

Online Supplement

Table 1s

Distribution of Propensity Scores

Variable	Mean	<i>SD</i>	Minimum	Maximum
Non-Support Cases	0.16	0.25	0.004	1.00
Non-Support Range	0.18	0.21	0.02	1.00
Support Range	0.56	0.23	0.00	0.97

Table 2s

Mean Balance by Covariate Type, Sample Size, and Number of Covariates (k)

Type of covariate	N	k	Conditioning method					
			No		Ignore	PS ANCOVA	Weighting	Stratification
			Caliper Match	Caliper Match				
Binary	50	3	0.00	0.09	0.11	-0.02	-0.02	0.86
		9	0.00	0.22	0.22	-0.04	-0.06	1.03
		15	-0.08	0.18	0.14	-4.37E+06	-0.13	1.04
		30	0.10	0.47	0.16	1.68E+16	0.17	0.96
	100	3	0.00	0.10	0.12	-0.01	-0.01	0.71
		9	0.01	0.26	0.26	0.00	-0.04	0.82
		15	0.01	0.31	0.32	-0.01	-0.05	0.86
		30	-0.34	0.09	0.08	-7.10E+07	-0.23	0.91
	250	3	0.00	0.11	0.12	0.00	0.00	0.47
		9	0.01	0.29	0.29	0.00	-0.02	0.50
		15	0.01	0.36	0.36	0.00	-0.03	0.53
		30	0.02	0.43	0.43	-0.01	-0.05	0.61
	500	3	0.00	0.11	0.13	0.00	0.00	0.34
		9	0.01	0.30	0.30	0.00	-0.01	0.35
		15	0.01	0.38	0.38	0.00	-0.02	0.37
		30	0.02	0.47	0.47	0.00	-0.04	0.43
1000	3	0.00	0.12	0.13	0.00	0.00	0.24	
	9	0.01	0.31	0.31	0.00	-0.01	0.24	
	15	0.01	0.39	0.39	0.00	-0.02	0.26	
	30	0.02	0.49	0.49	0.00	-0.03	0.31	
Continuous	50	2	0.00	0.12	0.14	0.00	-0.03	0.60
		6	0.00	0.16	0.17	-0.01	-0.03	0.72
		10	0.00	0.19	0.19	-0.77	0.01	1.50
		20	-0.01	0.23	0.18	-913.06	0.42	1.90
	100	2	0.00	0.13	0.16	0.00	-0.02	0.37
		6	0.00	0.19	0.20	0.00	-0.03	0.40
		10	0.01	0.24	0.25	0.00	-0.03	0.44
		20	0.00	0.25	0.26	-0.99	0.05	0.80
	250	2	0.00	0.15	0.18	0.00	-0.01	0.22
		6	0.00	0.22	0.23	0.00	-0.02	0.23
		10	0.01	0.28	0.28	0.00	-0.02	0.25
		20	0.01	0.32	0.33	-0.01	-0.02	0.31
	500	2	0.00	0.16	0.18	0.00	-0.01	0.15
		6	0.01	0.24	0.24	0.00	-0.01	0.16
		10	0.01	0.30	0.30	0.00	-0.02	0.17
		20	0.01	0.35	0.35	0.00	-0.02	0.21
1000	2	0.00	0.17	0.19	0.00	0.00	0.11	
	6	0.01	0.24	0.24	0.00	-0.01	0.11	
	10	0.01	0.31	0.31	0.00	-0.01	0.12	
	20	0.01	0.37	0.37	0.00	-0.01	0.15	

Table 3s

Mean Balance by Covariate Type and Reliability (ρ_{xx})

Type of covariate	ρ_{xx}	Conditioning method					
		No			PS ANCOVA	Weighting	Stratification
		Caliper Match	Caliper Match	Ignore			
Binary	0.4	0.00	0.17	0.16	0.00	-0.01	0.58
	0.6	-0.02	0.22	0.22	0.00	-0.03	0.59
	0.8	-0.02	0.29	0.28	0.00	-0.05	0.59
	1.0	0.00	0.41	0.39	0.00	-0.03	0.60
Continuous	0.4	0.00	0.18	0.18	0.00	0.01	0.48
	0.6	0.00	0.22	0.22	0.00	0.00	0.42
	0.8	0.00	0.25	0.26	0.00	0.01	0.42
	1	0.01	0.28	0.29	0.00	0.00	0.43

Note. $N = 50$ and $N = 100$ were removed for PS ANCOVA.

