Online-Appendix: Entropy Balancing for Causal Effects: A Multivariate Reweighting Method to Produce Balanced Samples in Observational Studies

June 2011

Abstract

This online appendix presents various additional results referenced in the main paper. In particular, it contains two additional simulations studies and two additional empirical applications.

I. Monte Carlo Simulations

In addition to the simulation experiment presented in the main paper we conducted two smaller simulation experiments which we describe in the sections below. We also show additional results for the third simulation study that is presented in the main text.

A. Monte Carlo Experiment I: EPBR Data

The first experiment follows the design presented in Diamond and Sekhon (2006) for a setting that meets the EPBR conditions. We use 50 treated observations and 100 control observations, with three baseline covariates X that are multivariate normal with zero covariances. For the treated observations the means of the X variables are .20 and for the control observations the means are equal to 0. We generate outcomes with a linear mapping $Y = X\beta + \epsilon$ where $\epsilon \sim N(0, .5)$ and $\beta = (1, 1, 1)'$. The true treatment effect is zero for all units. We consider three variants of this design:

- Design A: the variances of the covariates in X are unity in both groups; the easiest, but presumably most unrealistic case.

- Design B: the variances of the covariates in X are .5 in the control group and and 1.5 in the treatment group. This provides a somewhat more difficult but realistic case; in many empirical cases variances may differ between the two groups.

- Design C: the variances of the covariates in X are equal to unity, but we include all three squared terms in the preprocessing adjustment. Since these squared terms are omitted from the outcome equation (ie. the mapping from X to Y), this scenario mirrors a case where a researcher adjusts for three irrelevant covariates (as in Brookhart, Schneeweiss, Rothman, Glynn, Avorn and Sturmer (2006)). We include this case because adjusting for squared terms is often recommended in practice.

For the propensity score methods we use the correctly estimated propensity score (from a logistic regression that is linear in X), instead of the true score because the former is known to be more efficient. We run 1,000 simulations and report the mean estimate which indicates the bias (multiplied by 100) and the root mean squared error (MSE).

A.1. Results for Design A: Equal Variances

The results for design A (equal variances), design B (unequal variances), and design C (irrelevant covariates) are shown in the upper, middle, and lower panel of Table I respectively. We find that across all three designs the entropy balancing adjustment is unbiased and highly efficient; it outperforms all other propensity score methods and multivariate matching methods in terms of MSE. In particular for design A, the entropy balancing adjustment has an MSE that is more than four times lower than that of the conventional propensity score weighting estimator where a probit regression is used to estimate the score. This is expected because the entropy balancing adjustment fully incorporates the information about the known sample moments. The MSE of entropy balancing is 13 times lower compared to propensity score matching, about 8 times lower than Mahalanobis matching and about 11 times lower than the joint propensity score Mahalanobis distance matching. Consistent with Diamond and Sekhon (2006), Genetic matching dominates the other matching techniques in terms of MSE, but its MSE is still more than 3 times larger than that of entropy balancing. The fact that the matching adjustment (except Genetic matching) are generally less efficient is consistent with the results from Abadie and Imbens (2006). We also find that the multivariate matching methods are all biased. This is consistent with Abadie and Imbens (2006) who show that the bias of matching is of order $O(N^{-1/k})$ where k is the number of continuous covariates. The patterns are very similar for design B with unequal variances, except that the the differences in MSE are now amplified due to the higher variances in the treatment group. The results for design C (lower panel) shows that the inclusion of the irrelevant variables does not adversely affect entropy balancing, but the other methods now exhibit lower MSE compared to design A.

In summary, the first monte carlo experiment shows that in this setting where the conditions necessary for EPBR hold, entropy balancing outperforms the other preprocessing techniques.

B. Monte Carlo Experiment II: Non-EPBR Data

B.1. Design

The second experiment follows the second experiment in Diamond and Sekhon (2006). The covariates are taken from the Dehejia and Wahba (1999) experimental sample of the LaLonde (1986) data (see Diamond and Sekhon (2006) for details). The covariates are not ellipsoidally distributed and thus the EPBR conditions do not hold. Assuming a constant treatment effect of $1,000 the fictional earnings are a non-linear function of only two covariates:

$$Y = 1000D + .1\,exp[\,.7\,log(\texttt{re74} + .01)] + .7\,log(\texttt{re75} + .01)]) + \epsilon$$

where $\epsilon \sim N(0, .5)$, `re74` and `re75` are real earnings in 1974 and 1975 and D is the treatment indicator. The true propensity score is:

$$\pi_i = logit^{-1}[1 + .5.\mu + .01log(\texttt{age}^2) - .3log(\texttt{educ}^2) - .01log(\texttt{re74} + .01)^2 + .01log(\texttt{re75} + .01)^2]$$

where the linear predictor μ is obtained from regressing the actual treatment indicator on $\texttt{age}^2$, $\texttt{educ}^2$, `black`, `hispanic`, `married`, `nodegree`, $\texttt{re74}^2$, $\texttt{re75}^2$, `u74`, and `u75` in the Dehejia Wahba sample. So the true propensity score is a combination of this logistic regression plus the extra variables specified in the equation above.

In the Monte Carlo replications we use the following incorrect functional form to estimate the propensity score:

$$\begin{aligned}\hat{\mu} &= \alpha_0 + \alpha_1\texttt{age} + \alpha_2\texttt{educ} + \alpha_3\texttt{black} + \alpha_4\texttt{hispanic} + \\ &\quad \alpha_5\texttt{married} + \alpha_6\texttt{nodegree} + \alpha_7\texttt{re74} + \alpha_8\texttt{re75} + \alpha_9\texttt{u74} + \alpha_{10}\texttt{u75}\end{aligned}$$

We run 1,000 simulations and report the mean estimate which indicates the bias (multiplied by 100) and the root mean squared error (MSE). We also report the average computing time per simulation measured in seconds.

B.2. Results

The results are displayed in Table II. Entropy balancing achieves the second lowest bias and the lowest MSE across all adjustments. It is also much faster compared to Genetic

Matching, which achieves the second lowest MSE in this experiment. The propensity score methods perform badly given the incorrect specification of the propensity score model. This indicates that entropy balancing retains good finite sample properties in this situation where the EPRB conditions do not hold.

C. Additional Tables for Monte Carlo Experiment III

Tables III and IV display the results for the third simulation discussed in the main paper for the samples sizes $N = 600$ and $N = 1,500$ respectively. Overall the results are very similar to the case of $N = 300$ except that the propensity score methods (with a correctly estimated score) improve as the sample size grows. Entropy balancing achieves the lowest MSE across all simulations.

II. Additional Results for LaLonde Application

Table V presents additional statistics for the covariate balance in the LaLonde application described in the main text. We can see that both the means and the variances are much more similar after entropy balancing when comparing the treatment and control group in the preprocessed data (the last few columns present the balance results form propensity score weighting as a benchmark). Figure 1 shows the QQ-plots that compare the distributions for all four continuous variables: pretreatment earnings in 1975 and 1974, age, and education. The black dots represent empirical quantile estimates for the raw data. The gray dots represent quantile estimates for the reweighted data. The distributions of all four variables are much more similar after the entropy balancing adjustment.

III. Additional Results for News Media Persuasion Application

Table VI shows additional balance statistics for the data used in the Ladd and Lenz (2010) study on the effect of news media persuasion in the 1997 British general election. As discussed in the main text, entropy balancing exactly adjusts all the means and almost all of the variances in this data.

IV. Additional Application: The Fox News Effect

In this section we provide another application of entropy balancing by reanalyzing data from DellaVigna and Kaplan (2007) who study the effect of media bias on voting. The authors exploit the fact that between 1996 and 2000 the conservative Fox News Channel was introduced in the cable programming of about 20 percent of U.S. towns. Using voting data that is aggregated at the town level, the study compares the gain in Republican two-party vote share from the 1996 to 2000 Presidential election between towns that broadcasted Fox News by 2000 in their cable programming and towns that did not. The data covers 9,256 towns overall, 1,807 of which had Fox News availability by 2000 (the treated towns). Using various regression specifications, the authors find that the introduction of Fox News increased Republican vote share gains by 0.4 to 0.7 percentage points.

The authors control for a range of confounders that capture town characteristics measured in the 2000 census including the population size, median income, unemployment rate, and other socio-demographic characteristics for race, gender, urban, education, and married. They include additional controls that measure the trend in each of these variables from the 1990 to 2000 census. The authors also control for the number of cable channels and the number of potential cable subscribers in 2000. There are 26 covariates overall.[1]

Estimating the effect of Fox News from this data is difficult for various reasons. First, the effect is fairly small according to the original study. Second, all of the control variables are continuous and given the heterogeneity across towns their distributions are often heavily skewed as can be seeen in Figure 2 which visualizes the distributions using histograms. Third, since the introduction of Fox News into local cable markets was highly selective, the treatment and control towns strongly differ on many important characteristics. In particular, towns with Fox News availability in 2000 were much larger markets with more channels and potential cable subscribers; the standardized difference in means exceeds the

[1] Notice that the authors also control for House district or county fixed effects in some of their analysis, but many districts and counties have no variation on the treatment variable and including these fixed effects if anything lowers the positive effect of Fox News on vote shares according to the authors' specifications. Also notice that the authors weight many of their regressions by population size which we ignore here (they do state that their results are robust to using no weights and a linear regression of the outcomes on all covariates confirms this).

extreme level of one on both of these key variables. The areas with Fox News were also more urban, richer, and more highly educated. To correct for these imbalances we conduct entropy balancing on all covariates and their squared terms to equalize the means as well as the variances between the two groups (52 covariates overall). As in the other examples we also apply the other perprocessing methods and for this purpose estimate the propensity score with a logistic regression of the treatment indicator on the raw variables and squared terms (below we replicate the same analysis while omitting the squared terms).

Figure 3 displays the standardized means (left panel) and p-values for the difference in means tests (right panel) for the raw data and after the various adjustment methods. After entropy balancing the means between the two groups are equal for all 52 covariate combinations. Given that we also include squared terms for each of the raw covariates, the variances on these variables are also equal as can be seen in Table VII which shows additional balance statistics. According to these metrics the balance is higher than that produced by the other adjustment methods. The other methods often leave several of the most important covariates imbalanced (the differences in means remain large and significant) and in several cases the imbalance on key moments is actually increased over the unadjusted data which can be avoided in entropy balancing by including the relevant moments in the reweighting.

