

Table 1: Evaluation of CATE-based, dual, and primal tests on type I error and power across various data settings.

<i>N</i>	All binary				Y continuous				Y,X continuous			
	<i>Type I error</i>		<i>Power</i>		<i>Type I error</i>		<i>Power</i>		<i>Type I error</i>		<i>Power</i>	
	ADMG1	ADMG2	ADMG3	ADMG4	ADMG1	ADMG2	ADMG3	ADMG4	ADMG1	ADMG2	ADMG3	ADMG4
CATE-based test												
500	0.07	0.12	0.215	0.07	0.06	0.225	0.73	0.15	0.06	0.05	0.245	0.69
1000	0.03	0.06	0.455	0.06	0.035	0.16	0.98	0.33	0.06	0.065	0.395	0.935
2000	0.035	0.035	0.87	0.07	0.015	0.125	1	0.565	0.04	0.08	0.64	1
4000	0.02	0.055	0.995	0.195	0.035	0.05	1	0.93	0.05	0.055	0.93	1
10000	0.045	0.025	1	0.69	0.05	0.05	1	1	0.05	0.06	1	1
Dual test												
500	0.065	0.105	0.415	0.065	0.075	0.03	0.865	0.085	0.05	0.015	0.12	0.37
1000	0.065	0.035	0.805	0.095	0.035	0.075	0.995	0.11	0.04	0.065	0.14	0.765
2000	0.06	0.04	0.98	0.09	0.045	0.07	1	0.215	0.04	0.065	0.265	0.965
4000	0.035	0.09	1	0.135	0.04	0.05	1	0.505	0.02	0.04	0.525	1
10000	0.065	0.085	1	0.325	0.035	0.045	1	0.965	0.045	0.055	0.97	1
Primal test												
500	0.07	0.04	0.37	0.04	0.07	0.06	0.285	0.08	0.055	0.035	0.09	0.17
1000	0.025	0.01	0.73	0.055	0.055	0.025	0.515	0.05	0.07	0.06	0.09	0.22
2000	0.02	0.025	0.985	0.05	0.055	0.04	0.855	0.11	0.06	0.08	0.15	0.495
4000	0.025	0.015	1	0.13	0.055	0.02	0.995	0.1	0.04	0.045	0.345	0.65
10000	0.02	0	1	0.57	0.06	0.04	1	0.275	0.08	0.035	0.86	0.94