POLS 5385: Causal Inference - Homework 1

João Pedro Bastos

October 30, 2023

1) Compute the treatment effect (difference in $RE78$ between randomized treated and controls). Explain what your result means.

Since in this initial dataset the treatment was randomly assigned, it precludes selection into treatment and reassures of confidence that treated and controls are balanced across both observables and unobservables. Thus, we can find the average treatment effect on the treated (ATET) by computing a simple difference in means between treated and control units. The results are presented in Table 1.

Table 1: Average Treatment Effects of a Job Training Program on Earnings

Group	Obs	Mean	Std. err.	Std. dev.	[95% conf.	interval]
Control	260	4554.801	340.0931	5483.836	3885.102	5224.501
Treated	185	6349.144	578.4229	7867.402	5207.949	7490.338
Diff. (T - C)	445	1794.342	632.8534	-	550.5745	3038.11
p-value		(0.0024)				

Note:

The estimates indicate that this particular job training training program increased the earnings of its participants (treated) by $1,794.34 relative to the controls. The difference is significant the 1% confidence level.

2) Now drop the experimental controls (260 obs where treated = 0) and add in the PSID controls in the dataset `psid_controls.dta` . Run a regression of $RE78$ on treat and the other covariates in the dataset. What is your estimated treatment effect here?

In this alternative approach where we attempt to estimate the causal effects using a simple regression, at least two important problems arise. First, it is clear that the controls are different from the treated along a variety of characteristics, which makes them poor counterfactuals.

Second, by introducing several covariates, the regression will give extra weight to those observations where treatment status is unlikely given the characteristics. This is turn will bias our estimates even more by relying precisely on the observations that are the poorest counterfactuals to our treatment group.

The results in Table 2 reveals that while the point-estimates are positive, they are statistically indistinguishable from zero, which would prevent us from rejecting the null hypothesis that this job training program has no causal effect on earnings.

Table 2: Average Treatment Effect of a Job Training Program on Earnings, Simple Regression

RE78	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
$Treat$	751.9464	915.2572	0.82	0.411	-1042.74	2546.633
$re75$	0.568	0.027	20.61	0.000	0.514	0.622
$re74$	0.278	0.028	9.96	0.000	0.223	0.333
$Married$	1240.52	586.254	2.12	0.034	90.961	2390.078
$No\ degree$	590.467	646.784	0.91	0.361	-677.783	1858.717
$Hispanic$	2163.281	1092.29	1.98	0.048	21.459	4305.104
$Education$	592.610	103.303	5.74	0.000	390.0485	795.172
$Black$	-570.928	495.178	-1.15	0.249	-1541.899	400.044
Age	-83.566	20.814	-4.01	0.000	-124.3784	-42.753
Constant	-129.743	1688.517	-0.08	0.939	-3440.679	3181.194

3) Estimate a classic diff in diff (before is RE75, after is RE78). What is your estimated treatment effect now?

In Table 3, we obtain another estimate of the causal effect through a simple differences-in-differences (DiD) approach. In this case, the ATET of the job training program is estimated as a $2,326.51 increase in earnings.

Table 3: Average Treatment Effect of a Job Training Program on Earnings, Differences-in-Differences

	Real Earnings		
	Treated	Control	Treated - Control
Before (earnings in 1975)	1,532.055	19,063.34	-17,531.285
After (earnings in 1978)	6,349.144	21,553.92	-15,204.776
ΔEarnings (1978 - 1975)	4,817.089	2,490.58	2,326.509

Again, however, there are multiple concerns with this method. The first row makes it clear that the controls units are remarkably different from the treated units, with a mean earnings in 1975 more than twelve times larger. Even so, the DiD approach technically does not require similarity between treated and control, but only the assumption of parallel trends. That is, that we expect that the gap between treated and control would have remained the same over the treatment period if not by the treatment.

Nonetheless, real earnings of the control group faced mere 3% decline from 1974 to 1975, while the treatment group experience a sharp 27% decline in earnings over the same period. While the parallel trends assumption cannot be directly tested, the fact that they were in very different trends in the year just before treatment makes it very unconvincing that it actually holds over the treatment period.

While in this case we have estimates from a randomized experiment to compare, which suggests that this approach performs surprisingly well given this circumstances, a research would be in the dark as to whether this estimates are reliable in a normal context.

4) Using `psmatch2`, estimate propensity scores using the nearest neighbor and the treatment effect (restricted to the region of common support). Explain what variables you chose and why. Use `psgraph, bin(20)` to show the distribution of the propensity scores for the treated and controls. What is it telling you? Use `pstest` to look for covariate balance. What do you find? If you do not achieve balance, go back to the functional form of your PS equation and modify until you do. Once you have balance, compute the treatment effect on the treated

Because propensity score matching uses the probability of treatment given observable characteristics as the criteria for matching, it does not assure that covariate balance is achieved. Indeed, it took several attempts until a functional form generated almost perfect covariate balance.[1] Table A1 shows that we fail to reject the null hypothesis that the sample means for all covariates are different between treated and control groups at the 5% level.[2]

The results are displayed in Table 4. While the point-estimates suggest a causal effect of $1,249.587 in additional real earnings following the job training programs, the estimate is no statistically significant.