Table 4s

Mean Balance by Covariate Type and Correlation among the Covariates (ρ_{12})

Type of covariate	ρ_{12}	Conditioning method					
		Caliper Match	No Caliper Match	Ignore	PS ANCOVA	Weighting	Stratification
Binary	0.0	0.00	0.06	0.06	0.00	-0.01	0.57
	0.2	-0.01	0.26	0.26	0.00	-0.02	0.59
	0.5	-0.02	0.50	0.47	0.00	-0.06	0.61
Continuous	0.0	0.00	0.10	0.10	0.00	-0.01	0.39
	0.2	0.00	0.22	0.23	0.00	0.01	0.43
	0.5	0.01	0.38	0.39	0.00	0.02	0.50

Note. $N = 50$ and $N = 100$ were removed for PS ANCOVA.

Table 5s

Mean Balance by Trimming and Sample Size

Type of covariate	Trim	N	Conditioning method					
			Caliper Match	No Caliper Match	Ignore	PS ANCOVA	Weighting	Stratification
Binary	No	50	-0.03	0.37	0.38	-604.46	0.20	0.90
		100	-0.04	0.36	0.37	-0.76	0.08	0.78
		250	0.01	0.36	0.36	0.00	0.02	0.52
		500	0.01	0.36	0.36	0.00	0.01	0.37
		1000	0.01	0.36	0.36	0.00	0.01	0.27
	Yes	50	0.04	0.09	-0.07	8.09E+15	-0.23	1.05
		100	-0.12	0.02	0.02	-3.55E+07	-0.24	0.86
		250	0.01	0.24	0.24	0.00	-0.07	0.53
		500	0.01	0.28	0.28	0.00	-0.05	0.37
		1000	0.01	0.30	0.30	0.00	-0.03	0.26
Continuous	No	50	0.00	0.30	0.31	-426.28	0.15	0.58
		100	0.01	0.29	0.30	-0.48	0.07	0.41
		250	0.01	0.30	0.30	0.00	0.02	0.25
		500	0.01	0.30	0.30	0.00	0.01	0.18
		1000	0.01	0.30	0.30	0.00	0.01	0.13
	Yes	50	-0.01	0.05	0.03	0.01	0.02	1.77
		100	0.00	0.11	0.13	-0.02	-0.08	0.59
		250	0.01	0.19	0.20	0.00	-0.06	0.25
		500	0.01	0.23	0.24	0.00	-0.04	0.17
		1000	0.01	0.25	0.26	0.00	-0.03	0.12

Table 6s

Mean Bias of Treatment Effect by the Number of Covariates (k) and Reliability (ρ_{xx})

k	ρ_{xx}	Conditioning method						
		Caliper Match	No Caliper Match	Ignore	ANCOVA	PS ANCOVA	Weighting	Stratification
3	.40	0.05	0.08	0.09	0.05	0.05	0.05	0.05
	.60	0.03	0.08	0.08	0.03	0.03	0.03	0.04
	.80	0.02	0.07	0.08	0.02	0.02	0.01	0.02
	1.0	0.00	0.07	0.08	0.00	0.00	-0.01	0.01
9	.40	0.15	0.30	0.30	0.15	0.14	0.13	0.16
	.60	0.09	0.29	0.30	0.09	0.09	0.07	0.10
	.80	0.05	0.28	0.29	0.04	0.04	0.02	0.06
	1.0	0.01	0.27	0.28	0.00	0.00	-0.02	0.02
15	.40	0.28	0.63	0.64	0.27	0.27	0.24	0.29
	.60	0.17	0.61	0.61	0.14	0.15	0.13	0.19
	.80	0.08	0.59	0.60	0.08	0.07	0.04	0.11
	1.0	0.02	0.57	0.58	0.00	0.00	-0.03	0.04
30	.40	0.53	1.45	1.47	0.51	0.53	0.65	0.64
	.60	0.32	1.38	1.39	0.26	0.29	0.43	0.44
	.80	0.16	1.33	1.34	0.13	0.13	0.29	0.30
	1.0	0.05	1.30	1.31	0.05	-0.01	0.17	0.15

Note. $N = 50$ and $N = 100$ were removed for PS ANCOVA. The absolute mean bias less than .05 is in

bold.

