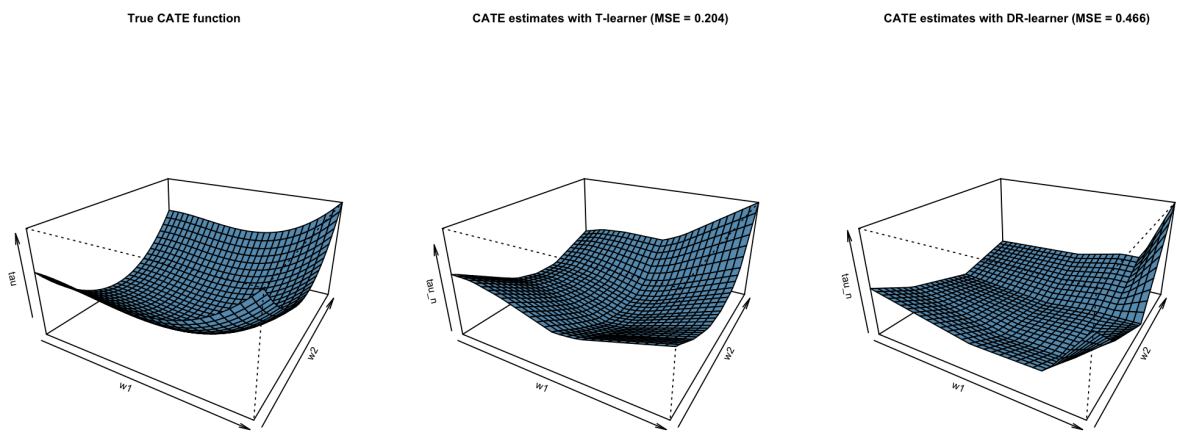

Figure 13: CATE estimation with HAL (n = 500, seed 2)

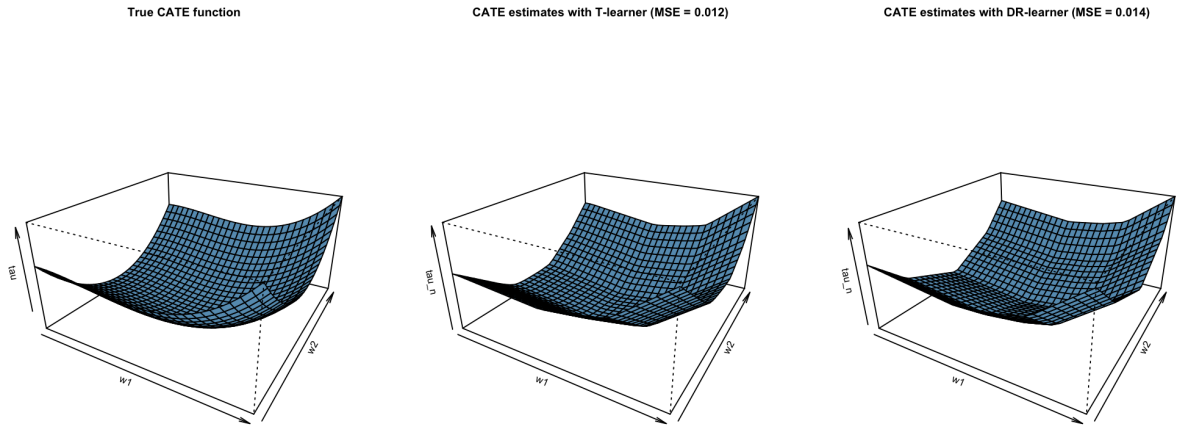

Figure 14: CATE estimation with HAL (n = 5000, seed 2)