A comparison of the average Republican two-party vote share between the treated and control towns in the preprocessed data yields an insignificant effect estimate that is very close to zero. To investigate the model dependency we again examine the effect estimates across a wide range of possible specifications. We create a dataset that includes all raw covariates, their squared terms, and all pairwise interactions (377 covariate combinations overall). We then fit one million regressions of the outcome on the treatment variable and covariates that we randomly sample from the set of all possible subsets of the covariates.[2] We fit each regression in the raw and the preprocessed data (regressions are weighted by the entropy balancing weights). Figure 4 shows the densities of the estimates of the Fox News effect across the regression specifications. In the raw data the estimates vary rather widely

[2]Notice that there are 3.078282×10^{113} possible subsets ($\sum_{i=1}^{377} \binom{377}{i}$).

within ± 1.5 percentage points of vote share, which may be expected given the limited overlap in the data. The model dependency is much reduced after the entropy balancing adjustment; the range of effect estimates is now narrowed down to $\pm .3$ percentage points.[3]

Finally, to more closely mirror common practice in applied work, Figures 5 and 6 show a replication of the balance figures for the same analysis where we now omit all squared terms from the preprocessing; the propensity score is estimated with a logistic regression of the treatment indicator on all raw covariates. Such simple propensity score models are widely used in practice where researchers often do not include all squared terms. We can see that the balancing property of the propensity score is now much worse. In fact, when squared terms are omitted the balance on many variables is significantly worse after the propensity score weighting adjustment compared to the raw data (the results look slightly better for propensity score matching). This shows that ill-estimated propensity scores can fail to produce good balance. In this case weighting on the logistic propensity score increases the imbalance over the unadjusted data on many covariates. This may be expected given that the simple model with only the raw covariates does a poor job of capturing the assignment process and the procedure also assigns some very extreme weights since the logistic propensity score is close to zero for some units (Rosenbaum, 1987).

Taken together the application suggests that entropy balancing delivers a high degree of balance in this dataset (as measured by standard metrics). Higher balance reduces model dependence for the estimation of causal effects. It is important to recognize that this replication is intended to simply illustrate the use of entropy balancing in an interesting dataset, it does not invalidate the results of the original study which contains many additional tests and evidence that we do not consider here. We are grateful to the authors for making their data freely available.

[3]Notice that the variability in the preprocessed estimates is entirely driven by the fact that the interaction terms are not included in the reweighting adjustment but only in the outcome regressions. Since the entropy balancing includes all raw covariates and their squared terms, the regression estimates are identical across all subsets that do not involve interaction terms.

V. Additional Application: The Financial Returns to Political Office

In this section we provide another application of entropy balancing by reanalyzing data from Eggers and Hainmueller (2010), who study the financial returns to serving in parliament using data on the estates of recently deceased British politicians. We focus on their sample of 223 conservative candidates that ran for the House of Commons during the 1950-1970 period (see the article for a detailed discussion of the data). The treatment variable is a binary indicator that is coded as one for the 104 candidates that ran successfully and served in parliament and zero for the 119 control candidates that lost and did not enter parliament. The outcome variable is logged wealth at death, which is measured using probate values that capture the value of the candidate's estate at the time of death (in real 2007 British Pounds). In order to account for the selection into political office the authors control for a variety of background covariates including the candidate's gender, year of birth, year of death, as well as educational, occupational, and aristocratic background. There are 18 covariates in total.

Columns 1-4 in Table VIII display the covariate balance in the unmatched data. As discussed by the authors there are important imbalances in this data. In particular, successful candidates are more likely than unsuccessful candidates to be male and to have aristocratic backgrounds and elite educations (Eton Schooling and Oxbridge Degrees). Successful candidates are also less likely to be in white-collar professions (engineering, accounting, or public relations), journalism, and teaching professions, and less likely to have business backgrounds. The standardized bias exceeds $|.1|$ for all but three of the covariates.

To correct for these imbalances we conduct entropy balancing and specify moment conditions to equalize the means of all 18 covariates between the treatment and the re-weighted control group. Columns 5-8 in Table VIII display the covariates means as well as the various balance metrics computed with the re-weighted control group. The mean differences are now reduced to zero on all covariates. Except for the year of birth and year of death measures, all variables are binary so by adjusting their means the variances are also adjusted. This constitutes a higher level of balance than previously achieved for these metrics in this dataset. The difference in means between the treatment group and the reweighted control group yields an average treatment effect on the treated of .99 with a t-statistic of about 2.8, indicating that at serving in Parliament considerably increased wealth at death from conservative MPs. This estimate is close to the magnitude estimated by the authors (the original study used genetic matching).

As a comparison the last few columns present similar balance measures when propensity score weighting, based on a score that is estimated with a logistic regression of the treatment indicator on all covariates, is applied to the same data. While propensity score weighting leads to some balance improvements, important imbalances remain one some key variables such as Oxbridge Degrees, Barrister and Solicitor, and aristocratic background an even worse one some covariates like White Collar professions and year of death the imbalance actually increases over the unmatched data. These imbalances may be corrected by tinkering with the propensity score specification. However, with 18 covariates is is difficult and tedious to find a model that jointly balance all covariates. This shows the benefits of entropy balancing which provides balance by construction of the moment conditions. Figures 7 and 8 show the standardized bias and p-value for the difference in means tests for each covariate when we apply various other matching methods to the same data. Entropy balancing improves on these balance metrics over all other methods including Mahalanobis distance matching, genetic matching, and matching or weighting on the logistic propensity score. Among the other methods, genetic matching does best although some imbalances remain on aristocratic backgrounds and Oxbridge degrees.

References

Abadie, A. and Imbens, G. (2006), 'Large Sample Properties of Matching Estimators for Average Treatment Effects', *Econometrica* **74**(2), 235–267.

Brookhart, M., Schneeweiss, S., Rothman, K., Glynn, R., Avorn, J. and Sturmer, T. (2006), 'Variable Selection for Propensity Score Models', *American Journal of Epidemiology* **163**(12), 1149–1156.

Diamond, A. J. and Sekhon, J. (2006), Genetic matching for causal effects: A general multivariate matching method for achieving balance in observational studies. Unpublished manuscript, Dept. of Political Science, UC Berkeley.

Rosenbaum, P. (1987), 'Model-based direct adjustment', *Journal of the American Statistical Association* **82**(398), 387–394.

VI. Tables

Table I: Results for First Monte Carlo Experiment

Estimator:	RAW	PSM	MD	PSMD	GM	PSW	EB
Design A: equal variances							
Bias	−60.19	−2.34	−9.66	−9.83	−11.36	−0.58	0.02
MSE	45.51	4.87	2.31	2.47	3.25	1.21	0.23
Bias / Bias EB	2746.19	106.78	440.73	448.53	518.36	26.31	1.00
MSE / MSE EB	195.78	20.94	9.93	10.61	13.99	5.22	1.00
Design B: unequal variances							
Bias	−59.79	−10.01	−20.84	−23.62	−13.37	−25.64	−0.11
MSE	46.58	5.17	7.07	8.78	3.88	9.38	0.25
Bias / Bias EB	564.52	94.54	196.76	222.99	126.19	242.09	1.00
MSE / MSE EB	188.78	20.94	28.64	35.59	15.73	38.04	1.00
Design C: equal variances and squared terms							
Bias	−63.24	−3.41	−10.10	−10.39	−11.11	−0.89	0.15
MSE	48.18	4.94	2.42	2.62	3.06	1.24	0.25
Bias / Bias EB	420.09	22.64	67.11	69.02	73.78	5.92	1.00
MSE / MSE EB	190.98	19.57	9.61	10.38	12.13	4.90	1.00

Note: Three independent normal covariates X drawn for 50 treated units with means .2 and for 100 control units with means 0. Outcome mapping is linear $Y = X\beta + \epsilon$ with $\beta = (1,1,1)'$ and $\epsilon \sim N(0,.5)$. The true treatment effect is zero for all units. Assumptions satisfy the conditions for EPBR. 1,000 simulations. *Design A*: Equal unit variances in both groups. *Design B*: Unequal variances: 1.5 in treatment and .5 in control group. *Design C*: Equal unit variances. Squared terms of X are included for all estimators, but not the outcome. This mirrors the situation of adjusting for additional irrelevant covariates. Raw: Difference of means; PSM: Propensity score matching; MD: Mahalanobis Distance Matching, PSMD: MD matching on the PS and orthogonalized covariates; GM: Genetic Matching; PSW: weighting on the PS; EB: entropy balancing. Matching is 1:1 pair matching. The propensity score is estimated with a linear logit in X.

Table II: Results for Second Monte Carlo Experiment

Estimator:	RAW	MD	GM	PS	PSMD	PSW	EB
Bias	−450	384	61	93	496	−183	−78
MSE	1632	690	518	1050	782	982	464
Time	0	0	23	0	0	0	0
Bias / Bias EB	5.8	4.9	0.8	1.2	6.4	2.3	1
MSE / MSE EB	3.5	1.5	1.1	2.3	1.7	2.1	1
Time / Time EB	0.0	1.0	1186.5	1.0	1.5	0.5	1

Note: $1,000 is the true effect for all units. The experiment follows the second experiment presented in Diamond and Sekhon (2006). The conditions do not satisfy EPBR and the mapping between the baseline covariates and the outcome is non-linear. The propensity score is misspecified.

Table III: Results for Monte Carlo Experiment (N=600)

Sample Design 1: Strong Separation and Normal Errors

MSE	RAW	MD	PSM1	PSM2	PSM3	PSMD1	PSMD2	PSMD3	PSW1	PSW2	PSW3	**EB**
Ratio CtoT 1 Y1	320	21	14	358	182	46	352	186	15	354	183	2
Ratio CtoT 1 Y2	497	15	7	153	524	49	164	449	19	152	529	2
Ratio CtoT 1 Y3	997	915	671	1326	1640	1764	1479	2026	742	1205	1526	172
Ratio CtoT 3 Y1	323	18	12	363	185	39	353	190	10	356	185	2
Ratio CtoT 3 Y2	499	13	5	153	525	41	155	459	12	156	532	2
Ratio CtoT 3 Y3	1041	807	599	1186	1675	1601	1357	2035	592	1123	1547	168
Ratio CtoT 5 Y1	324	17	14	368	186	33	352	190	8	360	182	2
Ratio CtoT 5 Y2	499	12	7	155	525	34	143	467	9	158	526	2
Ratio CtoT 5 Y3	1125	739	604	1133	1719	1431	1269	2022	627	1072	1589	186
BIAS	RAW	MD	PSM1	PSM2	PSM3	PSMD1	PSMD2	PSMD3	PSW1	PSW2	PSW3	**EB**
Ratio CtoT 1 Y1	177	42	6	185	132	65	185	135	7	186	134	0
Ratio CtoT 1 Y2	222	35	10	121	227	68	126	211	8	121	229	2
Ratio CtoT 1 Y3	294	292	72	333	393	412	374	443	51	329	384	69
Ratio CtoT 3 Y1	178	38	7	187	133	60	185	136	11	186	134	0
Ratio CtoT 3 Y2	222	31	8	120	227	61	121	213	14	123	229	2
Ratio CtoT 3 Y3	296	272	63	311	393	390	355	441	76	314	384	61
Ratio CtoT 5 Y1	177	35	5	187	131	54	183	135	12	187	133	−1
Ratio CtoT 5 Y2	222	28	7	120	226	54	115	213	17	123	227	1
Ratio CtoT 5 Y3	294	253	58	284	389	362	335	432	96	302	384	50

