Linear Model Simulations

Kellie Ottoboni

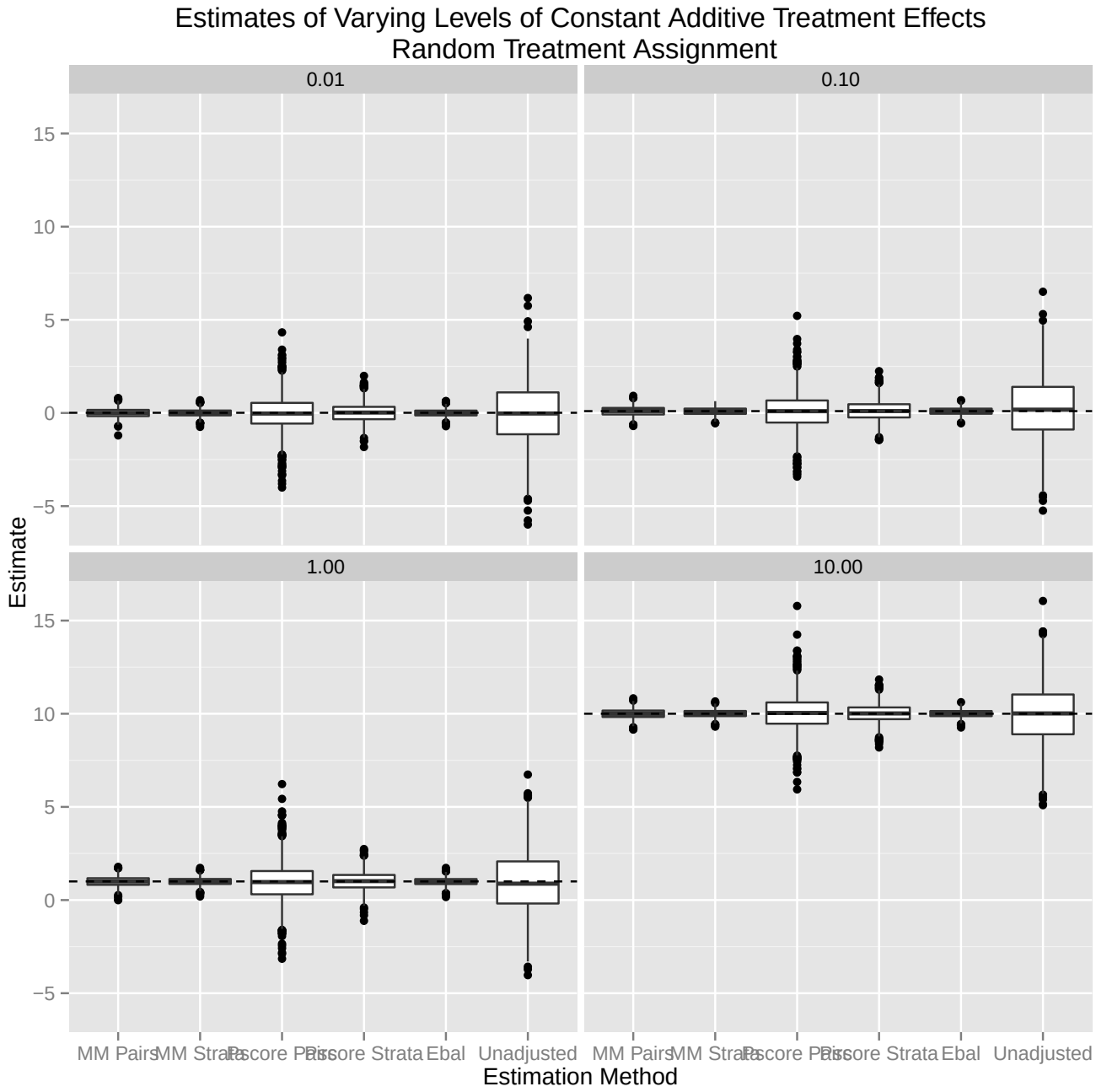
March 4, 2016

Set simulation parameters

```
set.seed(321)
gamma <- c(0.01, 0.1, 1, 10)
B <- 1000
N <- 100
```

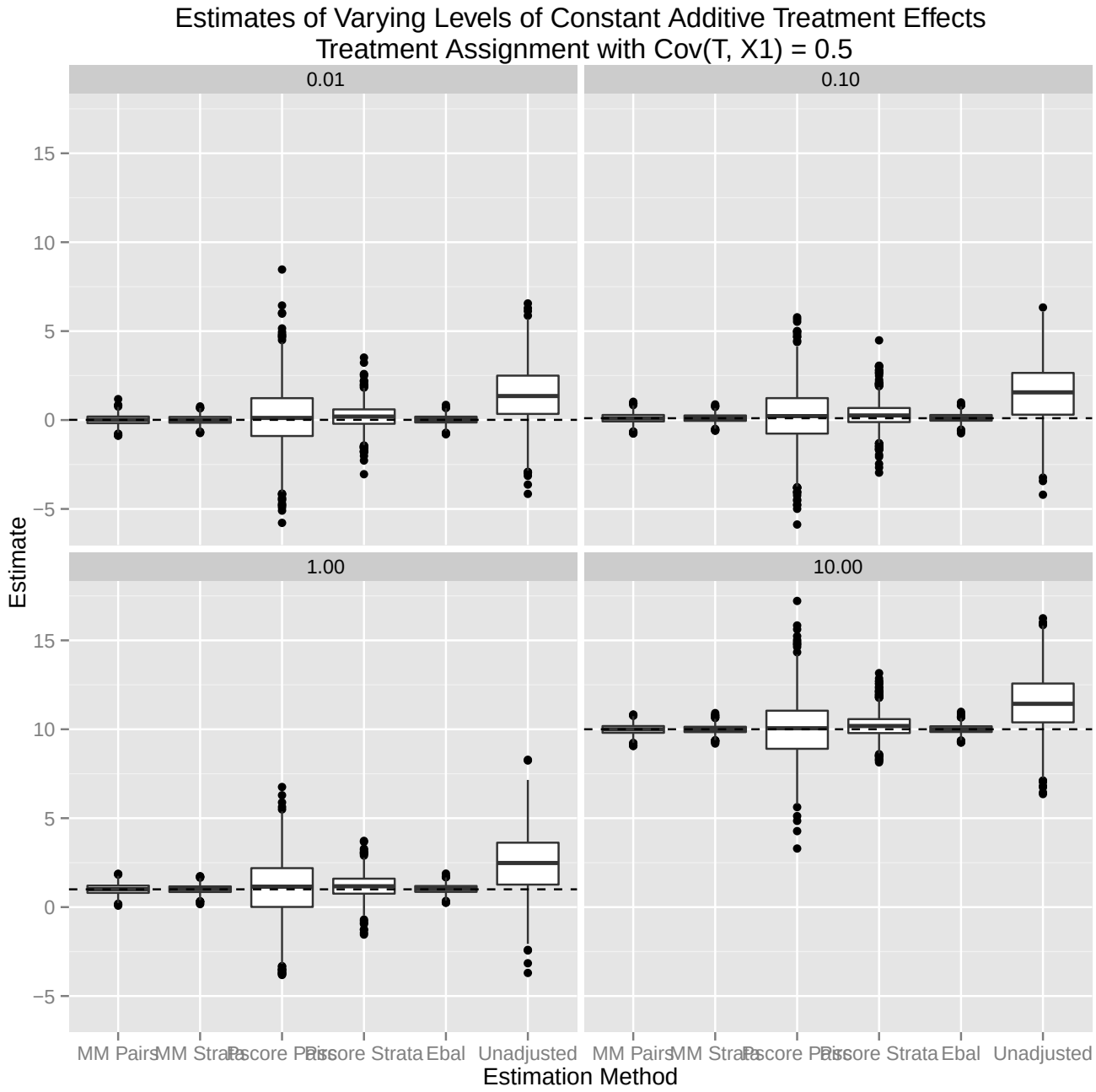
1 Comparing Model-based Matching to Other Methods