Table 7s

*The Eta-squared of Simulation Design Factors and the First-order Interaction Effects on Bias**Estimates*

Source	Conditioning method						
	Caliper Match	No Caliper Match	Ignore	ANCOVA	Propensity score ANCOVA	Weighting	Stratification
k	.08	.27	.27	.	.27	.07	.16
β_{yc}	.	.11	.11	.	.11	.	.07
$k * \beta_{yc}$.	.08	.08	.	.08	.	.
ρ_{12}	.	.10	.10	.	.10	.	.
$k * \rho_{12}$.	.09	.09	.	.09	.	.
ρ_{xx}	.0806
$k * \text{trimming}$09	.

Note. When $\eta^2 > .06$, η^2 is reported. k = number of covariates, β_{yc} = strength of relationship between

covariates and the outcome, ρ_{12} = strength of relationship between covariates, ρ_{xx} = reliability.

Table 8s

Mean RMSE by Study Factors and PS Conditioning Methods

Design Factor	Conditioning Methods						
	Caliper Match	No Caliper Match	Ignore	ANCOVA	Propensity Score ANCOVA	Weighting	Stratification
<i>N</i>							
50	1.80	1.96	1.74	207.21	891347651.2	1.80	1.56
100	0.57	1.35	1.29	7.62	7674.69	1.00	0.49
250	0.14	1.28	1.28	0.18	0.18	0.33	0.14
500	0.10	1.37	1.37	0.08	0.08	0.20	0.11
1000	0.08	1.45	1.45	0.06	0.07	0.14	0.09
Trim							
No	0.51	1.92	1.91	0.21	0.09	0.78	0.48
Yes	0.55	1.04	0.94	83.84	0.13	0.59	0.46
ρ_{xx}							
.40	0.66	1.54	1.49	5.10	0.19	0.69	0.57
.60	0.55	1.49	1.43	95.76	0.081	0.66	0.48
.80	0.48	1.47	1.41	60.39	0.043	0.68	0.45
1.0	0.45	1.42	1.37	5.50	0.13	0.72	0.40
<i>N</i> of Covariates							
3	0.01	0.02	0.02	0.01	0.00	0.01	0.01
9	0.06	0.20	0.20	0.05	0.02	0.05	0.06
15	0.20	0.83	0.82	153.29	0.06	0.25	0.23
30	1.89	4.91	4.71	12.41	0.36	2.46	1.61
β_{xc}							
0.025	0.42	0.83	0.78	2.09	0.03	0.38	0.35
0.050	0.56	1.53	1.47	42.44	0.07	0.67	0.49
0.100	0.63	2.08	2.03	81.21	0.23	1.02	0.59
β_{yc}							
0.025	0.09	0.22	0.21	1.12	0.01	0.11	0.08
0.050	0.32	0.86	0.82	78.43	0.12	0.40	0.27
0.100	1.19	3.36	3.24	45.49	0.20	1.55	1.07
ρ_{12}							
0	0.12	0.18	0.17	5.39	0.03	0.10	0.11
0.2	0.52	1.20	1.15	38.08	0.09	0.55	0.43
0.5	0.97	3.06	2.97	82.06	0.21	1.42	0.89

Note. Estimates are based on 5,000 samples of each condition. ρ_{xx} = reliability, β_{xc} = strength of relationship between covariates and treatment assignment, β_{yc} = strength of relationship between covariates and the outcome, ρ_{12} = strength of relationship between covariates.

Table 9s

Mean Type I Error Rates by Sample Size and Covariate Reliability (ρ_{xx})

N	ρ_{xx}	Analysis Method						
		Caliper Match	No Caliper Match	Ignore	ANCOVA	Propensity Score ANCOVA	Weighting	Stratification
50	.40	.09	.22	.25	.11	.08	.21	.16
	.60	.08	.22	.24	.09	.06	.20	.14
	.80	.07	.21	.23	.08	.05	.19	.12
	1.00	.07	.21	.22	.07	.05	.19	.12
100	.40	.11	.37	.41	.17	.12	.22	.17
	.60	.07	.35	.38	.10	.07	.17	.11
	.80	.05	.34	.35	.06	.04	.15	.07
	1.00	.04	.33	.34	.05	.04	.15	.05
250	.40	.28	.66	.68	.36	.31	.31	.36
	.60	.14	.65	.67	.19	.15	.18	.20
	.80	.06	.64	.65	.09	.05	.12	.09
	1.00	.03	.63	.64	.05	.03	.12	.04
500	.40	.47	.80	.81	.53	.50	.45	.56
	.60	.26	.80	.80	.31	.28	.25	.36
	.80	.11	.80	.80	.13	.10	.13	.17
	1.00	.04	.80	.80	.05	.03	.11	.05
1000	.40	.66	.89	.89	.70	.68	.62	.73
	.60	.44	.90	.90	.48	.45	.37	.56
	.80	.19	.90	.89	.20	.18	.17	.31
	1.00	.04	.90	.89	.05	.04	.11	.08

Note. Estimates are based on 5,000 samples of each condition.