Sample Design 2: Weaker Separation and Normal Errors

MSE	RAW	MD	PSM1	PSM2	PSM3	PSMD1	PSMD2	PSMD3	PSW1	PSW2	PSW3	**EB**
Ratio CtoT 1 Y1	144	8	5	126	82	18	126	88	3	121	78	1
Ratio CtoT 1 Y2	223	4	3	53	222	13	44	191	3	50	221	1
Ratio CtoT 1 Y3	514	440	273	458	823	859	722	1143	219	372	767	95
Ratio CtoT 3 Y1	145	8	5	126	85	16	124	89	2	123	79	1
Ratio CtoT 3 Y2	224	4	3	51	227	12	42	199	3	50	224	1
Ratio CtoT 3 Y3	557	387	259	431	861	775	659	1149	200	354	783	103
Ratio CtoT 5 Y1	146	8	8	131	86	14	123	92	2	124	80	1
Ratio CtoT 5 Y2	225	4	4	54	230	10	38	204	3	51	225	2
Ratio CtoT 5 Y3	629	362	387	458	939	708	630	1138	215	330	786	130
BIAS	RAW	MD	PSM1	PSM2	PSM3	PSMD1	PSMD2	PSMD3	PSW1	PSW2	PSW3	**EB**
Ratio CtoT 1 Y1	118	25	1	107	87	40	109	91	1	107	87	0
Ratio CtoT 1 Y2	148	15	2	68	147	33	63	136	1	68	147	1
Ratio CtoT 1 Y3	195	200	15	175	269	284	257	329	6	173	267	18
Ratio CtoT 3 Y1	117	23	2	107	87	37	107	91	3	108	87	0
Ratio CtoT 3 Y2	148	13	2	67	147	31	60	138	4	68	147	1
Ratio CtoT 3 Y3	199	184	18	161	271	266	241	326	22	165	267	14
Ratio CtoT 5 Y1	117	20	−1	107	86	32	104	91	2	107	86	−1
Ratio CtoT 5 Y2	147	10	1	67	147	26	54	138	4	67	147	−0
Ratio CtoT 5 Y3	198	171	6	146	270	246	226	316	22	154	261	3

Sample Design 3: Medium Separation and Leptokurtic Errors

MSE	RAW	MD	PSM1	PSM2	PSM3	PSMD1	PSMD2	PSMD3	PSW1	PSW2	PSW3	**EB**
Ratio CtoT 1 Y1	218	16	12	239	138	32	233	150	10	234	139	1
Ratio CtoT 1 Y2	343	10	6	103	398	31	108	346	15	115	409	2
Ratio CtoT 1 Y3	1049	794	583	843	1346	1512	1219	1721	506	862	1210	234
Ratio CtoT 3 Y1	216	14	11	237	138	27	227	148	8	234	139	1
Ratio CtoT 3 Y2	341	9	5	104	396	27	99	348	13	116	405	2
Ratio CtoT 3 Y3	1069	680	577	761	1399	1352	1044	1644	512	789	1218	222
Ratio CtoT 5 Y1	224	14	13	249	144	25	232	152	7	246	141	2
Ratio CtoT 5 Y2	343	9	6	108	402	23	92	358	11	118	407	2
Ratio CtoT 5 Y3	1172	655	820	723	1462	1244	1006	1670	524	780	1233	238
BIAS	RAW	MD	PSM1	PSM2	PSM3	PSMD1	PSMD2	PSMD3	PSW1	PSW2	PSW3	**EB**
Ratio CtoT 1 Y1	146	36	7	150	115	54	150	121	25	151	116	0
Ratio CtoT 1 Y2	184	29	11	98	198	53	101	184	35	105	201	4
Ratio CtoT 1 Y3	301	273	74	247	351	381	339	407	177	274	339	121
Ratio CtoT 3 Y1	145	32	6	149	114	50	147	119	24	151	116	0
Ratio CtoT 3 Y2	184	25	10	98	197	49	96	185	33	106	200	3
Ratio CtoT 3 Y3	302	250	63	217	357	358	309	395	181	263	339	117
Ratio CtoT 5 Y1	146	31	5	151	115	45	147	119	20	154	116	−0
Ratio CtoT 5 Y2	183	23	10	98	197	43	90	186	28	105	199	3
Ratio CtoT 5 Y3	297	237	42	195	351	335	292	388	167	245	333	103

Note: Results show MSE and Bias across 1,000 simulations. Six covariates with a mixture of continuous, binary, and categorical variables. Experimental factors are: 3 sample designs (sample design 1: strong separation and normal errors; sample design 2: weaker separation and normal errors; sample design 3: medium separation and leptokurtic errors), 3 outcome designs (Y1 linear: $Y1 = X_1 + X_2 + X_3 - X_4 + X_5 + X_6 + \eta$; Y2 somewhat non-linear $Y2 = X_1 + X_2 + 0.2\,X_3\,X_4 - \sqrt{X_5} + \eta$; Y3 highly non-linear: $Y3 = (X_1 + X_2 + X_5)^2 + \eta$), and 3 controls-to-treated ratios (Ratio CtoT 1, 3, and 5). Estimators are Raw: Difference of means; MD: Mahalanobis distance matching, GM: Genetic matching; PSM: Propensity score matching; PSMD: MD matching on the PS and orthogonalized covariates; PSW: weighting on the PS; EB: entropy balancing. All matching is 1:1 pair matching. We use three specifications (labeled with a 1, 2, or 3 postfix) for all propensity score based methods (PSM, PSW, PSMD). The first propensity score model is correct for sample designs 1 and 2, and slightly misspecified for sample design 3. Propensity score models 2 and 3 are increasing in misspecification (as measured by the linear correlation between the true and the estimated score). 1000 simulations for each scenario; the true treatment effect is zero.

Table IV: Results for Monte Carlo Experiment (N=1,500)

Sample Design 1: Strong Separation and Normal Errors

MSE	RAW	MD	PSM1	PSM2	PSM3	PSMD1	PSMD2	PSMD3	PSW1	PSW2	PSW3	**EB**
Ratio CtoT 1 Y1	318	11	6	343	180	28	327	180	8	342	181	1
Ratio CtoT 1 Y2	496	8	2	147	526	30	134	467	10	146	530	1
Ratio CtoT 1 Y3	952	562	301	1097	1540	1262	1078	1863	488	1060	1498	95
Ratio CtoT 3 Y1	317	9	5	343	181	23	321	180	4	341	180	1
Ratio CtoT 3 Y2	495	6	2	143	522	24	123	468	6	149	524	1
Ratio CtoT 3 Y3	936	475	258	970	1547	1099	960	1833	280	977	1511	75
Ratio CtoT 5 Y1	318	9	5	343	180	20	318	181	4	348	181	1
Ratio CtoT 5 Y2	495	6	2	143	522	20	115	473	5	155	523	1
Ratio CtoT 5 Y3	988	420	263	837	1603	972	875	1833	263	910	1548	81
BIAS	RAW	MD	PSM1	PSM2	PSM3	PSMD1	PSMD2	PSMD3	PSW1	PSW2	PSW3	**EB**
Ratio CtoT 1 Y1	178	31	3	184	133	52	180	133	3	184	134	0
Ratio CtoT 1 Y2	222	26	6	120	229	54	115	216	5	120	230	2
Ratio CtoT 1 Y3	300	232	44	318	388	352	323	429	30	316	384	54
Ratio CtoT 3 Y1	177	28	3	184	133	47	178	133	9	184	134	−0
Ratio CtoT 3 Y2	222	22	4	119	228	48	109	216	12	121	228	1
Ratio CtoT 3 Y3	296	213	32	297	388	327	304	424	61	304	385	41
Ratio CtoT 5 Y1	177	26	2	183	133	42	176	133	12	185	134	0
Ratio CtoT 5 Y2	222	20	4	118	227	43	105	216	17	123	228	1
Ratio CtoT 5 Y3	297	197	37	266	391	305	287	421	91	291	387	39

Sample Design 2: Weaker Separation and Normal Errors

MSE	RAW	MD	PSM1	PSM2	PSM3	PSMD1	PSMD2	PSMD3	PSW1	PSW2	PSW3	**EB**
Ratio CtoT 1 Y1	140	4	2	114	77	10	109	81	1	114	77	0
Ratio CtoT 1 Y2	221	2	1	47	219	7	32	197	2	47	220	0
Ratio CtoT 1 Y3	444	238	95	329	774	557	465	972	103	298	734	41
Ratio CtoT 3 Y1	140	4	2	115	78	9	104	82	1	115	77	0
Ratio CtoT 3 Y2	220	1	1	47	221	6	28	201	1	47	219	0
Ratio CtoT 3 Y3	448	206	104	289	754	495	409	954	80	277	731	42
Ratio CtoT 5 Y1	142	4	3	118	80	8	105	86	1	118	79	1
Ratio CtoT 5 Y2	221	2	1	48	222	5	28	205	1	49	220	1
Ratio CtoT 5 Y3	488	186	135	278	782	437	381	952	90	262	738	55
BIAS	RAW	MD	PSM1	PSM2	PSM3	PSMD1	PSMD2	PSMD3	PSW1	PSW2	PSW3	**EB**
Ratio CtoT 1 Y1	117	18	1	105	87	31	103	89	0	106	87	0
Ratio CtoT 1 Y2	148	10	1	67	147	25	55	139	1	67	148	0
Ratio CtoT 1 Y3	196	150	11	163	271	232	210	308	1	164	267	10
Ratio CtoT 3 Y1	117	16	−0	105	87	28	100	89	2	106	87	−0
Ratio CtoT 3 Y2	148	9	0	67	148	23	51	140	3	68	147	0
Ratio CtoT 3 Y3	196	139	6	152	267	218	195	304	18	158	266	8
Ratio CtoT 5 Y1	117	15	−0	106	87	25	100	91	3	107	87	−0
Ratio CtoT 5 Y2	148	7	0	67	147	20	49	141	4	68	147	−0
Ratio CtoT 5 Y3	196	127	0	136	266	200	184	299	26	150	264	2