1.1 Gaussian Errors in the Linear Model



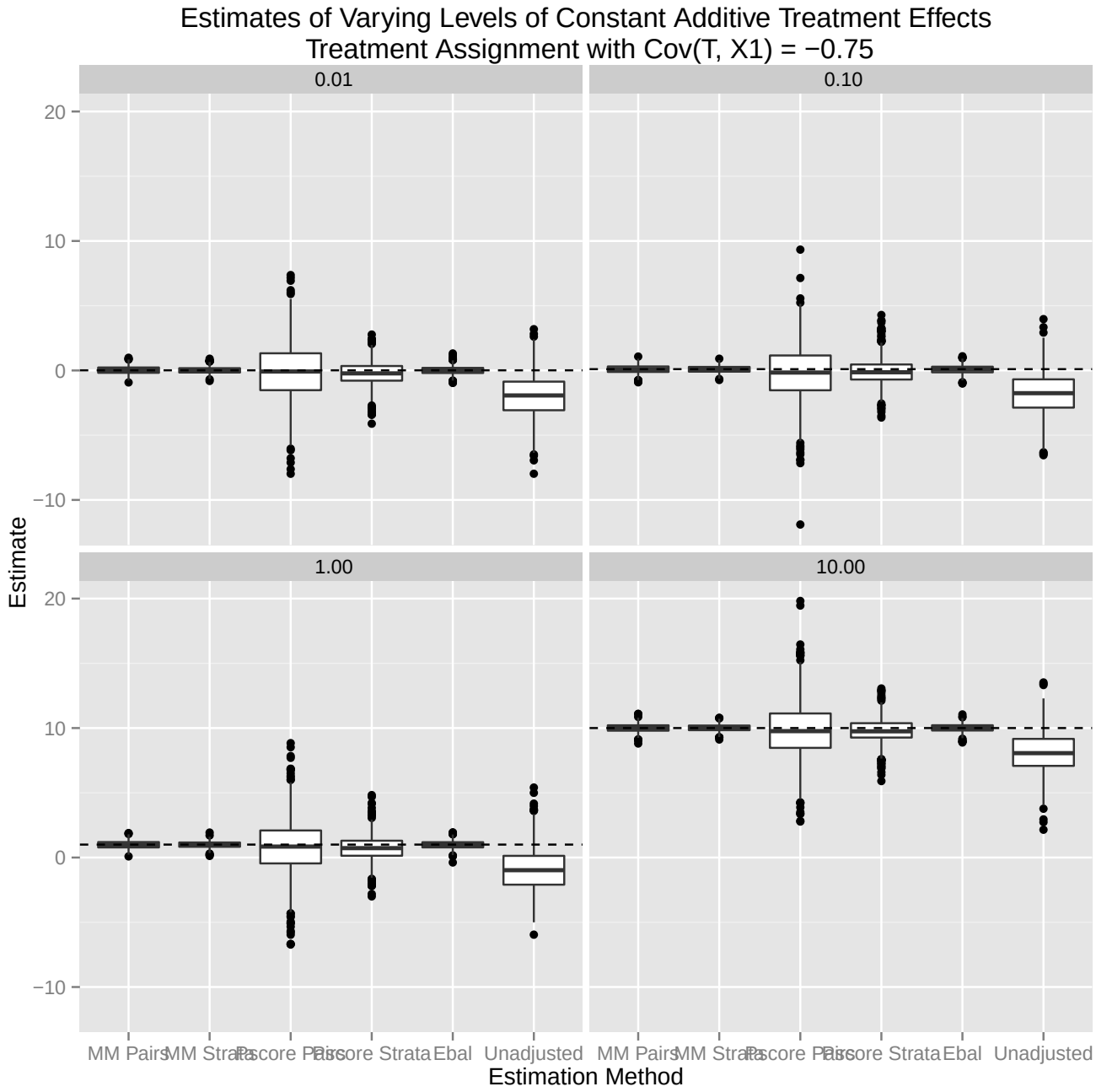
	0.01	0.1	1	10
MM Pairs	0.257	0.263	0.264	0.260
MM Strata	0.203	0.200	0.210	0.204
Pscore Pairs	1.030	1.033	1.139	1.043
Pscore Strata	0.520	0.548	0.519	0.512
Ebal	0.206	0.201	0.212	0.205
Unadjusted	1.683	1.705	1.646	1.579

Table 1: RMSE for various treatment effects; Random Treatment Assignment



	0.01	0.1	1	10
MM Pairs	0.287	0.291	0.296	0.296
MM Strata	0.239	0.234	0.234	0.237
Pscore Pairs	1.736	1.661	1.686	1.701
Pscore Strata	0.705	0.735	0.683	0.689
Ebal	0.250	0.243	0.251	0.248
Unadjusted	2.132	2.220	2.238	2.229

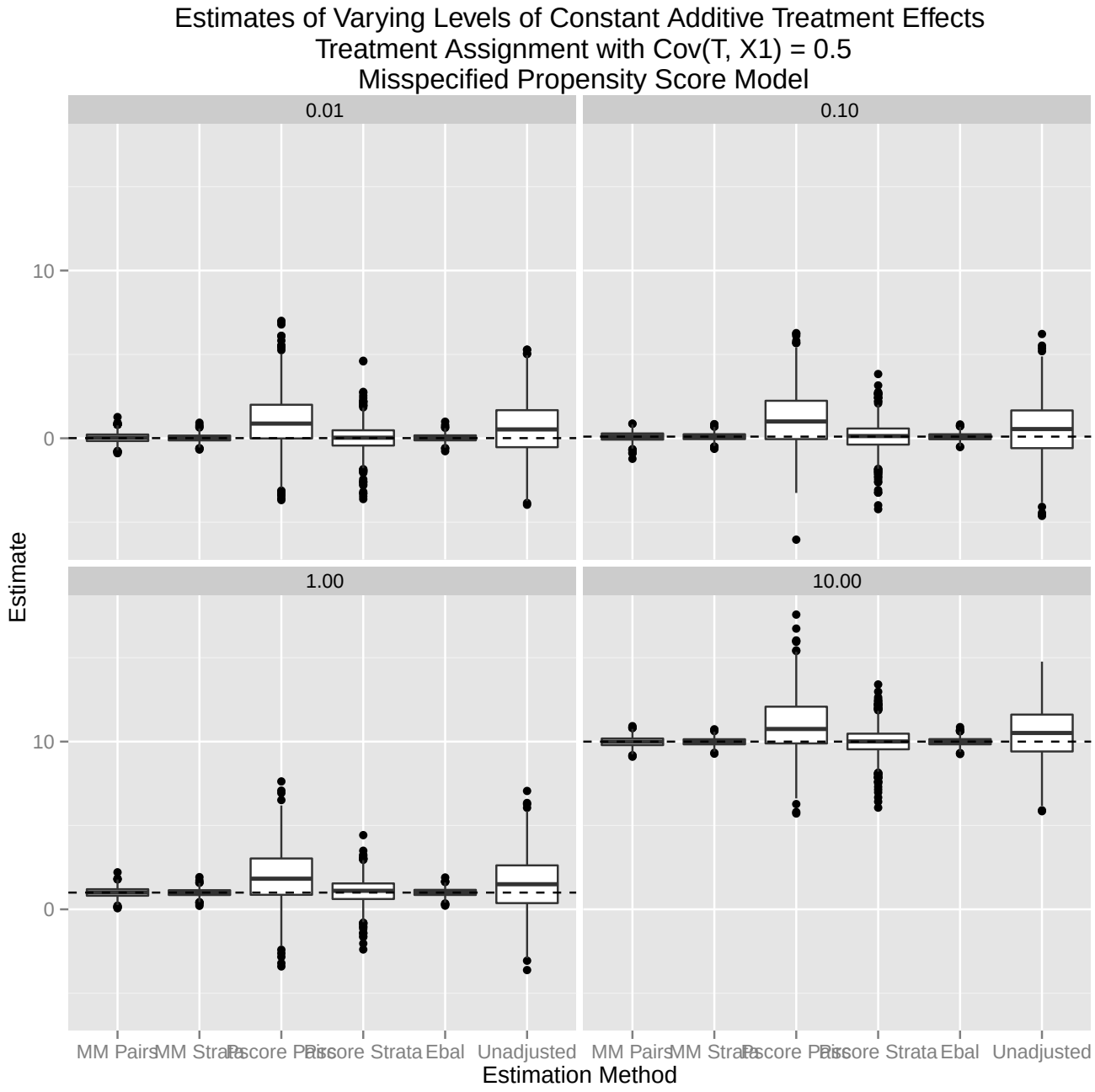
Table 2: RMSE for various treatment effects; Treatment Assignment with $cov(T, X_1) = 0.5$



	0.01	0.1	1	10
MM Pairs	0.302	0.313	0.292	0.304
MM Strata	0.256	0.266	0.245	0.262
Pscore Pairs	2.143	2.160	2.134	2.164
Pscore Strata	0.945	1.057	0.991	0.967
Ebal	0.310	0.319	0.300	0.308
Unadjusted	2.555	2.476	2.564	2.489

Table 3: RMSE for various treatment effects; Treatment Assignment with $cov(T, X_1) = -0.75$

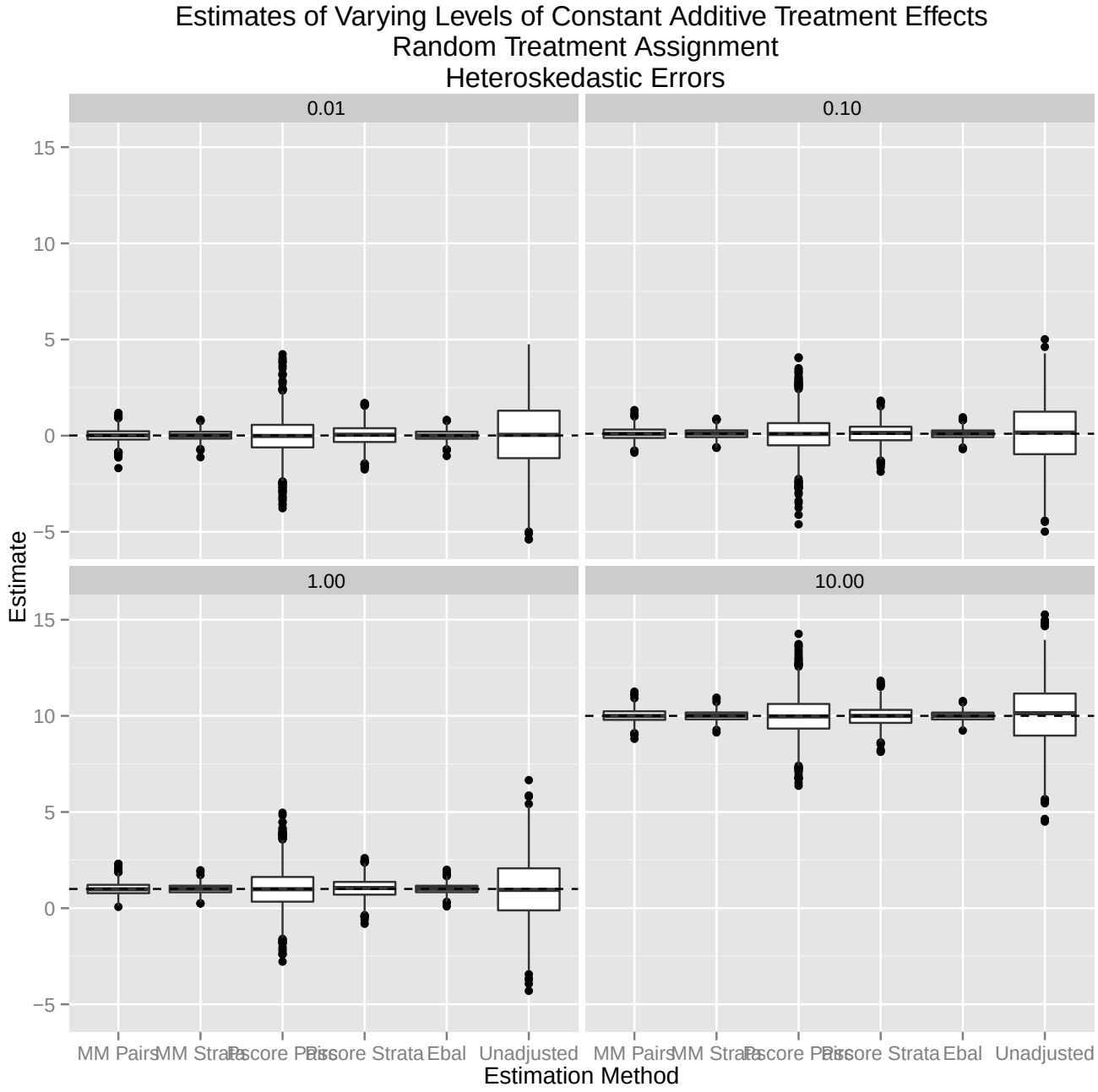
1.2 Misspecified Propensity Score Estimates