Table 10s

Mean Type I Error Rates by Sample Size and Strength of Relationship between Treatment Assignment and Covariates (β_{xc})

<i>N</i>	β_{xc}	Analysis Method						
		Caliper Match	No Caliper Match	Ignore	ANCOVA	Propensity Score ANCOVA	Weighting	Stratification
50	0.025	.04	.10	.11	.08	.04	.12	.11
	0.050	.08	.20	.22	.08	.06	.19	.13
	0.100	.12	.35	.37	.10	.09	.28	.16
100	0.025	.03	.17	.19	.06	.02	.08	.05
	0.050	.05	.33	.36	.08	.05	.15	.09
	0.100	.12	.54	.57	.15	.13	.28	.17
250	0.025	.04	.39	.40	.08	.04	.05	.06
	0.050	.10	.67	.68	.14	.11	.15	.15
	0.100	.24	.88	.89	.29	.26	.35	.32
500	0.025	.07	.58	.57	.11	.07	.06	.10
	0.050	.19	.84	.85	.23	.20	.20	.27
	0.100	.40	.98	.99	.43	.41	.45	.49
1000	0.025	.13	.74	.73	.16	.13	.09	.19
	0.050	.33	.95	.95	.36	.33	.29	.43
	0.100	.54	1.00	1.00	.55	.54	.56	.64

Note. Estimates are based on 5,000 samples of each condition.

Table 11s

Mean Type I Error Rates by Number of Covariates (k)

<i>k</i>	Analysis Method						
	Caliper Match	No Caliper Match	Ignore	ANCOVA	Propensity Score ANCOVA	Weighting	Stratification
3	.11	.37	.38	.14	.11	.09	.15
9	.15	.57	.58	.19	.16	.15	.21
15	.18	.66	.67	.21	.18	.23	.25
30	.22	.73	.75	.24	.21	.41	.27

Note. Estimates are based on 5,000 samples of each condition

Table 12s

Proportion of Conditions with Adequate Type I Error Control (Bradley Criterion) by Research Design

Factors

Design Factor	Analysis Method						
	Caliper Match	No Caliper Match	Ignore	ANCOVA	Propensity Score ANCOVA	Weighting	Stratification
<i>N</i>							
50	.64	.33	.30	.67	.42	.31	.20
100	.60	.11	.11	.60	.43	.33	.53
250	.48	.00	.00	.46	.43	.24	.41
500	.40	.00	.00	.36	.39	.22	.28
1000	.31	.00	.00	.29	.33	.15	.15
<i>Trim</i>							
No	.49	.03	.03	.50	.39	.21	.33
Yes	.48	.14	.14	.45	.41	.29	.29
<i>ρ_{xx}</i>							
.40	.24	.07	.06	.12	.24	.17	.07
.60	.42	.09	.08	.26	.38	.33	.19
.80	.57	.09	.09	.55	.46	.31	.37
1.00	.71	.10	.10	.96	.52	.20	.63
<i>k</i>							
3	.65	.15	.17	.58	.53	.40	.38
9	.49	.09	.08	.49	.36	.34	.32
15	.48	.07	.07	.44	.36	.19	.28
30	.33	.03	.01	.38	.34	.08	.26
<i>β_{xc}</i>							
0.025	.61	.17	.16	.67	.43	.36	.52
0.050	.51	.07	.07	.45	.42	.27	.29
0.100	.34	.01	.02	.31	.34	.13	.13
<i>β_{yc}</i>							
0.025	.49	.09	.08	.47	.40	.25	.31
0.050	.48	.08	.08	.47	.41	.25	.31
0.100	.49	.09	.08	.48	.40	.25	.32
<i>ρ_{12}</i>							
.00	.46	.12	.10	.44	.37	.26	.31
.20	.49	.08	.08	.47	.40	.26	.31
.50	.51	.06	.07	.52	.43	.23	.32

Note. Estimates are based on 5,000 samples of each condition. k = number of covariates, ρ_{xx} = reliability, β_{xc} = strength of relationship between covariates and treatment assignment, β_{yc} = strength of relationship between covariates and the outcome, ρ_{12} = strength of relationship between covariates.