Sample Design 3: Medium Separation and Leptokurtic Errors

MSE	RAW	MD	PSM1	PSM2	PSM3	PSMD1	PSMD2	PSMD3	PSW1	PSW2	PSW3	**EB**
Ratio CtoT 1 Y1	213	8	5	226	133	19	216	143	7	226	137	1
Ratio CtoT 1 Y2	339	5	3	98	390	19	87	354	13	112	403	1
Ratio CtoT 1 Y3	933	491	319	625	1264	1075	845	1468	379	745	1146	160
Ratio CtoT 3 Y1	217	7	5	230	134	17	212	144	6	231	137	0
Ratio CtoT 3 Y2	338	5	2	97	391	16	79	360	12	114	401	1
Ratio CtoT 3 Y3	967	426	291	557	1266	955	719	1450	386	700	1163	154
Ratio CtoT 5 Y1	216	7	6	229	135	15	208	144	5	235	136	1
Ratio CtoT 5 Y2	338	4	2	96	393	13	72	365	9	113	399	1
Ratio CtoT 5 Y3	1004	387	407	463	1307	850	640	1470	397	621	1189	166
BIAS	RAW	MD	PSM1	PSM2	PSM3	PSMD1	PSMD2	PSMD3	PSW1	PSW2	PSW3	**EB**
Ratio CtoT 1 Y1	145	27	4	149	114	43	146	119	24	149	116	−0
Ratio CtoT 1 Y2	184	21	7	98	197	42	92	188	34	105	200	3
Ratio CtoT 1 Y3	296	218	44	228	350	324	286	380	174	265	335	111
Ratio CtoT 3 Y1	146	25	3	150	115	40	144	119	23	151	117	−0
Ratio CtoT 3 Y2	183	19	7	97	197	39	87	189	33	106	200	3
Ratio CtoT 3 Y3	301	202	34	210	349	305	262	376	178	257	337	109
Ratio CtoT 5 Y1	146	24	4	149	114	36	142	118	20	152	116	−0
Ratio CtoT 5 Y2	183	18	6	96	197	35	82	190	29	105	198	3
Ratio CtoT 5 Y3	298	190	29	176	351	284	242	375	176	239	338	107

Note: Results show MSE and Bias across 1,000 simulations. Six covariates with a mixture of continuous, binary, and categorical variables. Experimental factors are: 3 sample designs (sample design 1: strong separation and normal errors; sample design 2: weaker separation and normal errors; sample design 3: medium separation and leptokurtic errors), 3 outcome designs (Y1 linear: $Y1 = X_1 + X_2 + X_3 - X_4 + X_5 + X_6 + \eta$; Y2 somewhat non-linear $Y2 = X_1 + X_2 + 0.2\, X_3\, X_4 - \sqrt{X_5} + \eta$; Y3 highly non-linear: $Y3 = (X_1 + X_2 + X_5)^2 + \eta$), and 3 controls-to-treated ratios (Ratio CtoT 1, 3, and 5). Estimators are Raw: Difference of means; MD: Mahalanobis distance matching, GM: Genetic matching; PSM: Propensity score matching; PSMD: MD matching on the PS and orthogonalized covariates; PSW: weighting on the PS; EB: entropy balancing. All matching is 1:1 pair matching. We use three specifications (labeled with a 1, 2, or 3 postfix) for all propensity score based methods (PSM, PSW, PSMD). The first propensity score model is correct for sample designs 1 and 2, and slightly misspecified for sample design 3. Propensity score models 2 and 3 are increasing in misspecification (as measured by the linear correlation between the true and the estimated score). 1000 simulations for each scenario; the true treatment effect is zero.

Table V: Covariate Balance in Lalonde Data

	Raw Data					Entropy Balancing				PS Weighting			
	Means		Std.	Var	T	Mean	Std.	Var	T pval	Mean	Std.	Var	T
	Treated	Controls	Diffs	Ratio	pval	Controls	Diffs	Ratio	pval	Controls	Diffs	Var Ratio	pval
Age	25.82	33.23	−0.95	0.42	0.00	25.82	0	1.00	1	27.76	−0.37	0.92	0.01
Schooling	10.35	12.03	−0.83	0.49	0.00	10.35	0	1.00	1	10.22	0.09	1.18	0.51
Black	0.84	0.07	3.96	1.95	0.00	0.84	0	1.00	1	0.88	−0.13	1.21	0.36
Hispanic	0.06	0.07	−0.07	0.84	0.47	0.06	0	1.00	1	0.04	0.10	1.35	0.47
Married	0.19	0.71	−1.62	0.75	0.00	0.19	0	1.00	1	0.16	0.12	1.16	0.39
HS Dropout	0.71	0.30	1.27	1.00	0.00	0.71	0	1.00	1	0.74	−0.10	1.07	0.49
Earnings 1974	2095.57	14024.14	−1.75	0.26	0.00	2095.57	0	1.15	1	1684.68	0.13	1.36	0.37
Earnings 1975	1532.06	13642.53	−1.84	0.12	0.00	1532.06	0	0.90	1	1307.01	0.10	1.07	0.48
Unemployed 1974	0.71	0.12	2.50	1.97	0.00	0.71	0	1.00	1	0.76	−0.17	1.13	0.24
Unemployed 1975	0.60	0.11	2.18	2.48	0.00	0.60	0	1.00	1	0.66	−0.18	1.07	0.22
Age*Age	717.39	1225.91	−0.92	0.30	0.00	717.39	0	1.13	1	826.13	−0.36	1.03	0.01
Schooling*Age	266.98	395.54	−1.17	0.35	0.00	266.98	0	1.02	1	282.20	−0.24	1.09	0.09
Schooling*Schooling	111.06	152.90	−0.88	0.34	0.00	111.06	0	1.01	1	107.82	0.12	1.16	0.39
Black*Age	21.91	2.40	2.97	1.65	0.00	21.91	0	1.00	1	24.65	−0.33	1.00	0.02
Black*Schooling	8.70	0.81	3.58	1.99	0.00	8.70	0	1.03	1	8.90	−0.07	1.25	0.62
Hispanic*Age	1.36	2.38	−0.16	0.37	0.01	1.36	0	0.97	1	0.96	0.11	1.36	0.45
Hispanic*Schooling	0.58	0.73	−0.07	0.70	0.40	0.58	0	0.94	1	0.42	0.10	1.28	0.48
Married*Age	5.56	25.85	−1.55	0.42	0.00	5.56	0	1.04	1	4.71	0.10	1.15	0.46
Married*Schooling	1.96	8.56	−1.55	0.49	0.00	1.96	0	0.98	1	1.58	0.14	1.20	0.34
Married*Black	0.16	0.05	0.74	3.05	0.00	0.16	0	1.00	1	0.13	0.11	1.17	0.44
Married*Hispanic	0.02	0.05	−0.23	0.32	0.00	0.02	0	1.00	1	0.01	0.03	1.21	0.81
HS Dropout*Age	17.97	10.09	0.65	0.59	0.00	17.97	0	0.98	1	20.42	−0.26	0.89	0.07
HS Dropout*Black	0.61	0.03	4.25	7.74	0.00	0.61	0	1.00	1	0.66	−0.16	1.07	0.26
HS Dropout*Hispanic	0.05	0.04	0.07	1.25	0.53	0.05	0	1.00	1	0.03	0.10	1.40	0.47
HS Dropout*Married	0.14	0.20	−0.20	0.77	0.03	0.14	0	1.00	1	0.11	0.11	1.19	0.44
Earnings 1974*Age	54074.04	509069.25	−1.57	0.10	0.00	54074.04	0	1.14	1	43947.39	0.12	1.33	0.41
Earnings 1974*Schooling	22898.73	171241.56	−1.62	0.20	0.00	22898.73	0	1.23	1	18132.78	0.13	1.52	0.37
Earnings 1974*Black	1817.20	840.59	0.36	1.54	0.01	1817.20	0	1.21	1	1451.45	0.12	1.46	0.40
Earnings 1974*Hispanic	151.40	893.68	−0.26	0.09	0.00	151.40	0	0.79	1	126.10	0.03	0.94	0.83
Earnings 1974*Married	760.63	11809.15	−1.48	0.12	0.00	760.63	0	1.37	1	608.57	0.07	1.73	0.64
Earnings 1974*HS Dropout	1094.15	3432.61	−0.45	0.21	0.00	1094.15	0	1.06	1	896.94	0.09	1.26	0.54
Earnings 1975*Age	41167.28	489047.95	−1.61	0.06	0.00	41167.28	0	1.00	1	35832.82	0.08	1.14	0.58
Earnings 1975*Schooling	15880.57	167310.76	−1.69	0.07	0.00	15880.57	0	0.93	1	13332.83	0.11	1.14	0.44
Earnings 1975*Black	1257.04	804.32	0.17	0.69	0.05	1257.04	0	0.97	1	1082.19	0.08	1.14	0.56
Earnings 1975*Hispanic	153.73	884.98	−0.27	0.07	0.00	153.73	0	0.65	1	122.50	0.04	0.88	0.76
Earnings 1975*Married	654.34	11366.04	−1.47	0.08	0.00	654.34	0	1.14	1	549.80	0.06	1.42	0.69
Earnings 1975*HS Dropout	1134.96	3290.78	−0.44	0.18	0.00	1134.96	0	0.89	1	983.57	0.07	1.06	0.60
Unemployed 1974*Age	18.78	3.60	1.97	1.63	0.00	18.78	0	0.99	1	21.92	−0.32	0.95	0.02
Unemployed 1974*Schooling	7.26	1.42	2.04	1.58	0.00	7.26	0	1.01	1	7.71	−0.13	1.17	0.35
Unemployed 1974*Black	0.60	0.01	6.44	23.63	0.00	0.60	0	1.00	1	0.68	−0.22	1.10	0.11
Unemployed 1974*Hispanic	0.03	0.01	0.36	3.74	0.07	0.03	0	1.00	1	0.02	0.10	1.55	0.47
Unemployed 1974*Married	0.11	0.06	0.33	1.83	0.02	0.11	0	1.00	1	0.09	0.10	1.20	0.49
Unemployed 1974*HS Dropout	0.52	0.05	2.93	5.28	0.00	0.52	0	1.00	1	0.59	−0.17	1.03	0.22
Unemployed 1974*Earnings 1975	307.44	175.27	0.14	0.69	0.12	307.44	0	0.50	1	310.15	0.00	0.51	0.98
Unemployed 1975*Age	15.98	3.57	1.58	1.73	0.00	15.98	0	0.99	1	19.22	−0.31	0.91	0.03
Unemployed 1975*Schooling	6.15	1.33	1.72	1.84	0.00	6.15	0	1.01	1	6.71	−0.15	1.11	0.28
Unemployed 1975*Black	0.52	0.01	5.51	22.19	0.00	0.52	0	1.00	1	0.60	−0.22	1.04	0.11
Unemployed 1975*Hispanic	0.03	0.01	0.32	3.61	0.10	0.03	0	1.00	1	0.02	0.10	1.59	0.49
Unemployed 1975*Married	0.09	0.06	0.12	1.31	0.30	0.09	0	1.00	1	0.07	0.07	1.17	0.61
Unemployed 1975*HS Dropout	0.43	0.04	2.84	7.08	0.00	0.43	0	1.00	1	0.50	−0.19	0.98	0.19
Unemployed 1975*Earnings 1974	43.85	203.65	−0.14	0.08	0.00	43.85	0	0.67	1	35.68	0.02	0.77	0.87
Unemployed 1975*Unemployed 1974	0.59	0.07	2.69	3.56	0.00	0.59	0	1.00	1	0.65	−0.18	1.07	0.20

Note: Std. Diffs: Standardized difference in means. Var ratio: Ratio of variances. T-pval: p-value from difference of means t-test