	0.01	0.1	1	10
MM Pairs	0.293	0.288	0.292	0.288
MM Strata	0.222	0.219	0.220	0.221
Pscore Pairs	1.892	1.953	1.887	1.918
Pscore Strata	0.850	0.839	0.780	0.829
Ebal	0.221	0.217	0.222	0.223
Unadjusted	1.766	1.694	1.769	1.725

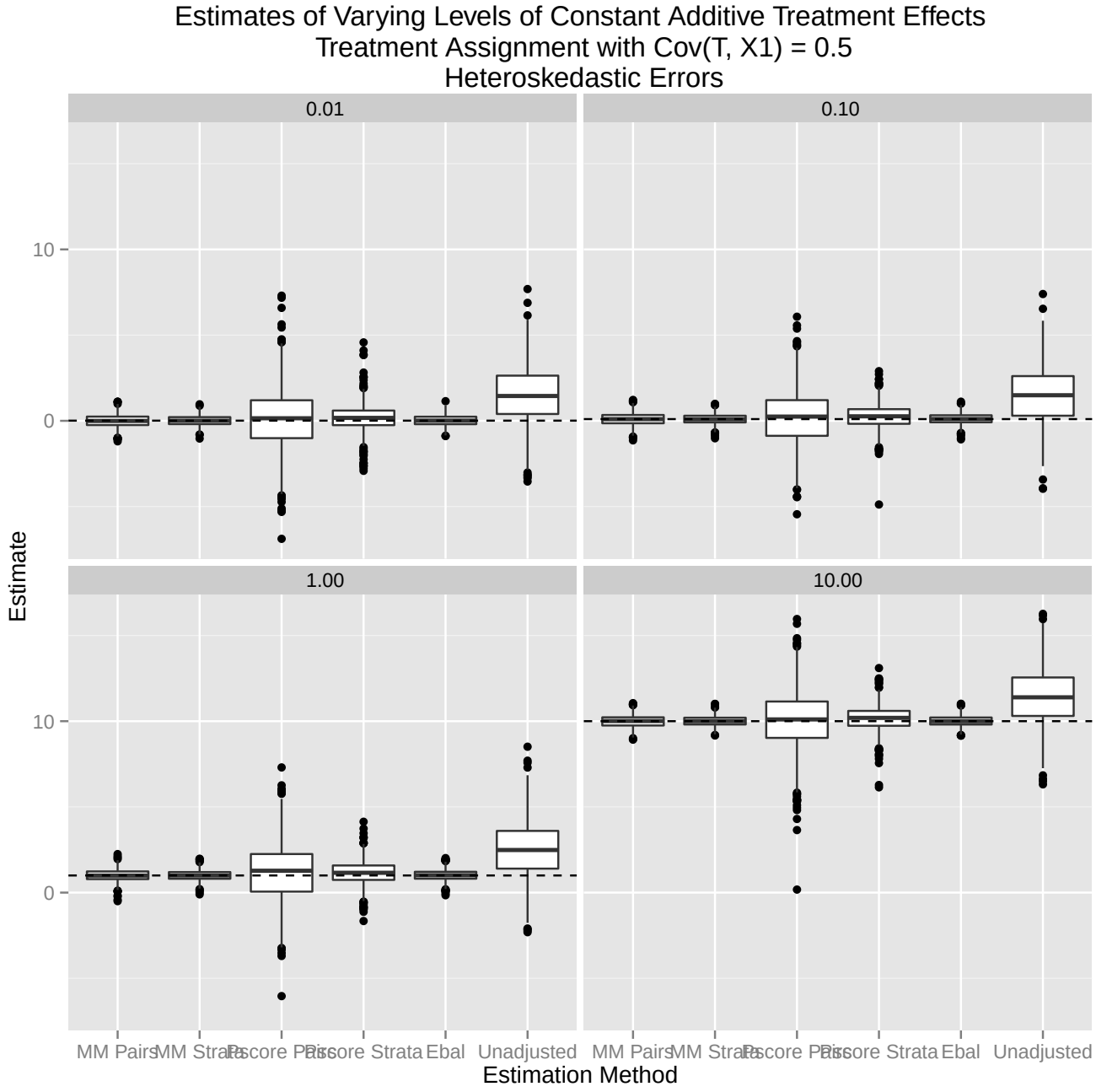
Table 4: RMSE for various treatment effects; Treatment Assignment with $cov(T, X_1) = 0.5$; Misspecified Propensity Score Model

1.3 Heteroskedastic Linear Model Errors



	0.01	0.1	1	10
MM Pairs	0.335	0.330	0.322	0.339
MM Strata	0.267	0.262	0.251	0.256
Pscore Pairs	1.077	1.088	1.146	1.091
Pscore Strata	0.548	0.548	0.521	0.538
Ebal	0.268	0.263	0.259	0.256
Unadjusted	1.753	1.640	1.641	1.627

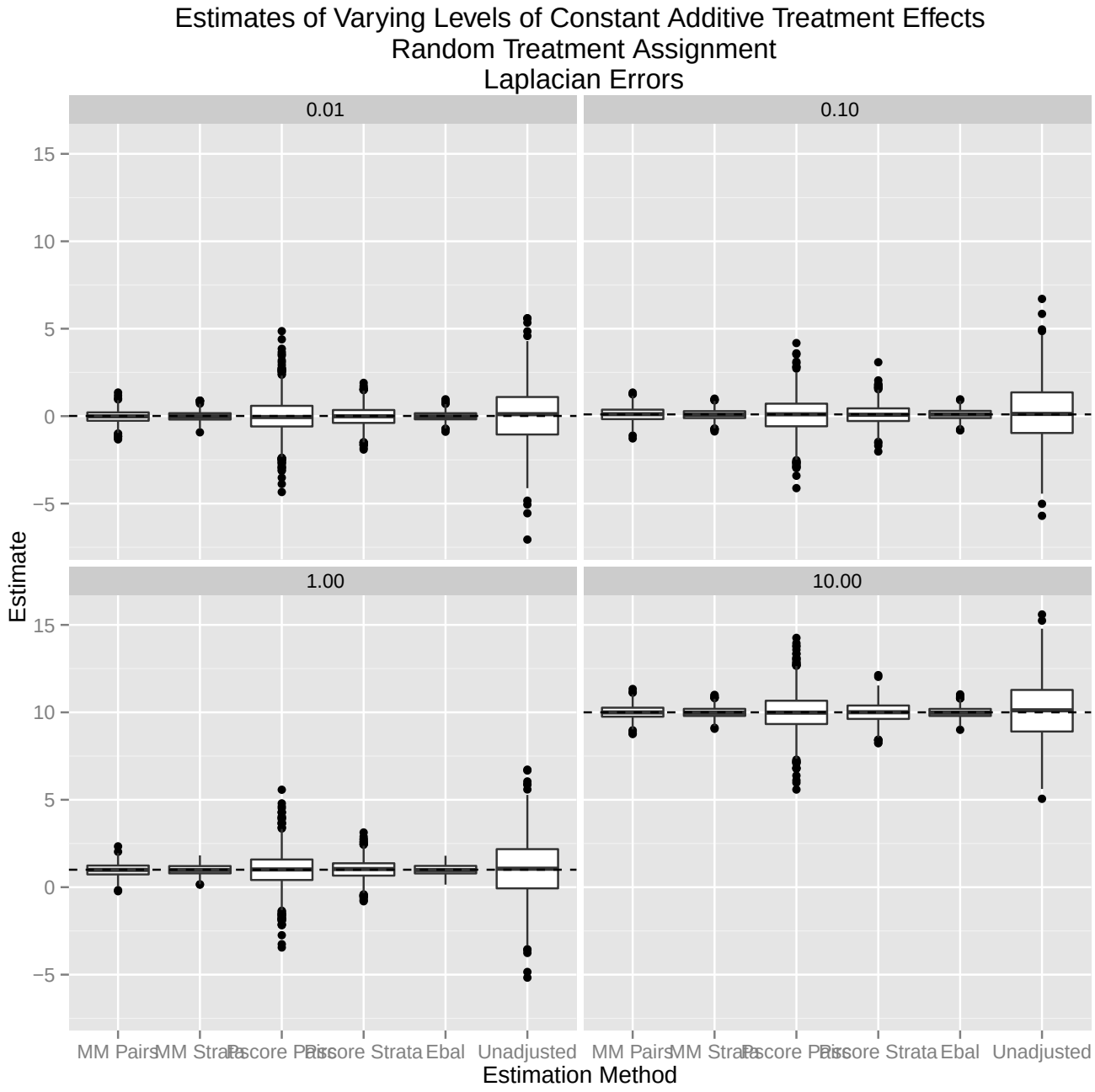
Table 5: RMSE for various treatment effects; Random Treatment Assignment; Heteroskedastic Errors



	0.01	0.1	1	10
MM Pairs	0.369	0.371	0.359	0.357
MM Strata	0.299	0.297	0.297	0.296
Pscore Pairs	1.789	1.679	1.692	1.735
Pscore Strata	0.782	0.706	0.692	0.735
Ebal	0.312	0.313	0.313	0.309
Unadjusted	2.246	2.099	2.257	2.205