Table 13s

Mean Confidence Interval Coverage Estimates by Sample Size and Covariate Reliability(ρ_{xx})

<i>N</i>	ρ_{xx}	Analysis Method						
		Caliper Match	No Caliper Match	Ignore	ANCOVA	Propensity Score ANCOVA	Weighting	Stratification
50	.40	.90	.79	.78	.89	.93	.79	.84
	.60	.91	.79	.78	.92	.94	.80	.86
	.80	.92	.80	.79	.92	.95	.81	.87
	1.00	.92	.80	.80	.93	.95	.82	.88
100	.40	.90	.66	.63	.86	.89	.79	.84
	.60	.94	.68	.66	.91	.94	.83	.90
	.80	.95	.70	.68	.94	.96	.85	.93
	1.00	.96	.71	.69	.95	.96	.85	.94
250	.40	.76	.39	.37	.70	.73	.72	.69
	.60	.88	.40	.38	.84	.87	.83	.82
	.80	.94	.42	.40	.92	.95	.88	.91
	1.00	.96	.43	.42	.95	.97	.88	.96
500	.40	.59	.25	.24	.53	.56	.60	.50
	.60	.78	.25	.24	.74	.77	.77	.69
	.80	.91	.25	.24	.89	.91	.87	.85
	1.00	.96	.25	.25	.95	.97	.89	.95
1000	.40	.40	.15	.14	.36	.38	.43	.33
	.60	.62	.14	.14	.59	.61	.66	.51
	.80	.84	.14	.15	.83	.85	.84	.74
	1.00	.95	.14	.15	.95	.96	.89	.92

Note. Estimates are based on 5,000 samples of each condition.

Table 14s

Mean Confidence Interval Coverage Estimates by Strength of Relationship between Covariates (β_{xc}) and Treatment Assignment and Covariate Reliability (ρ_{xx})

β_{xc}	ρ_{xx}	Analysis Method						
		Caliper Match	No Caliper Match	Ignore	ANCOVA	Propensity Score ANCOVA	Weighting	Stratification
0.025	.40	.88	.63	.62	.85	.88	.87	.83
	.60	.93	.63	.62	.91	.94	.92	.89
	.80	.95	.63	.63	.94	.97	.94	.93
	1.00	.96	.63	.63	.94	.97	.95	.95
0.050	.40	.72	.44	.42	.68	.71	.69	.65
	.60	.85	.44	.43	.82	.85	.81	.78
	.80	.92	.44	.44	.91	.94	.87	.88
	1.00	.95	.45	.45	.95	.96	.88	.93
0.100	.40	.52	.28	.26	.47	.50	.45	.44
	.60	.69	.29	.27	.66	.68	.61	.60
	.80	.86	.30	.29	.86	.87	.74	.78
	1.00	.94	.32	.30	.95	.95	.76	.91

Note. Estimates are based on 5,000 samples of each condition.

Table 15s

Mean Confidence Interval Width by Sample Size (N) and Number of Covariates (k)

<i>N</i>	<i>k</i>	Analysis Method						
		Caliper Match	No Caliper Match	Ignore	ANCOVA	Propensity Score ANCOVA	Weighting	Stratification
50	3	0.73	0.64	0.61	0.54	0.60	0.60	0.56
	9	1.54	1.25	1.13	1.05	1.22	1.10	1.14
	15	3.29	2.43	2.00	2.24	6.12	1.86	2.20
	30	11.16	6.35	5.90	5.79	18407.81	5.18	8.06
100	3	0.45	0.42	0.42	0.37	0.40	0.41	0.39
	9	0.88	0.77	0.74	0.66	0.75	0.74	0.74
	15	1.48	1.22	1.14	1.08	1.25	1.13	1.21
	30	5.06	3.44	2.96	2.95	9.65	2.73	3.35
250	3	0.26	0.25	0.26	0.23	0.24	0.26	0.24
	9	0.49	0.46	0.46	0.40	0.44	0.46	0.44
	15	0.79	0.70	0.69	0.63	0.70	0.70	0.69
	30	1.66	1.33	1.24	1.27	1.46	1.23	1.42
500	3	0.18	0.17	0.18	0.16	0.17	0.18	0.17
	9	0.33	0.32	0.32	0.28	0.30	0.32	0.30
	15	0.52	0.48	0.48	0.44	0.47	0.49	0.47
	30	1.06	0.89	0.84	0.85	0.94	0.85	0.92
1000	3	0.12	0.12	0.13	0.11	0.12	0.13	0.12
	9	0.22	0.22	0.23	0.20	0.20	0.23	0.21
	15	0.36	0.34	0.34	0.30	0.32	0.35	0.32
	30	0.71	0.61	0.58	0.59	0.63	0.60	0.62