Table VI: Covariate Balance in News Media Persuasion Data

	Raw Data					Entropy Balancing				PS Weighting			
	Means		Std.	Var	T	Mean	Std.	Var	T pval	Mean	Std.	Var	T
	Treated	Controls	Diffs	Ratio	pval	Controls	Diffs	Ratio	pval	Controls	Diffs	Var Ratio	pval
Prior Conservative Identification	0.41	0.42	−0.02	1.00	0.86	0.41	0	1.00	1	0.40	0.04	1.01	0.76
Prior Labour Identification	0.34	0.31	0.07	1.04	0.53	0.34	0	1.00	1	0.36	−0.06	0.97	0.66
Prior Liberal Identification	0.13	0.15	−0.08	0.89	0.40	0.13	0	1.00	1	0.14	−0.01	0.98	0.93
White	0.99	0.98	0.09	0.60	0.29	0.99	0	1.00	1	0.98	0.02	0.92	0.91
Working-Class	0.72	0.58	0.39	0.84	0.00	0.72	0	1.00	1	0.72	−0.01	1.01	0.92
Parents Voted Labour	0.44	0.35	0.24	1.08	0.03	0.44	0	1.00	1	0.44	−0.01	1.00	0.96
Prior Ideological Moderation	0.65	0.65	−0.02	0.97	0.88	0.65	0	1.00	1	0.65	0.01	1.01	0.95
Prior Labour Vote	0.39	0.32	0.20	1.09	0.07	0.39	0	1.00	1	0.40	−0.03	0.99	0.83
Prior Conservative Vote	0.39	0.40	−0.04	0.99	0.68	0.39	0	1.00	1	0.37	0.04	1.01	0.76
Prior Liberal Vote	0.16	0.19	−0.12	0.87	0.24	0.16	0	1.00	1	0.16	−0.01	0.99	0.96
Prior Labour Party Support	0.49	0.46	0.11	1.02	0.31	0.49	0	0.96	1	0.49	−0.03	0.97	0.82
Prior Conservative Party Support	0.52	0.52	0.01	1.00	0.90	0.52	0	0.99	1	0.51	0.05	1.00	0.73
Prior Political Knowledge	0.55	0.67	−0.49	1.14	0.00	0.55	0	1.01	1	0.55	0.00	0.99	0.99
Prior Television Viewer	0.22	0.29	−0.22	0.83	0.02	0.22	0	1.00	1	0.21	0.03	1.03	0.84
Prior Daily Newspaper Reader	0.93	0.66	0.82	0.30	0.00	0.93	0	1.00	1	0.93	0.01	0.98	0.94
Prior Ideology	0.55	0.54	0.11	0.85	0.27	0.55	0	1.00	1	0.55	−0.01	1.02	0.93
Authoritarianism	0.58	0.57	0.09	1.29	0.44	0.58	0	1.47	1	0.58	−0.01	1.44	0.96
Prior Trade Union Member	0.22	0.24	−0.07	0.94	0.47	0.22	0	1.00	1	0.22	0.01	1.01	0.96
Prior Coping Mortgage	0.71	0.68	0.08	0.46	0.33	0.71	0	0.51	1	0.71	0.00	0.45	0.99
Prior Education	0.75	0.64	0.42	0.81	0.00	0.75	0	1.09	1	0.75	0.00	1.09	1.00
Prior Income	1.20	1.25	−0.03	1.07	0.81	1.20	0	1.06	1	1.20	0.00	1.06	0.99
Prior Age	0.51	0.61	−0.14	0.39	0.06	0.51	0	1.25	1	0.51	0.00	1.14	0.99
Male	0.45	0.56	−0.29	1.01	0.01	0.45	0	1.00	1	0.45	0.02	1.00	0.86
North West	0.12	0.09	0.16	1.34	0.16	0.12	0	1.00	1	0.13	−0.04	0.93	0.76
Yorks	0.08	0.07	0.03	1.08	0.76	0.08	0	1.00	1	0.07	0.04	1.09	0.78
West Midlands	0.08	0.08	−0.03	0.94	0.76	0.08	0	1.00	1	0.07	0.02	1.05	0.88
East Midlands	0.07	0.07	−0.01	0.97	0.90	0.07	0	1.00	1	0.07	−0.01	0.98	0.95
East Anglia	0.01	0.03	−0.16	0.44	0.05	0.01	0	1.00	1	0.01	0.00	1.03	0.97
SW England	0.07	0.08	−0.05	0.90	0.63	0.07	0	1.00	1	0.07	0.01	1.03	0.93
SE England	0.17	0.17	0.00	1.00	1.00	0.17	0	1.00	1	0.16	0.04	1.06	0.76
Greater London	0.11	0.07	0.23	1.57	0.06	0.11	0	1.00	1	0.12	−0.04	0.93	0.78
Wales	0.06	0.03	0.17	1.64	0.17	0.06	0	1.00	1	0.05	0.06	1.18	0.68
Scotland	0.17	0.26	−0.29	0.74	0.00	0.17	0	1.00	1	0.17	−0.01	0.99	0.95
Profession: Large Employer	0.11	0.16	−0.17	0.76	0.07	0.11	0	1.00	1	0.12	−0.01	0.99	0.96
Profession: Small Employer	0.02	0.05	−0.17	0.50	0.03	0.02	0	1.00	1	0.02	0.00	1.00	1.00
Profession: Self Employed	0.29	0.41	−0.33	0.86	0.00	0.29	0	1.00	1	0.29	0.01	1.01	0.93
Profession: Employee	0.05	0.04	0.06	1.21	0.59	0.05	0	1.00	1	0.05	0.02	1.05	0.91
Profession: Temporary Worker	0.45	0.28	0.55	1.24	0.00	0.45	0	1.00	1	0.46	0.00	1.00	0.98
Profession: Junior	0.04	0.05	−0.08	0.77	0.41	0.04	0	1.00	1	0.04	−0.02	0.93	0.88

Note: Std. Diffs: Standardized difference in means. Var ratio: Ratio of variances. T-pval: p-value from difference of means t-test

Table VII: Covariate Balance in Fox News Data

	Raw Data					Entropy Balancing				PS Weighting			
	Means		Std.	Var	T	Mean	Std.	Var	T pval	Mean	Std.	Var	T
	Treated	Controls	Diffs	Ratio	pval	Controls	Diffs	Ratio	pval	Controls	Diffs	Var Ratio	pval
No. of Cable Channels 2000	4.45	2.47	1.91	2.00	0.00	4.45	0	1.00	1	4.51	−0.05	0.95	0.24
No. of Cable Channels 2000 Sq.	22.37	7.40	1.78	4.75	0.00	22.37	0	1.14	1	23.06	−0.06	1.06	0.22
Population 2000	1.15	0.92	0.10	0.98	0.01	1.15	0	1.00	1	1.18	−0.01	0.97	0.77
Population 2000 Sq.	11.84	11.53	0.00	0.17	0.95	11.84	0	0.29	1	12.23	0.00	0.14	0.96
No. of Potential Cable Subscribers 2000	16.36	5.74	1.01	5.72	0.00	16.36	0	1.00	1	13.12	0.22	2.11	0.00
No. of Potential Cable Subscribers 2000 Sq.	875.80	139.28	0.57	42.33	0.00	875.80	0	4.61	1	459.82	0.20	12.99	0.00
Fraction w. HS Degree 2000	0.36	0.37	−0.18	1.20	0.00	0.36	0	1.00	1	0.36	−0.06	1.11	0.20
Fraction w. HS Degree 2000 Sq.	0.14	0.15	−0.14	1.20	0.00	0.14	0	1.03	1	0.14	−0.04	1.15	0.43
Fraction w. Some College 2000	0.26	0.26	0.05	0.97	0.14	0.26	0	1.00	1	0.26	0.04	0.91	0.36
Fraction w. Some College 2000 Sq.	0.07	0.07	0.05	0.95	0.18	0.07	0	0.87	1	0.07	0.02	0.70	0.61
Fraction w. College Degree 2000	0.22	0.19	0.33	1.31	0.00	0.22	0	1.00	1	0.21	0.09	1.13	0.05
Fraction w. College Degree 2000 Sq.	0.07	0.05	0.29	1.48	0.00	0.07	0	0.98	1	0.06	0.10	1.13	0.04
Fraction Male 2000	0.49	0.49	−0.09	0.86	0.01	0.49	0	1.00	1	0.49	0.04	1.00	0.36
Fraction Male 2000 Sq.	0.24	0.25	−0.09	0.85	0.01	0.24	0	1.07	1	0.24	0.04	1.06	0.39
Fraction Black 2000	0.03	0.03	−0.05	0.70	0.11	0.03	0	1.00	1	0.03	0.02	1.23	0.60
Fraction Black 2000 Sq.	0.01	0.01	−0.08	0.61	0.01	0.01	0	0.96	1	0.01	0.04	1.30	0.36
Fraction Hispanic 2000	0.03	0.03	0.08	0.98	0.02	0.03	0	1.00	1	0.03	0.02	0.95	0.61
Fraction Hispanic 2000 Sq.	0.01	0.01	0.01	0.82	0.86	0.01	0	1.07	1	0.01	−0.01	0.96	0.87
Fraction Employed 2000	0.61	0.61	0.02	0.96	0.53	0.61	0	1.00	1	0.60	0.18	0.97	0.00
Fraction Employed 2000 Sq.	0.38	0.38	0.02	0.96	0.60	0.38	0	1.02	1	0.37	0.19	1.03	0.00
Unemployment Rate 2000	0.05	0.05	−0.03	1.03	0.39	0.05	0	1.00	1	0.05	−0.13	0.91	0.00
Unemployment Rate 2000 Sq.	0.00	0.00	−0.01	1.71	0.88	0.00	0	1.40	1	0.00	−0.07	1.46	0.14
Fraction Married 2000	0.61	0.61	−0.04	1.19	0.32	0.61	0	1.00	1	0.61	0.00	1.02	0.93
Fraction Married 2000 Sq.	0.38	0.38	−0.02	1.17	0.54	0.38	0	1.04	1	0.38	0.00	1.02	0.96
Median Income 2000	4.44	3.99	0.35	1.47	0.00	4.44	0	1.00	1	4.25	0.13	1.22	0.01
Median Income 2000 Sq.	24.03	18.90	0.32	2.02	0.00	24.03	0	1.02	1	21.63	0.13	1.29	0.01
Fraction Urban 2000	0.54	0.37	0.53	1.08	0.00	0.54	0	1.00	1	0.53	0.03	1.00	0.54
Fraction Urban 2000 Sq.	0.49	0.32	0.54	1.18	0.00	0.49	0	1.01	1	0.48	0.03	1.02	0.50
Population 2000-1990	0.08	0.07	0.03	1.25	0.38	0.08	0	1.00	1	0.06	0.08	1.41	0.09
Population 2000-1990 Sq.	0.15	0.12	0.03	0.79	0.47	0.15	0	1.13	1	0.10	0.04	1.25	0.36
Fraction w. HS Degree 2000-1990	−0.01	0.00	−0.11	0.97	0.00	−0.01	0	1.00	1	−0.01	−0.02	1.07	0.71
Fraction w. HS Degree 2000-1990 Sq.	0.00	0.00	0.00	0.53	0.93	0.00	0	0.54	1	0.00	0.03	0.82	0.52
Fraction w. Some College 2000-1990	0.03	0.04	−0.20	0.90	0.00	0.03	0	1.00	1	0.04	−0.03	1.11	0.47
Fraction w. Some College 2000-1990 Sq.	0.00	0.00	−0.16	0.67	0.00	0.00	0	0.92	1	0.00	0.04	1.11	0.45
Fraction w. College Degree 2000-1990	0.04	0.04	0.16	1.14	0.00	0.04	0	1.00	1	0.04	0.05	1.12	0.27
Fraction w. College Degree 2000-1990 Sq.	0.00	0.00	0.15	1.40	0.00	0.00	0	1.09	1	0.00	0.08	1.20	0.11
Fraction Male 2000-1990	0.00	0.00	−0.05	0.97	0.16	0.00	0	1.00	1	0.00	0.02	1.29	0.75
Fraction Male 2000-1990 Sq.	0.00	0.00	−0.01	0.75	0.82	0.00	0	1.05	1	0.00	0.06	1.89	0.21
Fraction Black 2000-1990	0.00	0.00	−0.04	0.93	0.28	0.00	0	1.00	1	0.00	−0.01	1.23	0.86
Fraction Black 2000-1990 Sq.	0.00	0.00	−0.02	0.83	0.66	0.00	0	1.05	1	0.00	0.04	1.84	0.39
Fraction Hispanic 2000-1990	0.01	0.01	0.14	1.35	0.00	0.01	0	1.00	1	0.01	0.04	1.06	0.43
Fraction Hispanic 2000-1990 Sq.	0.00	0.00	0.10	1.63	0.02	0.00	0	1.04	1	0.00	0.02	1.33	0.63
Fraction Employed 2000-1990	0.01	0.01	−0.09	1.09	0.02	0.01	0	1.00	1	0.01	0.00	1.18	0.94
Fraction Employed 2000-1990 Sq.	0.00	0.00	0.02	1.05	0.53	0.00	0	0.98	1	0.00	0.06	1.65	0.19
Unemployment Rate 2000-1990	−0.01	−0.01	0.08	0.92	0.03	−0.01	0	1.00	1	−0.01	−0.07	1.05	0.16
Unemployment Rate 2000-1990 Sq.	0.00	0.00	−0.02	0.47	0.52	0.00	0	0.56	1	0.00	0.02	0.83	0.74
Fraction Married 2000-1990	−0.02	−0.02	0.05	0.95	0.14	−0.02	0	1.00	1	−0.02	0.05	1.11	0.29
Fraction Married 2000-1990 Sq.	0.00	0.00	−0.02	0.57	0.57	0.00	0	0.83	1	0.00	0.02	0.85	0.67
Median Income 2000-1990	1.32	1.23	0.17	1.30	0.00	1.32	0	1.00	1	1.24	0.15	1.30	0.00
Median Income 2000-1990 Sq.	2.45	2.05	0.15	1.58	0.00	2.45	0	0.98	1	2.07	0.14	1.68	0.00
Fraction Urban 2000-1990	0.08	0.08	−0.03	1.01	0.49	0.08	0	1.00	1	0.09	−0.04	1.01	0.37
Fraction Urban 2000-1990 Sq.	0.06	0.06	0.00	1.08	0.98	0.06	0	1.02	1	0.06	0.00	1.04	0.94