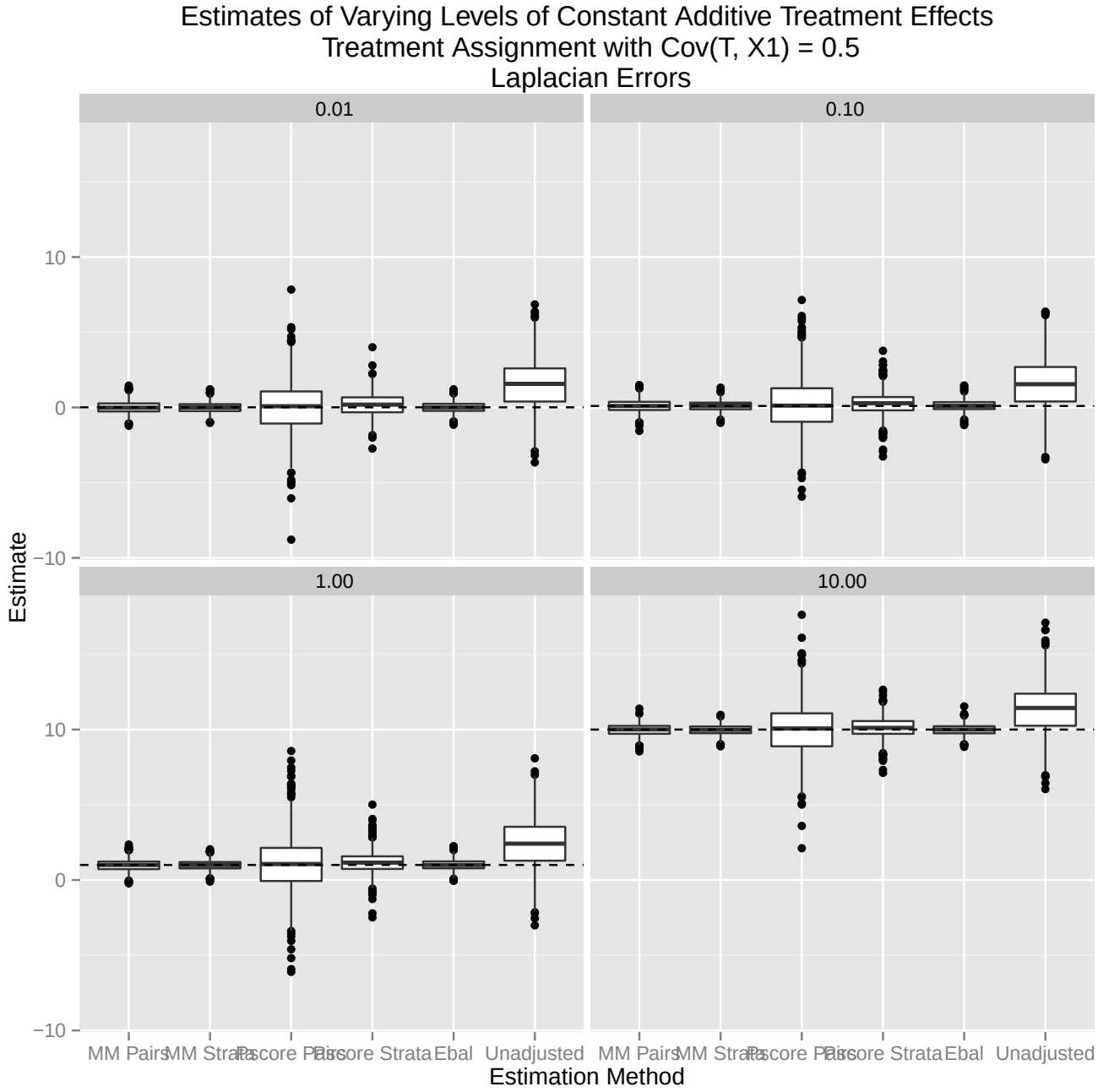
Table 6: RMSE for various treatment effects; Treatment Assignment with $cov(T, X_1) = 0.5$; Heteroskedastic Errors

1.4 Heavy-tailed Linear Model Errors



	0.01	0.1	1	10
MM Pairs	0.372	0.386	0.378	0.380
MM Strata	0.282	0.299	0.293	0.303
Pscore Pairs	1.076	1.043	1.063	1.134
Pscore Strata	0.569	0.572	0.566	0.573
Ebal	0.276	0.296	0.291	0.303
Unadjusted	1.670	1.700	1.694	1.722

Table 7: RMSE for various treatment effects; Random Treatment Assignment; Laplacian Errors



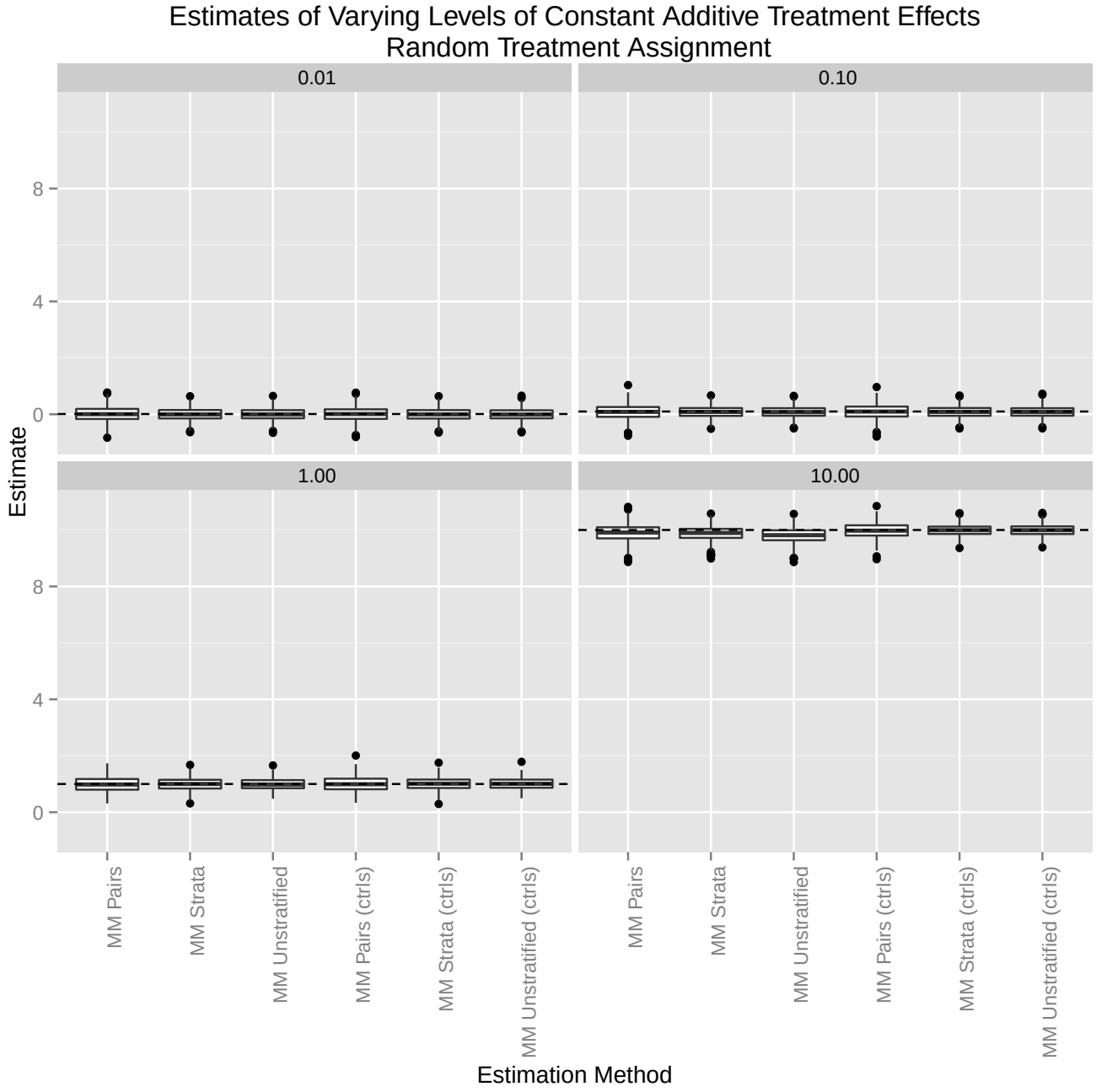
	0.01	0.1	1	10
MM Pairs	0.402	0.411	0.397	0.398
MM Strata	0.337	0.333	0.330	0.332
Pscore Pairs	1.688	1.752	1.831	1.721
Pscore Strata	0.729	0.767	0.775	0.682
Ebal	0.362	0.355	0.353	0.351
Unadjusted	2.263	2.211	2.205	2.153

Table 8: RMSE for various treatment effects; Treatment Assignment with $cov(T, X_1) = 0.5$; Laplacian Errors

2 Comparing Variants of Model-based Matching

One consideration is how we fit the model: should we fit it to controls only, fit it to all individuals, or something else?

First, consider estimation. From the three simulations below, it's clear that fitting to all individuals introduces bias in the estimate when treatment assignment is correlated with X. We ought to fit to controls only.