Note. Estimates are based on 5,000 samples of each condition.

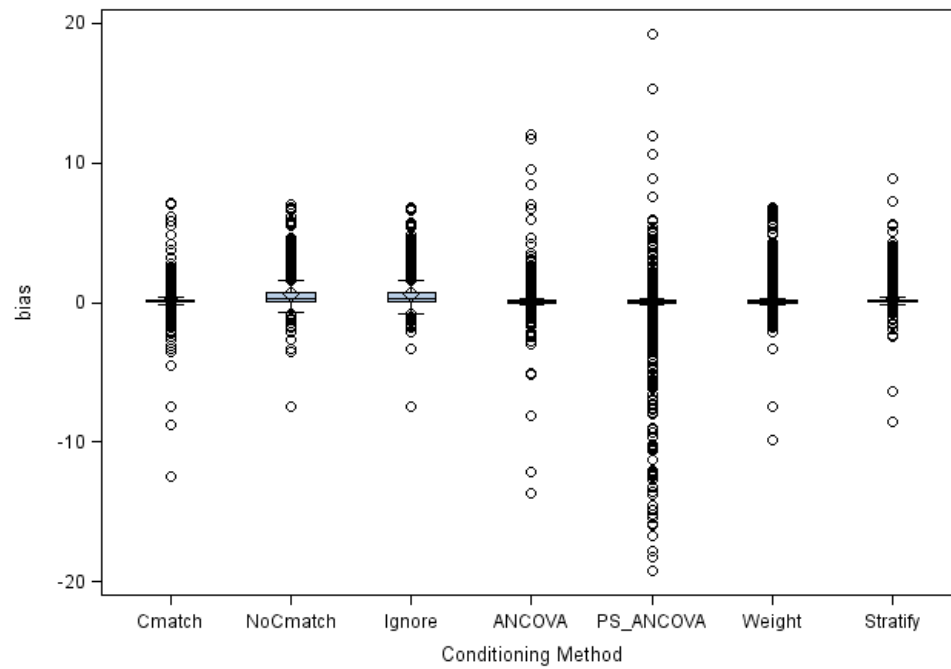
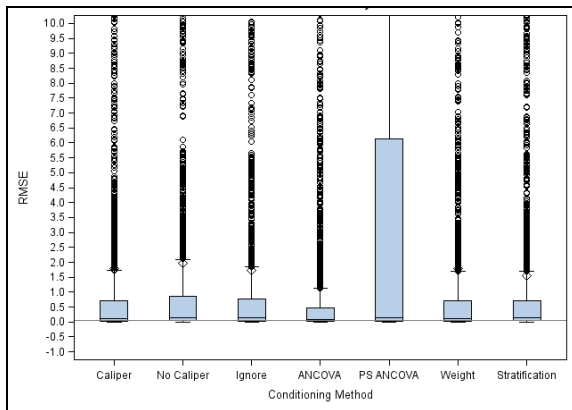
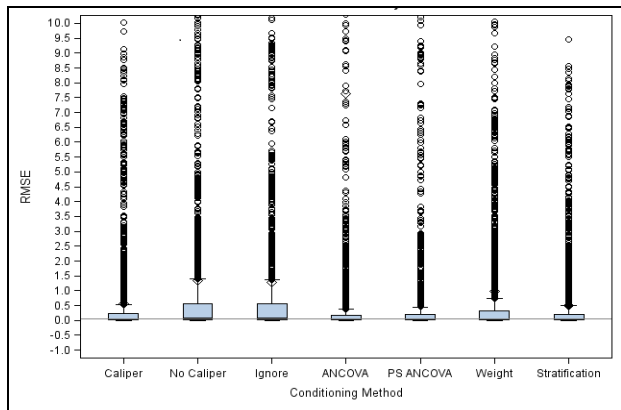