Note: Std. Diffs: Standardized difference in means. Var ratio: Ratio of variances. T-pval: p-value from difference of means t-test

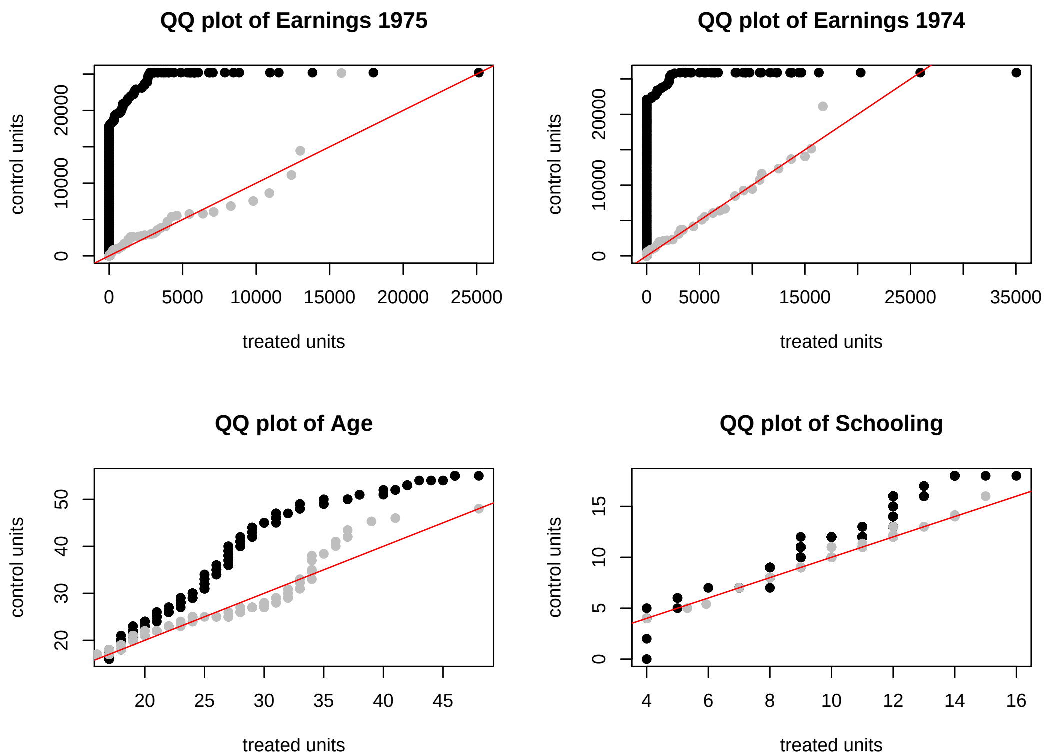
Table VIII: Covariates Balance in Conservative Candidates Data

	Raw Data					Entropy Balancing				PS Weighting			
	Means		Std.	Var	T	Mean	Std.	Var	T pval	Mean	Std.	Var	T
	Treated	Controls	Diffs	Ratio	pval	Controls	Diffs	Ratio	pval	Controls	Diffs	Var Ratio	pval
Year of Birth	1916.57	1918.34	−0.29	0.87	0.13	1916.57	0	1.12	1	1916.22	0.06	1.03	0.77
Year of Death	1995.48	1995.03	0.10	0.98	0.60	1995.48	0	0.94	1	1996.00	−0.11	0.94	0.56
Female	0.03	0.06	−0.21	0.50	0.27	0.03	0	1.00	1	0.04	−0.07	0.77	0.71
Teacher	0.02	0.05	−0.24	0.39	0.20	0.02	0	1.00	1	0.02	−0.04	0.82	0.83
Barrister	0.12	0.15	−0.11	0.85	0.55	0.12	0	1.00	1	0.10	0.11	1.22	0.56
Solicitor	0.06	0.08	−0.15	0.70	0.43	0.06	0	1.00	1	0.04	0.15	1.57	0.46
Doctor	0.02	0.03	−0.13	0.58	0.50	0.02	0	1.00	1	0.03	−0.09	0.66	0.64
Civil Servant	0.02	0.01	0.13	2.25	0.50	0.02	0	1.00	1	0.01	0.07	1.46	0.73
Local Politician	0.21	0.20	0.03	1.03	0.88	0.21	0	1.00	1	0.20	0.05	1.05	0.79
Business	0.14	0.26	−0.41	0.64	0.03	0.14	0	1.00	1	0.13	0.07	1.11	0.73
White Collar	0.11	0.14	−0.13	0.81	0.50	0.11	0	1.00	1	0.15	−0.20	0.73	0.31
Journalist	0.05	0.08	−0.21	0.59	0.27	0.05	0	1.00	1	0.04	0.09	1.33	0.65
Schooling: Eton	0.20	0.03	0.75	4.93	0.00	0.20	0	1.00	1	0.21	−0.02	0.98	0.93
Schooling: Public	0.38	0.47	−0.26	0.95	0.18	0.38	0	1.00	1	0.42	−0.10	0.97	0.62
Schooling: Regular	0.19	0.38	−0.59	0.66	0.00	0.19	0	1.00	1	0.18	0.04	1.05	0.83
Schooling: Not reported	0.22	0.11	0.42	1.76	0.03	0.22	0	1.00	1	0.19	0.10	1.10	0.63
University: Oxbridge	0.37	0.30	0.21	1.11	0.28	0.37	0	1.00	1	0.33	0.11	1.05	0.57
University: Degree	0.28	0.31	−0.11	0.94	0.57	0.28	0	1.00	1	0.33	−0.16	0.91	0.41
University: Not reported	0.36	0.39	−0.10	0.96	0.60	0.36	0	1.00	1	0.34	0.05	1.02	0.81
Aristocrat	0.10	0.02	0.49	5.22	0.01	0.10	0	1.00	1	0.07	0.13	1.32	0.51

Note: Std. Diffs: Standardized difference in means. Var ratio: Ratio of variances. T-pval: p-value from difference of means t-test

VII. Figures

Figure 1: Covariate Balance: QQ plots of Continuous Covariates



Note: QQ plots of pretreatment earnings in 1975 and 1974, age, and education. The black dots represent empirical QQ estimates for the raw data. The gray dots represent QQ estimates for the matched data. The superimposed 45-degree line indicates identical distributions for the treatment and control group.