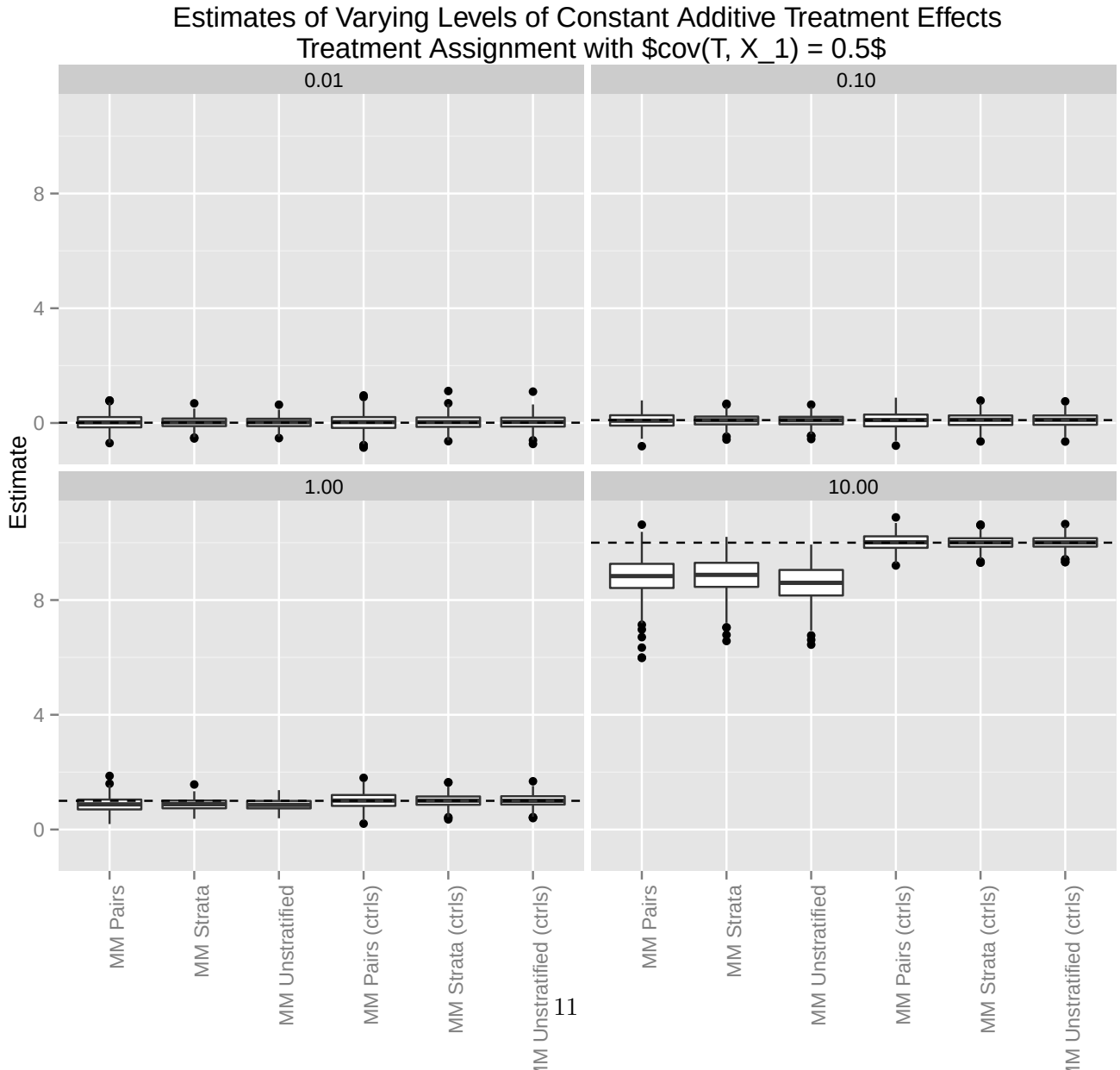


	0.01	0.1	1	10
MM Pairs	0.272	0.270	0.272	0.324
MM Strata	0.210	0.201	0.209	0.273
MM Unstratified	0.205	0.198	0.199	0.328
MM Pairs (ctrls)	0.268	0.270	0.267	0.268
MM Strata (ctrls)	0.214	0.205	0.210	0.204
MM Unstratified (ctrls)	0.211	0.204	0.204	0.203

Table 9: RMSE for various treatment effects; Random Treatment Assignment

	0.01	0.1	1	10
MM Pairs	-0.008	-0.026	-0.017	-0.109
MM Strata	-0.007	-0.018	0.001	-0.120
MM Unstratified	-0.006	-0.021	-0.008	-0.196
MM Pairs (ctrls)	-0.005	-0.016	-0.001	-0.021
MM Strata (ctrls)	-0.009	-0.016	0.011	-0.011
MM Unstratified (ctrls)	-0.007	-0.018	0.011	-0.010

Table 10: Bias for estimators; Random Treatment Assignment

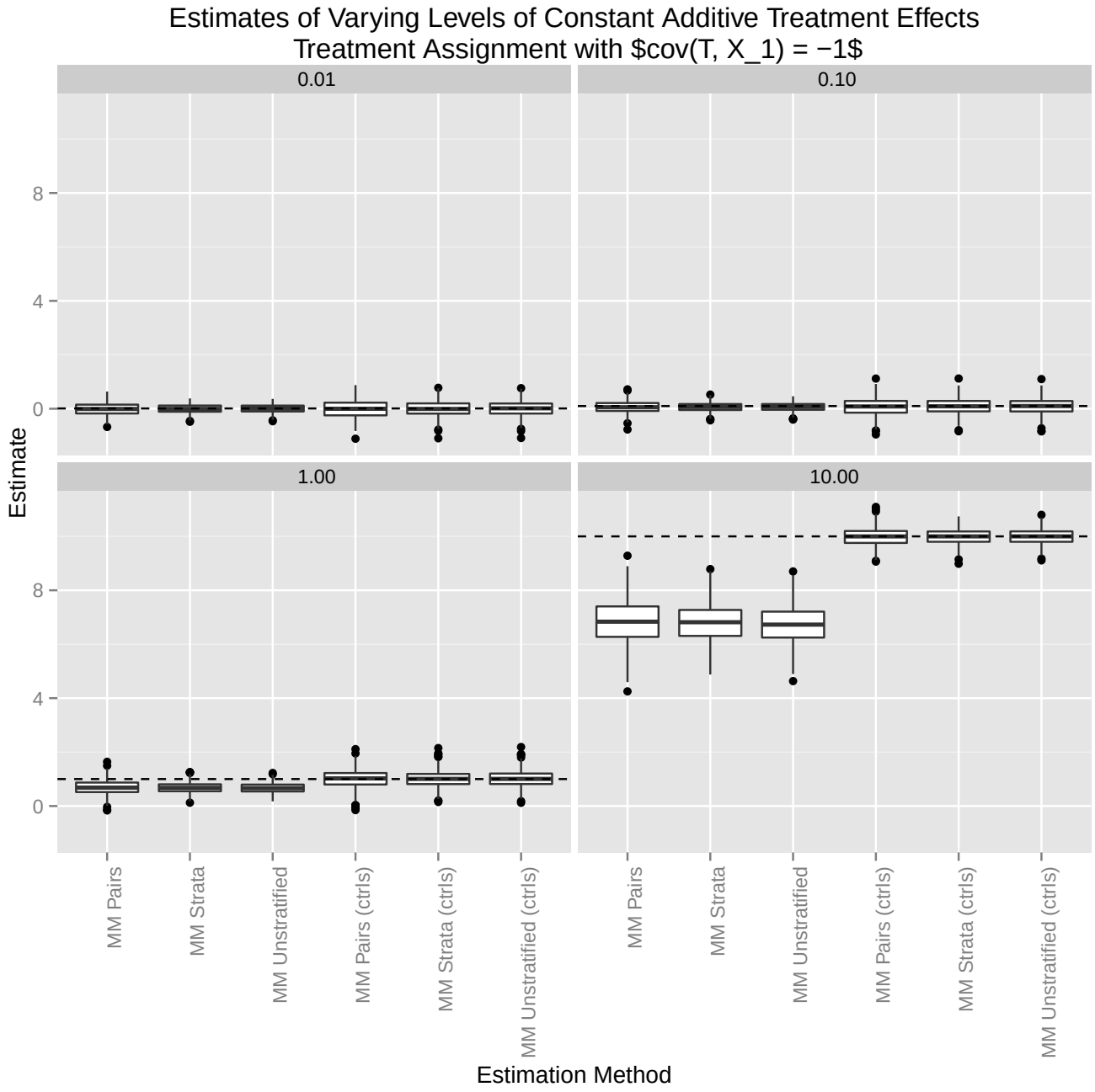


	0.01	0.1	1	10
MM Pairs	0.264	0.259	0.291	1.358
MM Strata	0.198	0.196	0.227	1.307
MM Unstratified	0.189	0.187	0.228	1.557
MM Pairs (ctrls)	0.299	0.283	0.267	0.278
MM Strata (ctrls)	0.245	0.234	0.218	0.220
MM Unstratified (ctrls)	0.241	0.231	0.216	0.218

Table 11: RMSE for various treatment effects; Treatment Assignment with $cov(T, X_1) = 0.5$

	0.01	0.1	1	10
MM Pairs	0.016	-0.015	-0.119	-1.185
MM Strata	0.011	-0.010	-0.123	-1.167
MM Unstratified	0.013	-0.010	-0.138	-1.424
MM Pairs (ctrls)	0.020	-0.005	0.012	0.012
MM Strata (ctrls)	0.013	0.000	0.010	0.000
MM Unstratified (ctrls)	0.017	0.002	0.011	0.001

Table 12: Bias for estimators; Treatment Assignment with $cov(T, X_1) = 0.5$



	0.01	0.1	1	10
MM Pairs	0.264	0.259	0.291	1.358
MM Strata	0.198	0.196	0.227	1.307
MM Unstratified	0.189	0.187	0.228	1.557
MM Pairs (ctrls)	0.299	0.283	0.267	0.278
MM Strata (ctrls)	0.245	0.234	0.218	0.220
MM Unstratified (ctrls)	0.241	0.231	0.216	0.218