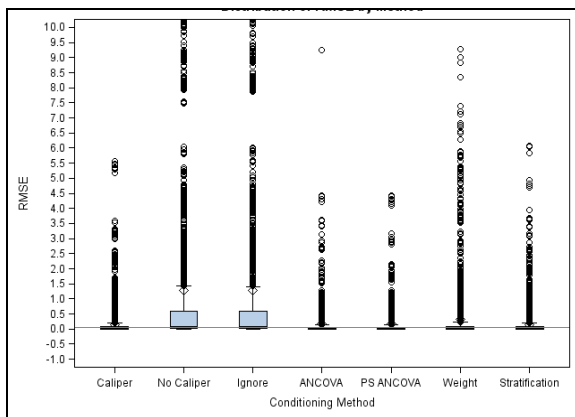
Figure 1s. Distributions of estimated bias across all simulation conditions. Cmatch = caliper matching, NoCmatch = no caliper matching, Ignore = ignoring the covariates.



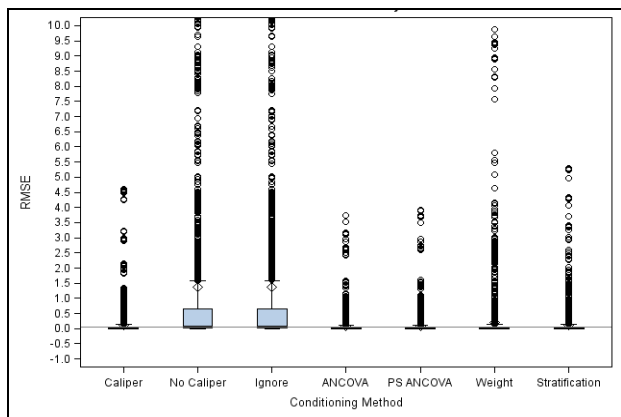
N=50



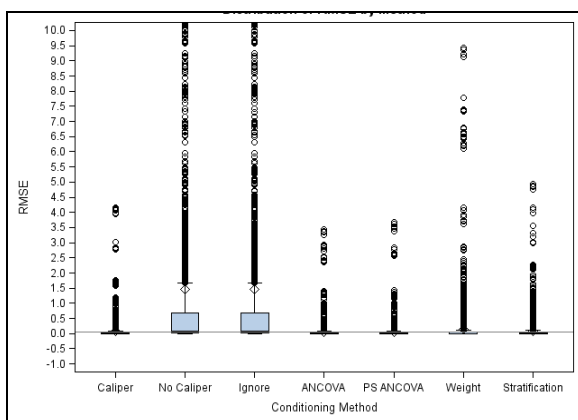
N=100



N=250



N=500



N=1000

Figure 2s. Mean RMSE distributions by method and sample size.