Figure 2: Histogram for Selected Covariates in Fox News Data

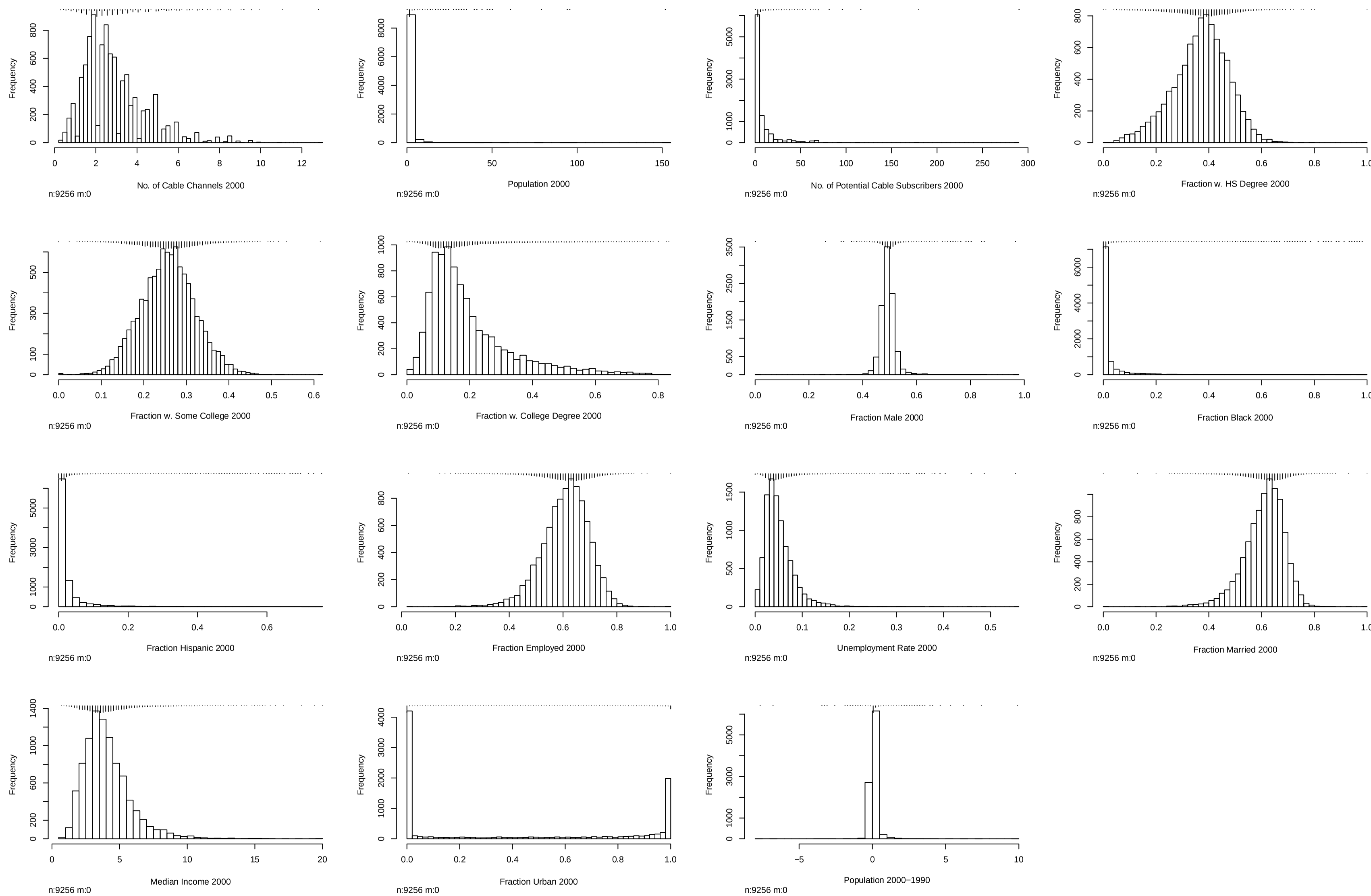
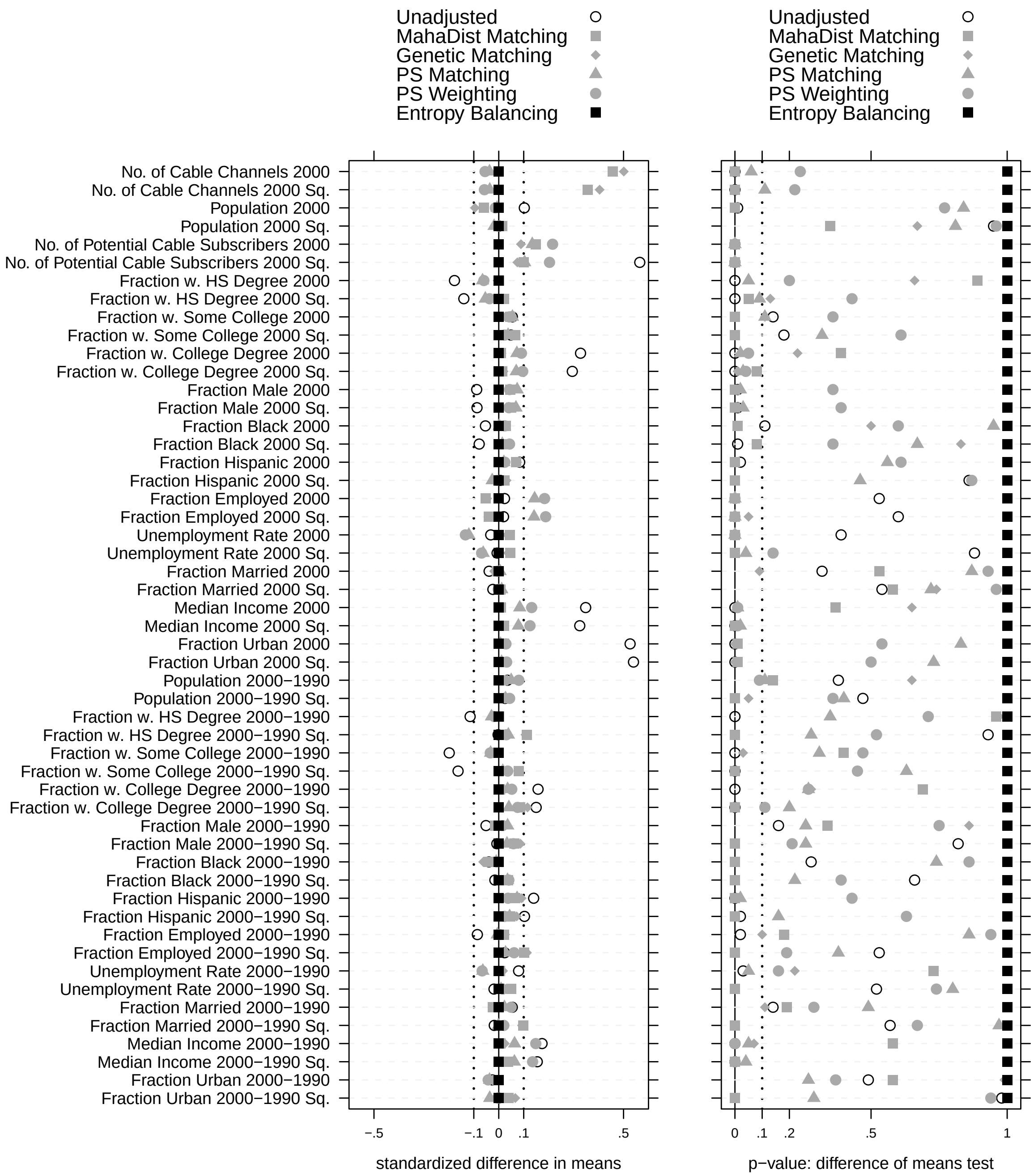
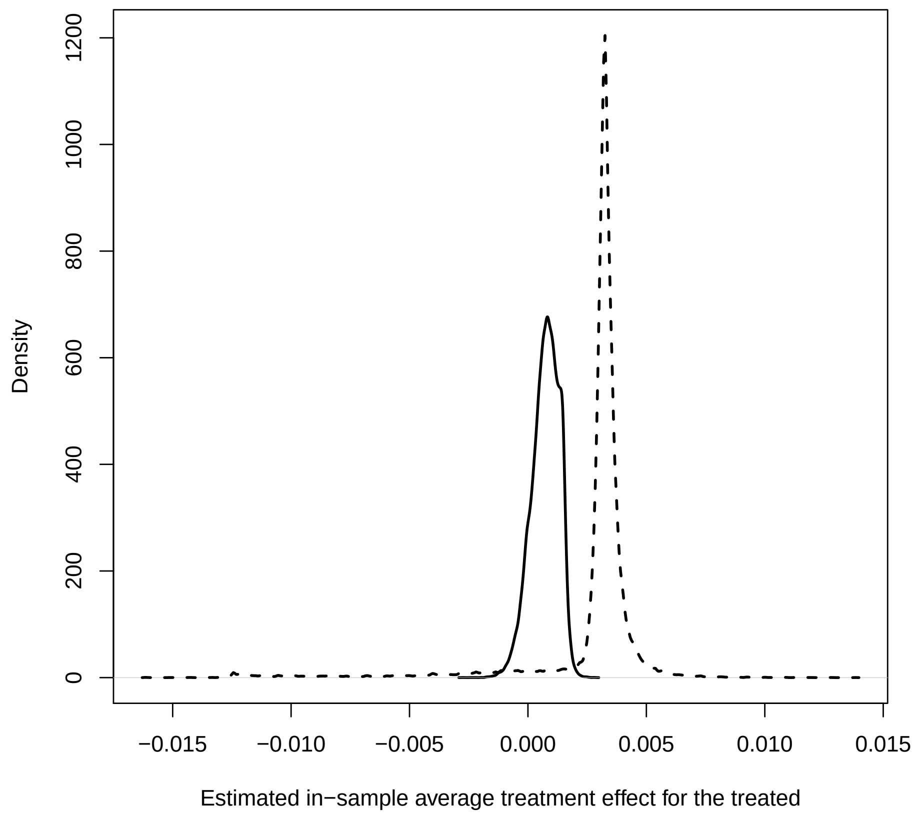