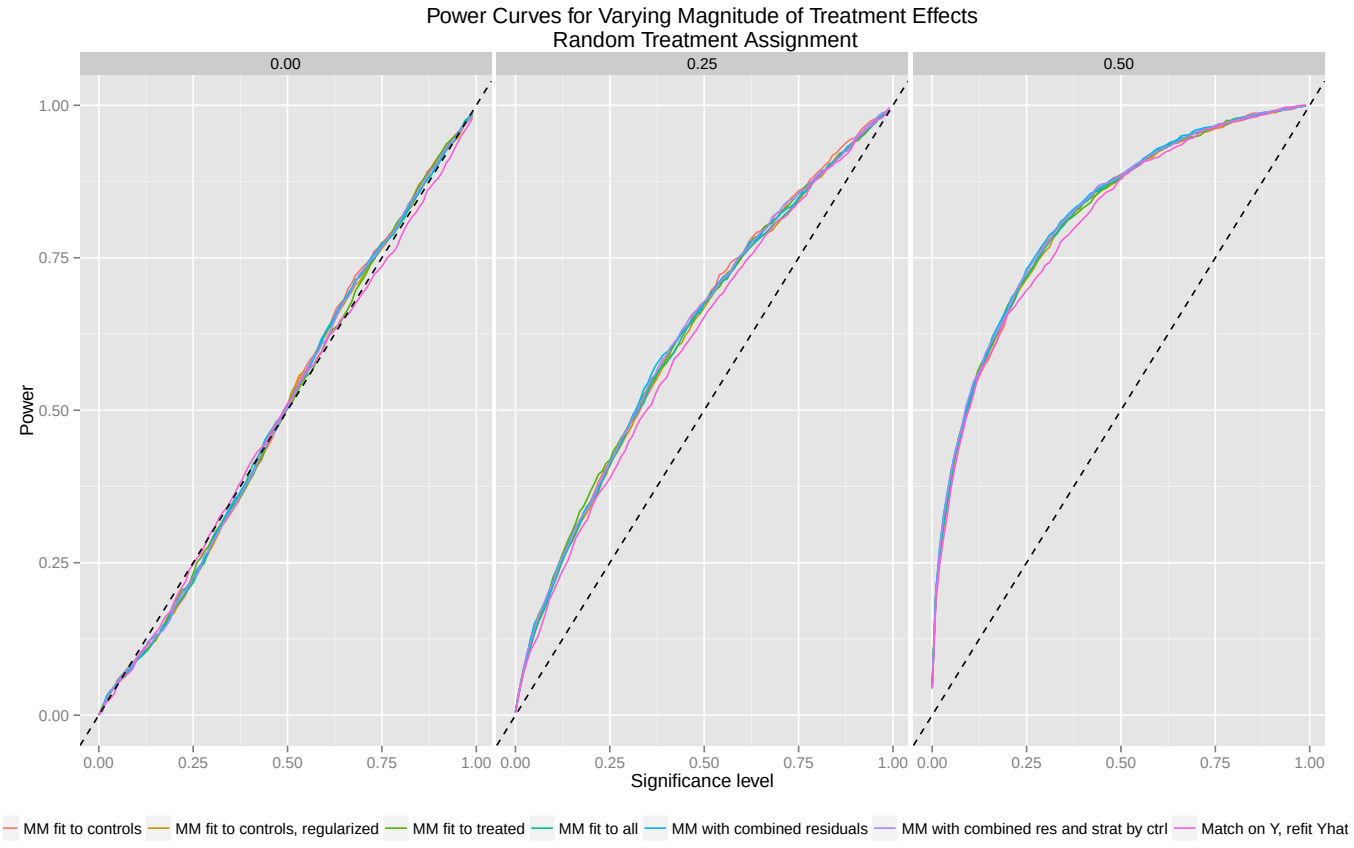
Table 13: RMSE for various treatment effects; Treatment Assignment with $cov(T, X_1) = -1$

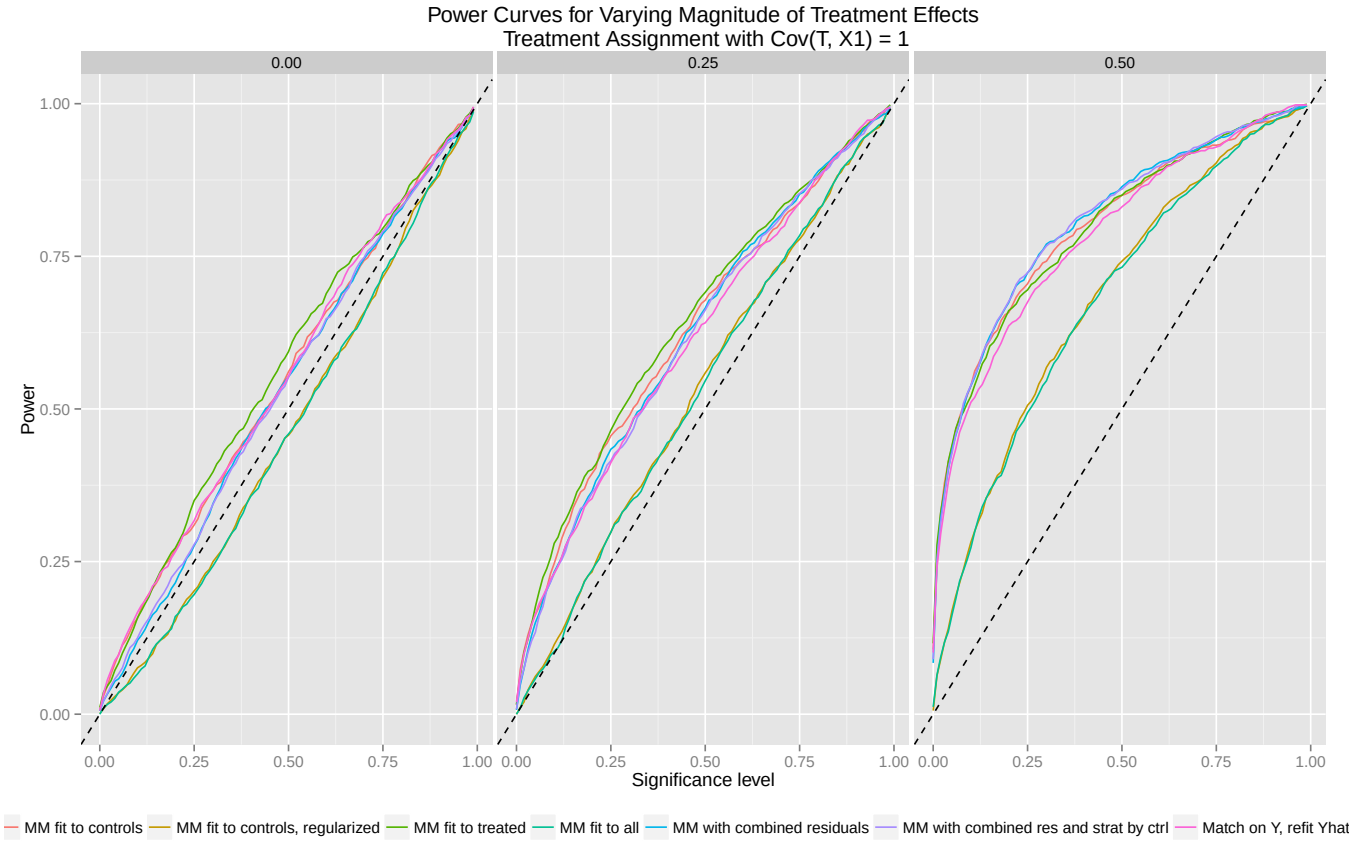
Next, consider testing. We fix the number of strata at 5 according to quintiles of $\hat{Y}$. We try several different methods: fitting $\hat{Y}$ to controls only and stratifying on it, fitting a regularized model to controls only and stratifying on the predictions, fitting $\hat{Y}$ to treated only and stratifying on it, fitting $\hat{Y}$ to all units

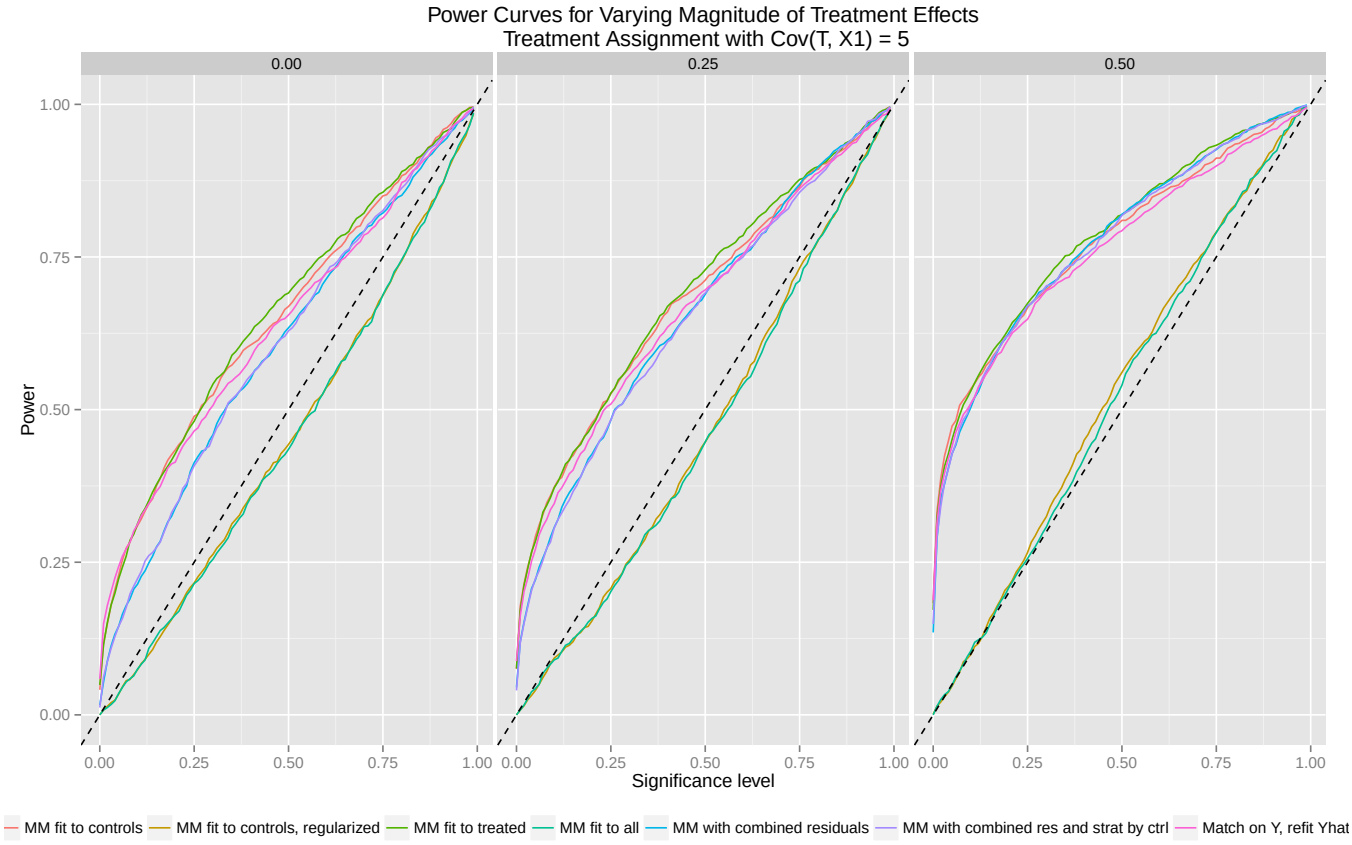
	0.01	0.1	1	10
MM Pairs	-0.014	-0.042	-0.304	-3.186
MM Strata	-0.009	-0.036	-0.315	-3.201
MM Unstratified	-0.008	-0.035	-0.325	-3.284
MM Pairs (ctrls)	-0.009	-0.014	0.016	-0.012
MM Strata (ctrls)	-0.009	-0.009	0.012	-0.009
MM Unstratified (ctrls)	-0.008	-0.008	0.016	-0.011