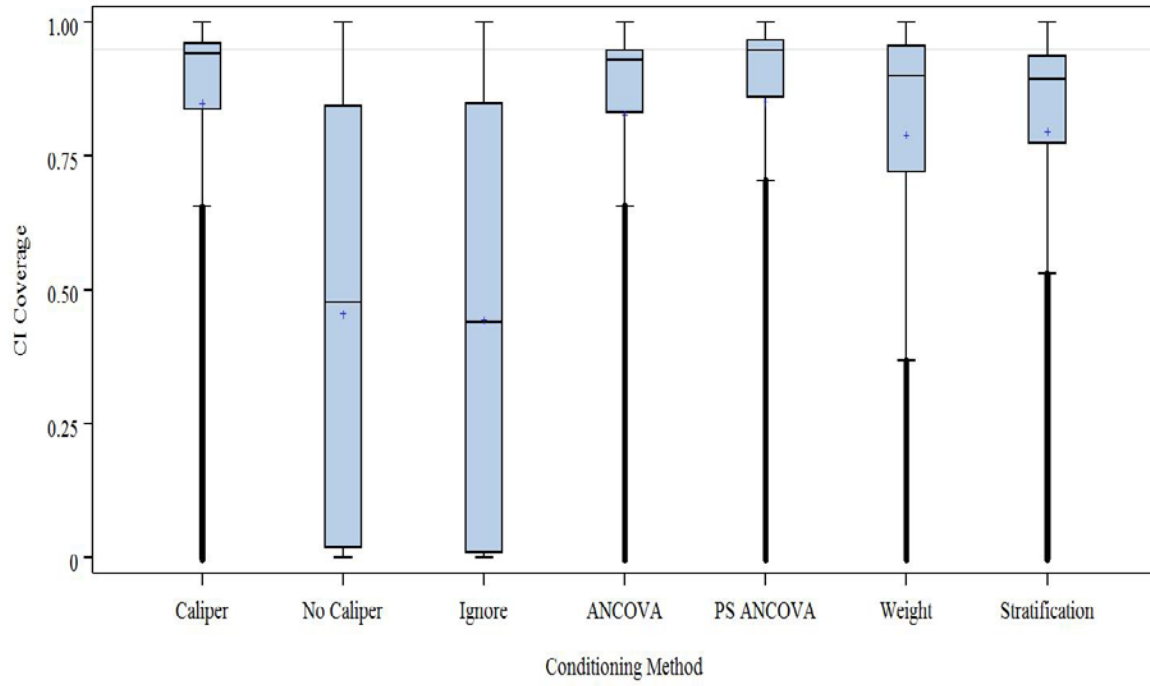


Figure 3s. Distributions of estimated confidence interval coverage across all conditions.

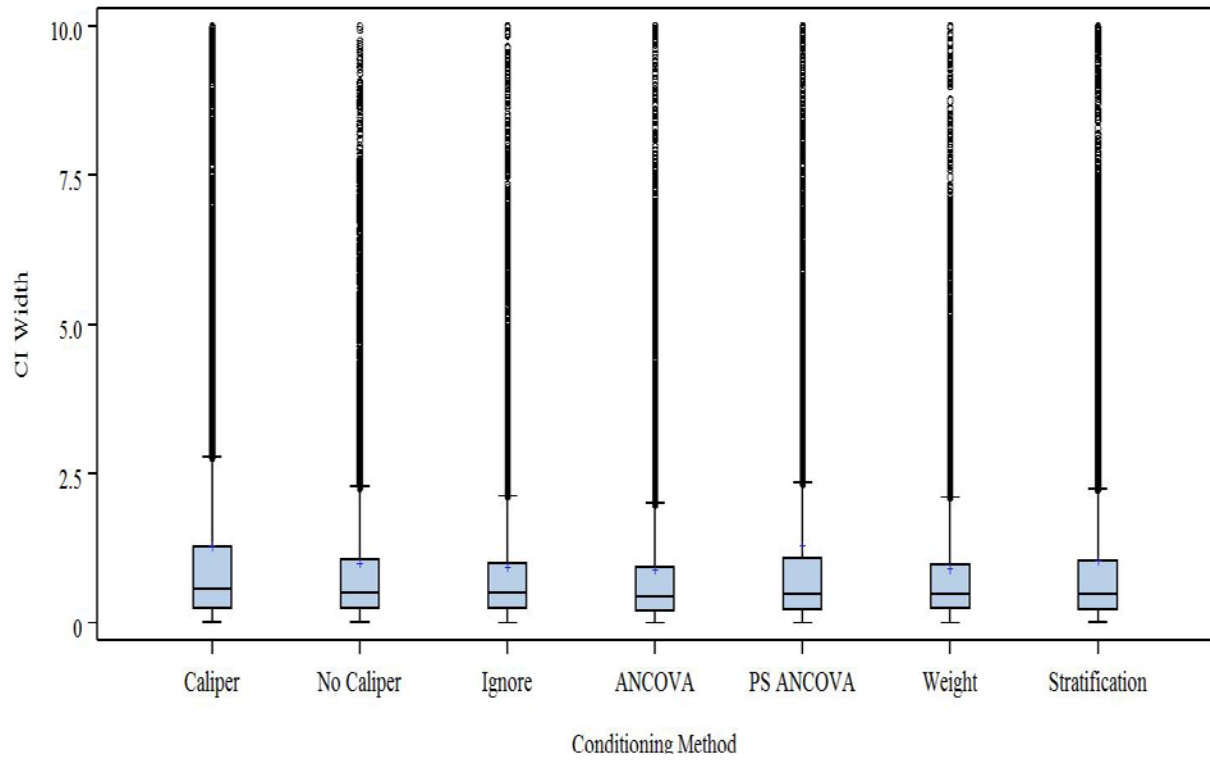


Figure 4s. Distributions of estimated confidence interval width across all conditions.