Figure 3: Covariate Balance in Fox News Data



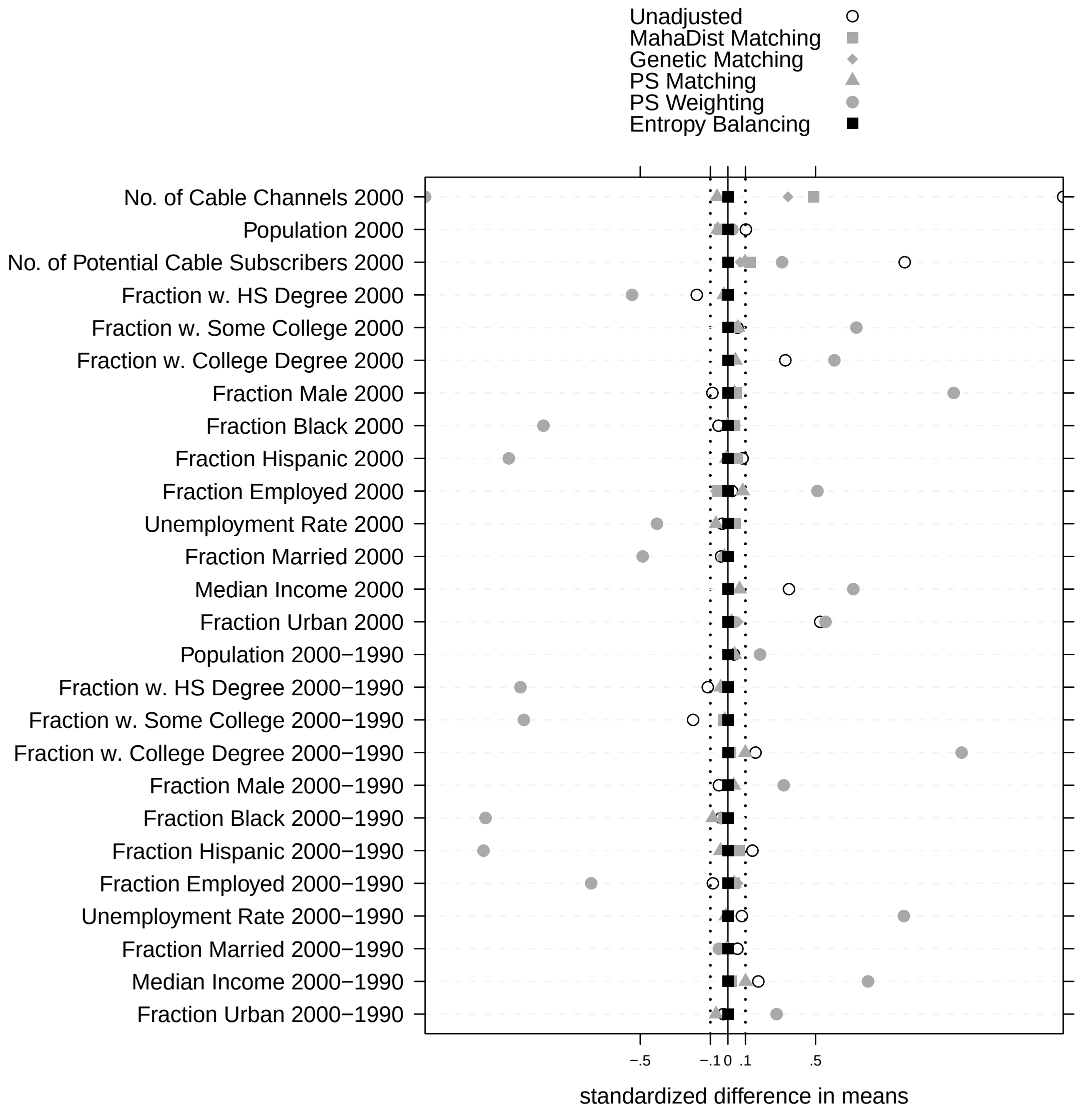
Note: Left panel shows plot of covariate-by-covariate standardized bias in the unadjusted data and after the various preprocessing methods. The standardized bias measures the difference in means between the treatment and control group (scaled by the standard deviation). Zero bias indicates identical means, dots to the right (left) of zero indicate a higher mean among the treatment (control) group. The right panel shows the p-value for a covariate-by-covariate t-test for the differences in means after the unadjusted data and after the various preprocessing methods.

Figure 4: Model Dependency in Fox News Data



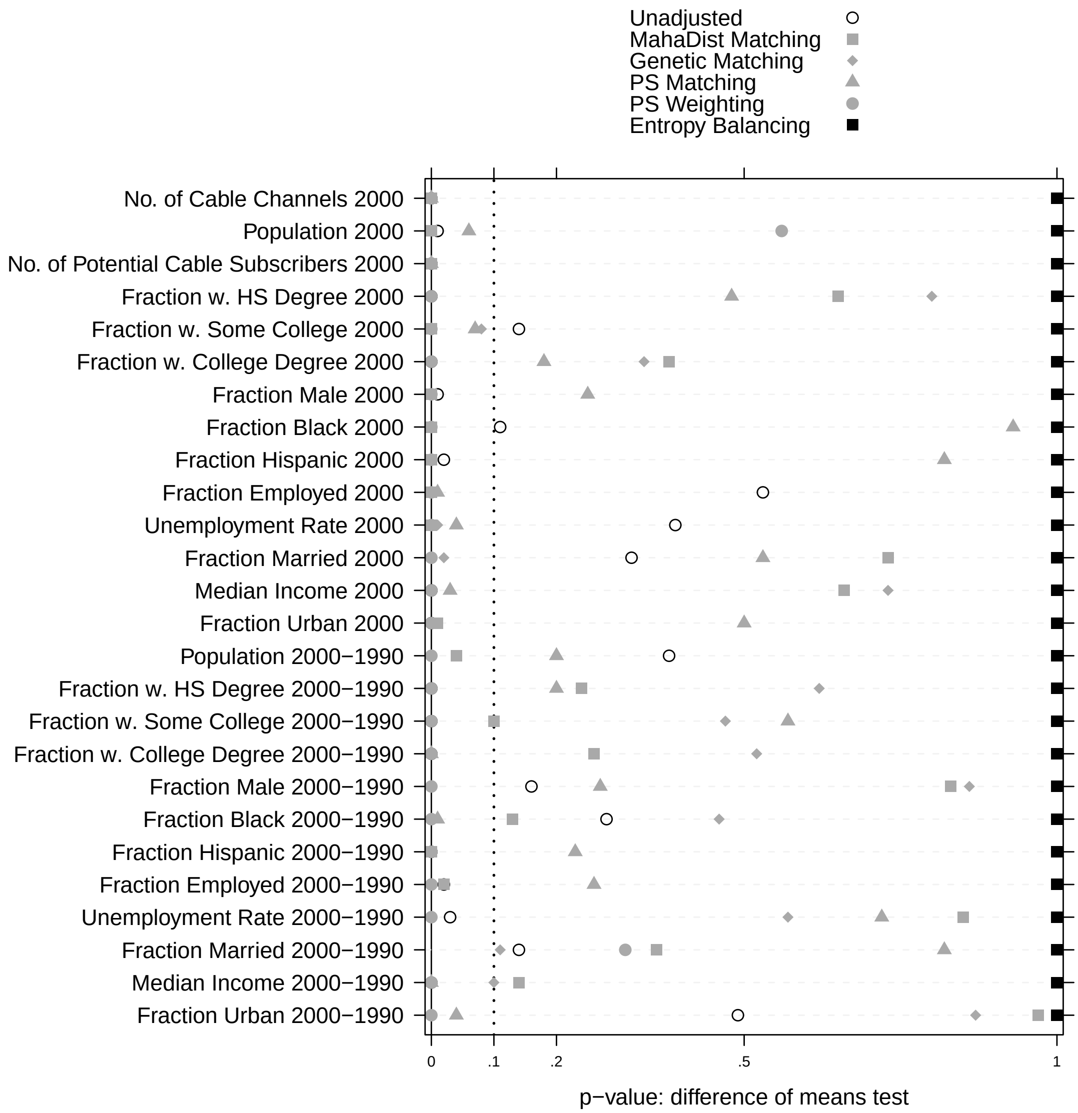
Note: Density of estimated treatment effects across one million randomly samples model specifications in the unadjusted data (dashed line) and the data preprocessed with entropy balancing (solid line).

Figure 5: Covariate Balance in Fox News Data - Using Only the Raw Covariates (Standardized Bias)



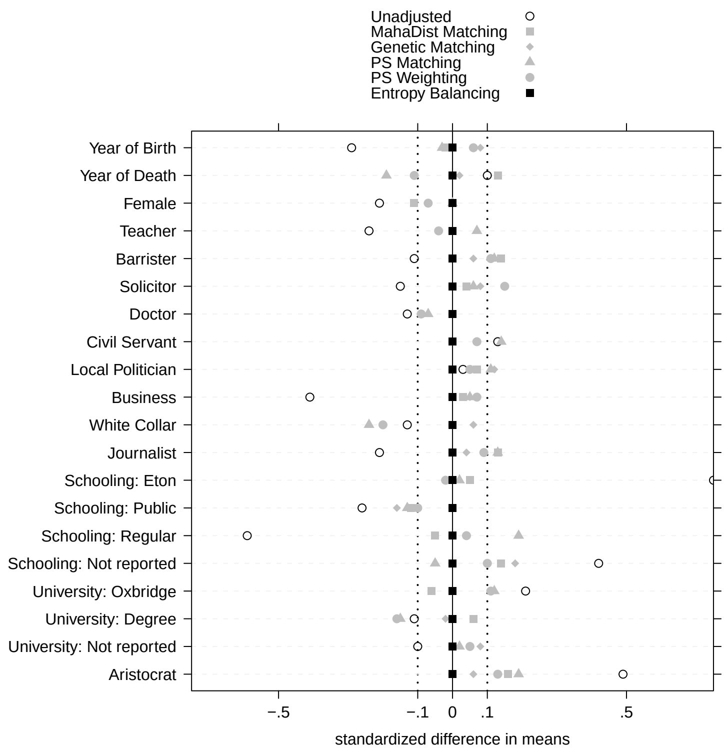
Note: Covariate-by-covariate standardized bias in the unadjusted data and after the various preprocessing methods. The standardized bias measures the difference in means between the treatment and control group (scaled by the standard deviation). Zero bias indicates identical means, dots to the right (left) of zero indicate a higher mean among the treatment (control) group.

Figure 6: Covariate Balance in Fox News Data - Using only the Raw Covariates (p-values)



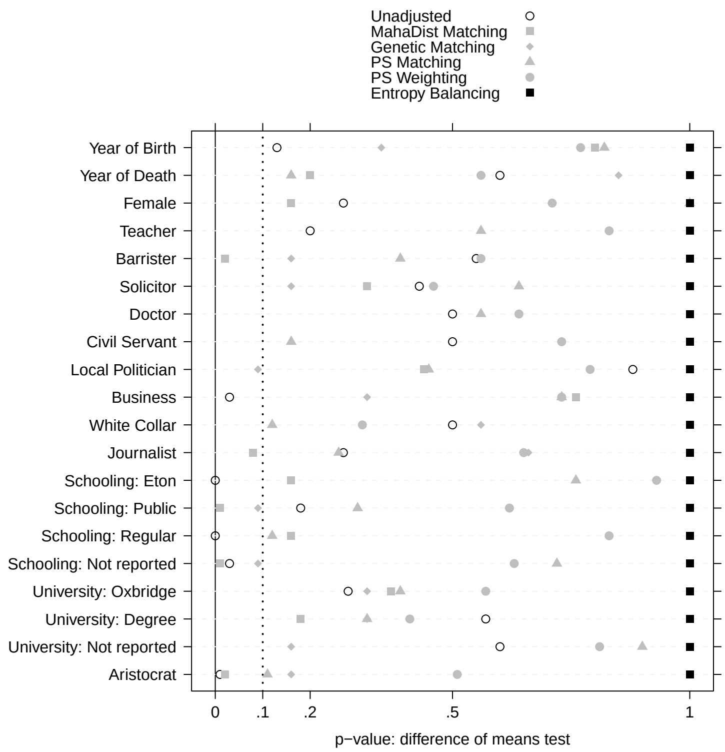
Note: p-values for a covariate-by-covariate t-test for the differences in means in the unadjusted data and after the various preprocessing methods.

Figure 7: Covariates Balance in British MPs Data (Standardized Bias)



Note: Covariate-by-covariate standardized bias in the unadjusted data and after the various preprocessing methods. The standardized bias measures the difference in means between the treatment and control group (scaled by the standard deviation). Zero bias indicates identical means, dots to the right (left) of zero indicate a higher mean among the treatment (control) group.

Figure 8: Covariates Balance in British MPs Data (p-values)



Note: p-values for a covariate-by-covariate t-test for the differences in means after the unadjusted data and after the various preprocessing methods.