Table 14: Bias for estimators; Treatment Assignment with $cov(T, X_1) = -1$

and stratifying on it, fitting to treated and controls and using the average of the two predictions as our $\hat{Y}$ and our stratifying score, using this combined residual but stratifying based on the control prediction only, and stratifying on observed Y and refitting the model to controls at each permutation.
It appears that only two methods have the correct level: fitting a regularized model to controls and fitting to all observations. We've uncovered an interesting problem: if we our model to get $\hat{Y}$ systematically overfits to one group, we get anti-conservative tests. The actual level of our permutation tests is much higher than the nominal level.
Thus, we must be clear about whether our goal is testing or estimation.

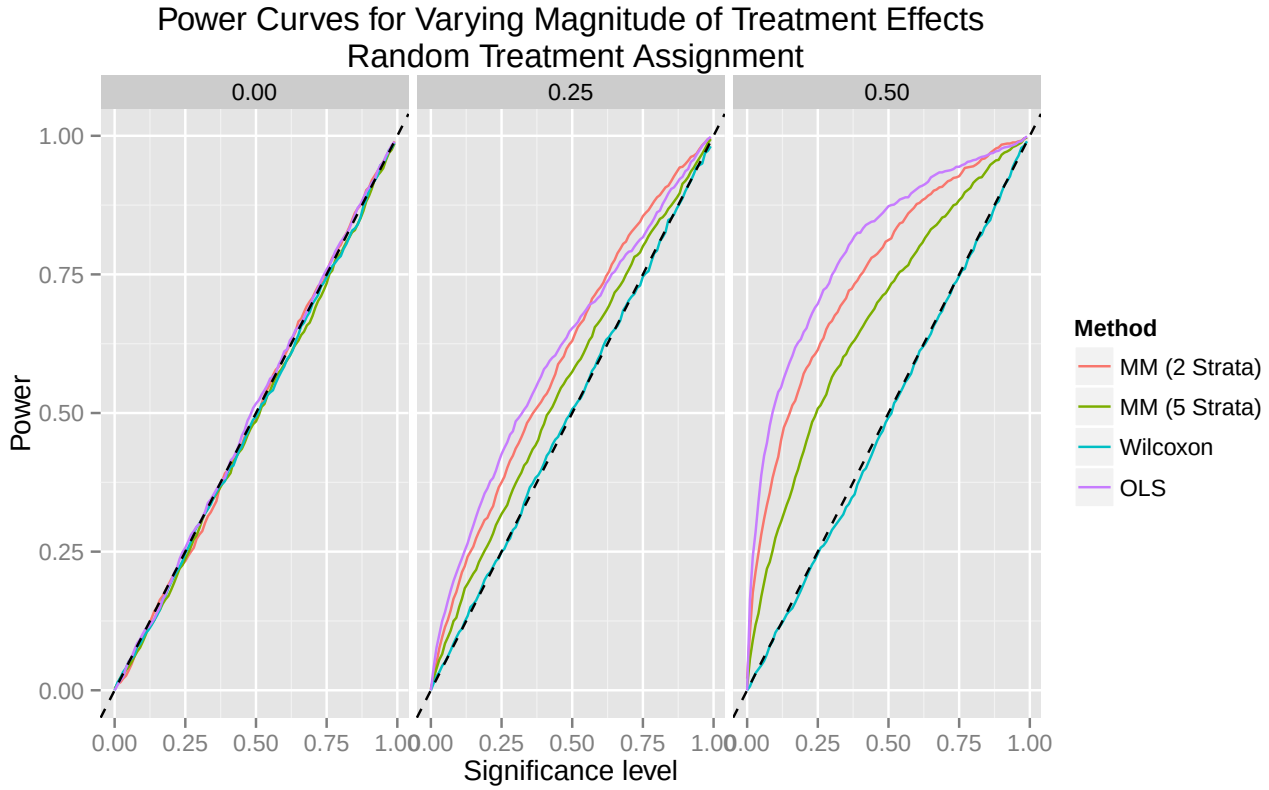


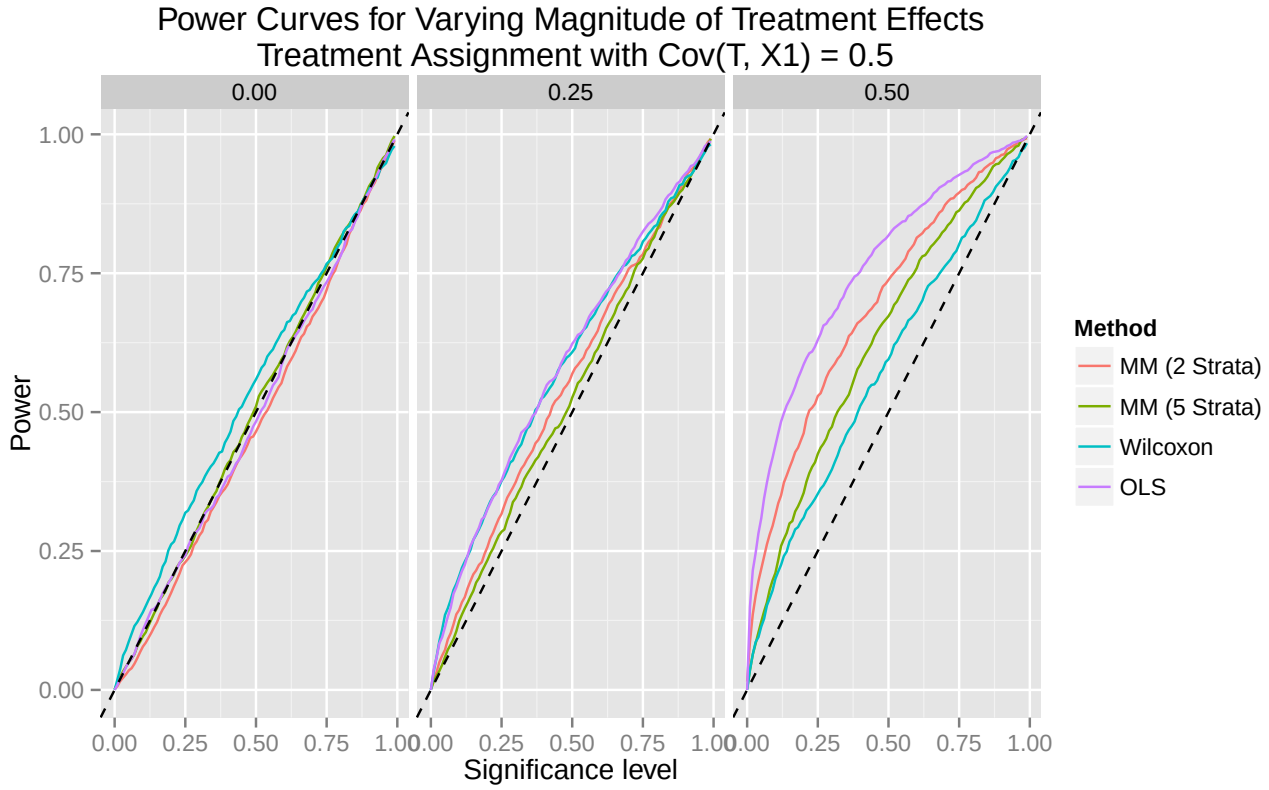

Power Curves for Varying Magnitude of Treatment Effects
Treatment Assignment with Cov(T, X1) = 1
0.00
0.25
0.50
Power
1.00
0.75
0.50
0.25
0.00
0.00
0.25
0.50
0.75
1.00
Significance level
MM fit to controls
MM fit to controls, regularized
MM fit to treated
MM fit to all
MM with combined residuals
MM with combined res and strat by ctrl
Match on Y, refit Yhat

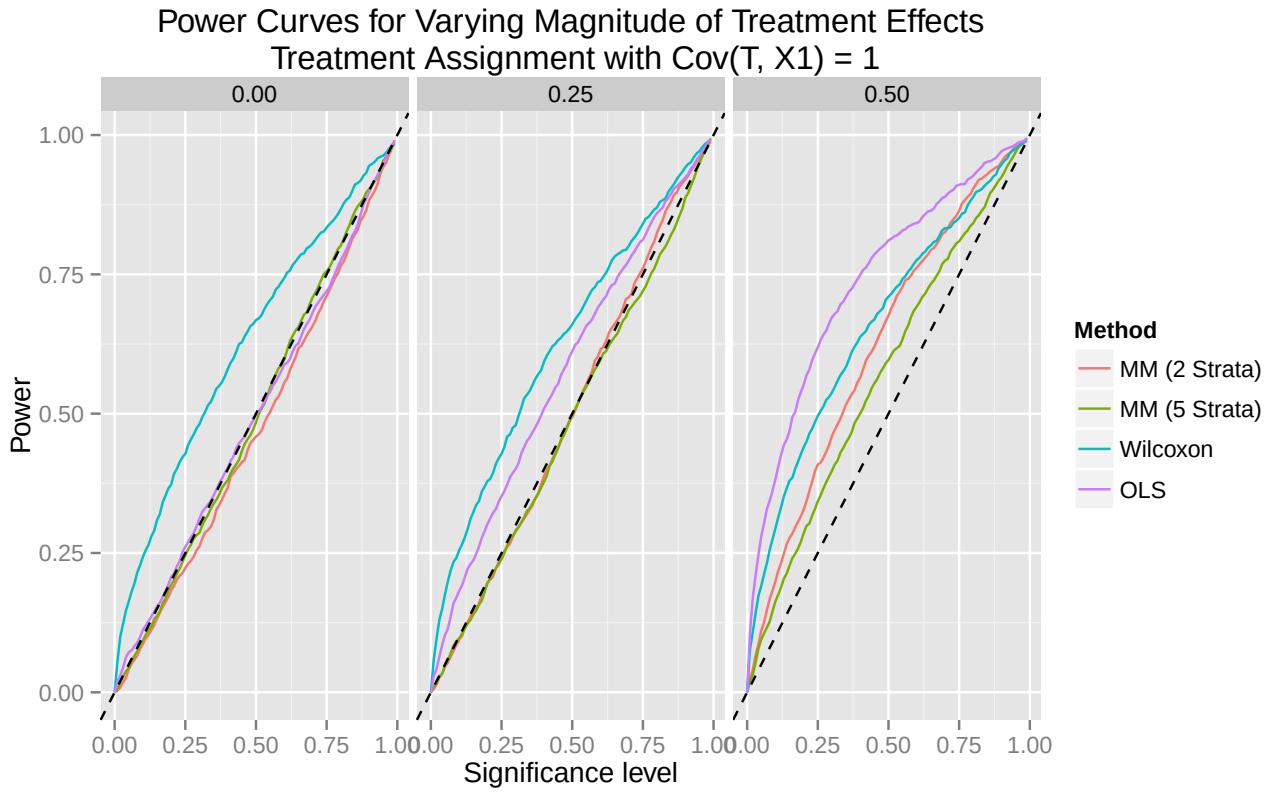

Power Curves for Varying Magnitude of Treatment Effects
Treatment Assignment with Cov(T, X1) = 5
0.00
0.25
0.50
Power
1.00
0.75
0.50
0.25
0.00
0.00
0.25
0.50
0.75
1.00
Significance level
MM fit to controls
MM fit to controls, regularized
MM fit to treated
MM fit to all
MM with combined residuals
MM with combined res and strat by ctrl
Match on Y, refit Yhat

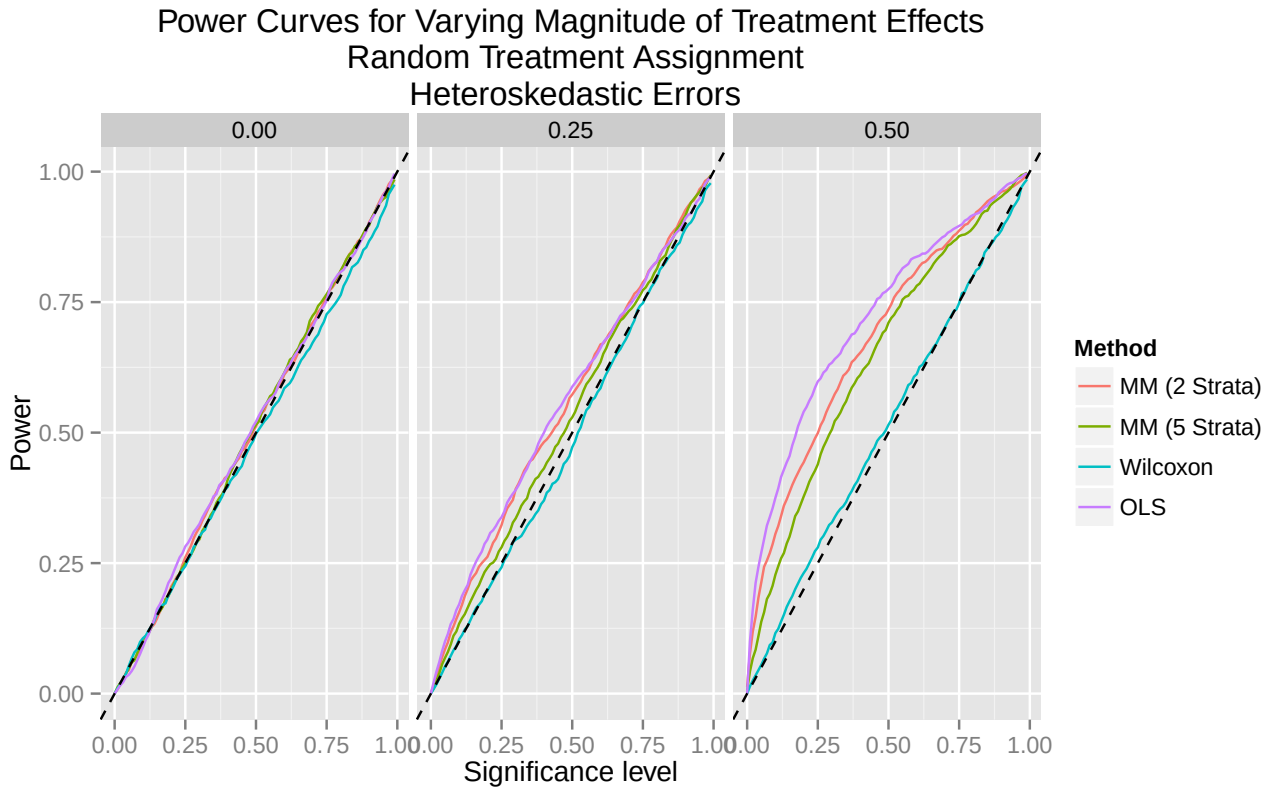
3 Hypothesis Testing

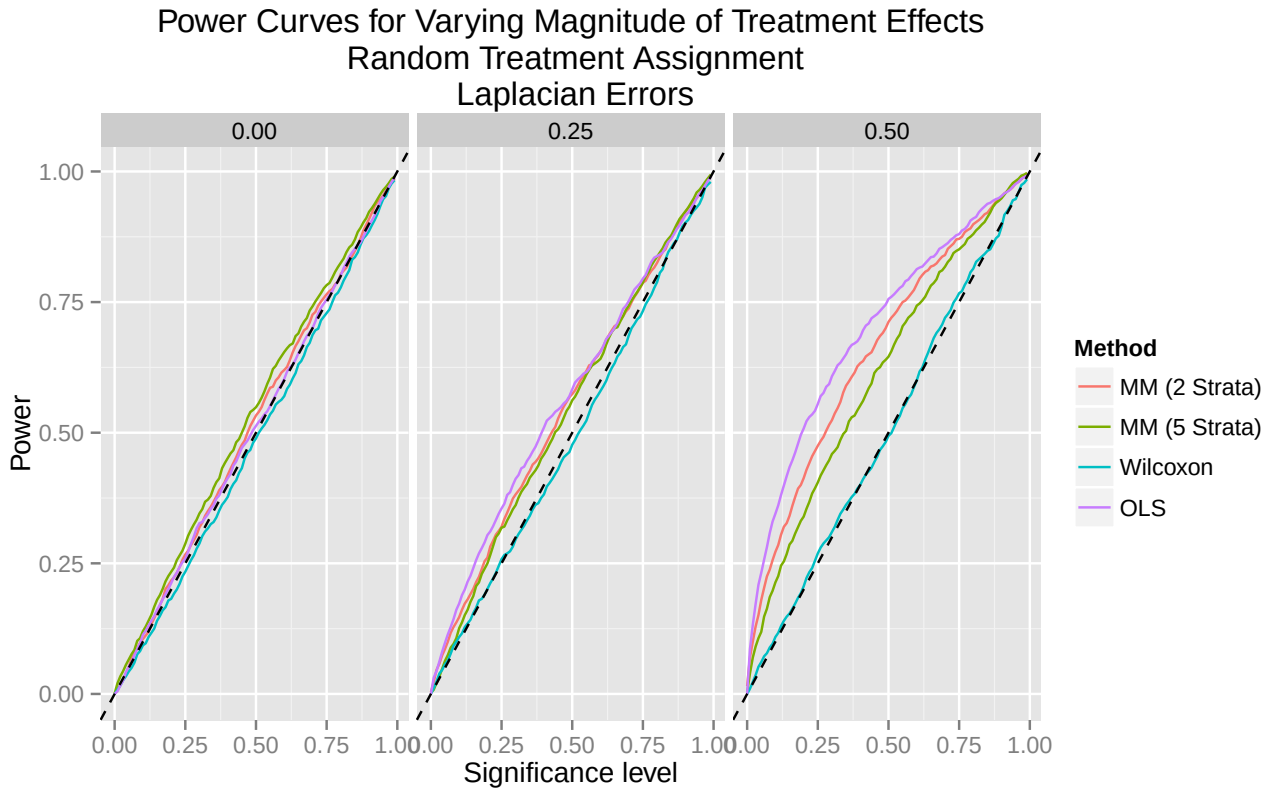
We compare the power of stratified permutation tests using model-based matching with different numbers of strata to the t-test from OLS and the Wilcoxon rank sum test.






Power Curves for Varying Magnitude of Treatment Effects
Treatment Assignment with Cov(T, X1) = 1
0.00
0.25
0.50
Power
1.00
0.75
0.50
0.25
0.00
0.00 0.25 0.50 0.75 1.00
Significance level
Method
MM (2 Strata)
MM (5 Strata)
Wilcoxon
OLS


Power Curves for Varying Magnitude of Treatment Effects
Random Treatment Assignment
Heteroskedastic Errors
0.00
0.25
0.50
Power
1.00
0.75
0.50
0.25
0.00
0.00 0.25 0.50 0.75 1.00
Significance level
Method
MM (2 Strata)
MM (5 Strata)
Wilcoxon
OLS


Power Curves for Varying Magnitude of Treatment Effects
Random Treatment Assignment
Laplacian Errors
0.00
0.25
0.50
Power
1.00
0.75
0.50
0.25
0.00
0.00
0.25
0.50
0.75
1.00
Significance level
Method
MM (2 Strata)
MM (5 Strata)
Wilcoxon
OLS