Table 4: Average Treatment Effect of a Job Training Program on Earnings, propensity score matching with 1 nearest-neighbor

Sample	Treated	Controls	Difference	S.E.	T-stat
Unmatched	6349.144	21553.921	-15204.777	1154.614^a	-13.17
ATET	6443.625	5194.038	1249.587 (0.405)	1120.220^b	1.12

Notes: P-value in parenthesis. a: Sample standard errors. b: Bootstrapped standard errors, 250 replications.

[1] To estimate the propensity scores, I used all the available covariates plus squared terms for the continuous variables, given by: $Pr(T=1|X) = \alpha + \beta_1 Age + \beta_2 Age^2 + \beta_3 Educ + \beta_4 Edu^2 + \beta_5 RE74 + \beta_6 RE74^2 + \beta_7 RE75 + \beta_8 RE75^2 + \beta_9 Black + \beta_{10} Hispanic + \beta_{11} Married + \beta_{12} NoDegree$

[2] Only $RE74$ can be rejected at the 10% level.

Figure A1 plots the distribution of common support status by treatment assignment along different values of the propensity score. We can see that there is a clustering of controls units that are off support at very low propensity scores. These are the numerous potential controls are not similar enough to any treated unit, and so are discarded from the estimation. In total, there are 1,525 potential controls outside of common support.

By the same token, the far right of the graph shows 12 treated units that are outside of common support, which likewise are not used for the estimation because we cannot find control units that are similar enough. The final estimation relies on 173 of the 185 treated units and 965 out of 2,490 controls that in the common support group.

5) Use `ebalance` to achieve covariate balance. Then run a weighted regression of RE78 on treat. What is the estimated treatment effect now? Compare it to your answers from parts 1 and 2 above.

In Table 5, we estimate the effect again with a simple weighted regression. The results are satisfactory to the extent that they are contained in the confidence interval of the "golden-standard" estimates of Table 1, where treatment is randomly assigned, and thus are statistically indistinguishable from each other.

Table 5: Average Treatment Effect of a Job Training Program on Earnings, Weighted Regression

RE78	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
Treat	2056.32	889.4193	2.31	0.021	312.301	3800.339
Constant	4292.824	676.8883	6.34	0.000	2965.546	5620.101

Note: Standard errors obtained by bootstrapping with 250 replications.

6) Do a classic diff-in-diff using the weighted data and compare results to the discussion in question 5.

In Table 6 we re-estimate the causal effect using a DiD approach, but now using weights obtained through entropy balancing. The results are nearly identical to those obtained through the weighted regression in Table 5, suggesting that the ATET of the job training program on earnings is $ 2,058.09.

Table 6: Average Treatment Effect of a Job Training Program on Earnings, Weighted Differences-in-Differences

	Real Earnings		
	Treated	Control	Treated - Control
Before (earnings in 1975)	1,532.055	1,533.827	-17,531.285
After (earnings in 1978)	6,349.144	4,292.824	-15,204.776
ΔEarnings (1978 - 1975)	4,817.089	2,758.997	2,058.092

Note: Weights obtained by entropy balancing.

Appendix

A Figures and Tables

Table A1: Covariate Balance, Propensity Score Matching

Variable	Unmatched Matched	Mean Treated	Mean Control	% Bias	% Reduct. Bias	T-test t	T-test $p > \|t\|$
Age	U	25.816	34.851	-100.9		-11.57	0.000
	M	25.751	25.532	2.5	97.6	0.30	0.768
Age^2	U	717.39	1323.5	-97.1		-10.59	0.000
	M	717.24	692.99	3.9	96.0	0.53	0.600
Educ	U	10.346	12.117	-68.1		-7.69	0.000
	M	10.434	10.197	9.1	86.6	1.08	0.282
Educ^2	U	111.06	156.32	-78.5		-8.52	0.000
	M	112.99	108.17	8.4	89.4	1.13	0.260
*RE*74	U	2095.6	19429	-171.8		-17.50	0.000
	M	2240.9	3164.4	-9.2	94.7	-1.86	0.063
*RE*75	U	1532.1	19063	-177.4		-17.50	0.000
	M	1638.3	2059.5	-4.3	97.6	-1.30	0.196
*RE*74^2	U	2.8e+07	5.6e+08	-85.7		-8.30	0.000
	M	3.0e+07	2.7e+07	0.5	99.4	0.28	0.776
*RE*75^2	U	1.3e+07	5.5e+08	-82.9		-7.98	0.000
	M	1.4e+07	1.2e+07	0.3	99.6	0.41	0.681
Black	U	.84324	.2506	148.0		18.13	0.000
	M	.84393	.89595	-13.0	91.2	-1.44	0.151
Hispanic	U	.05946	.03253	12.9		1.94	0.053
	M	.05202	.03468	8.3	35.6	0.79	0.430
Married	U	.18919	.86627	-184.2		-25.81	0.000
	M	.20231	.15607	12.6	93.2	1.12	0.263
No Degree	U	.70811	.30522	87.9		11.49	0.000
	M	.68786	.72254	-7.6	91.4	-0.71	0.481

Sample	PS R^2	LR χ^2	$p>\chi^2$	MeanBias	MedBias	B	R
Unmatched	0.665	894.83	0.000	107.9	92.5	280.0	0.16
Matched	0.072	34.63	0.001	6.6	7.9	57.1	2.33

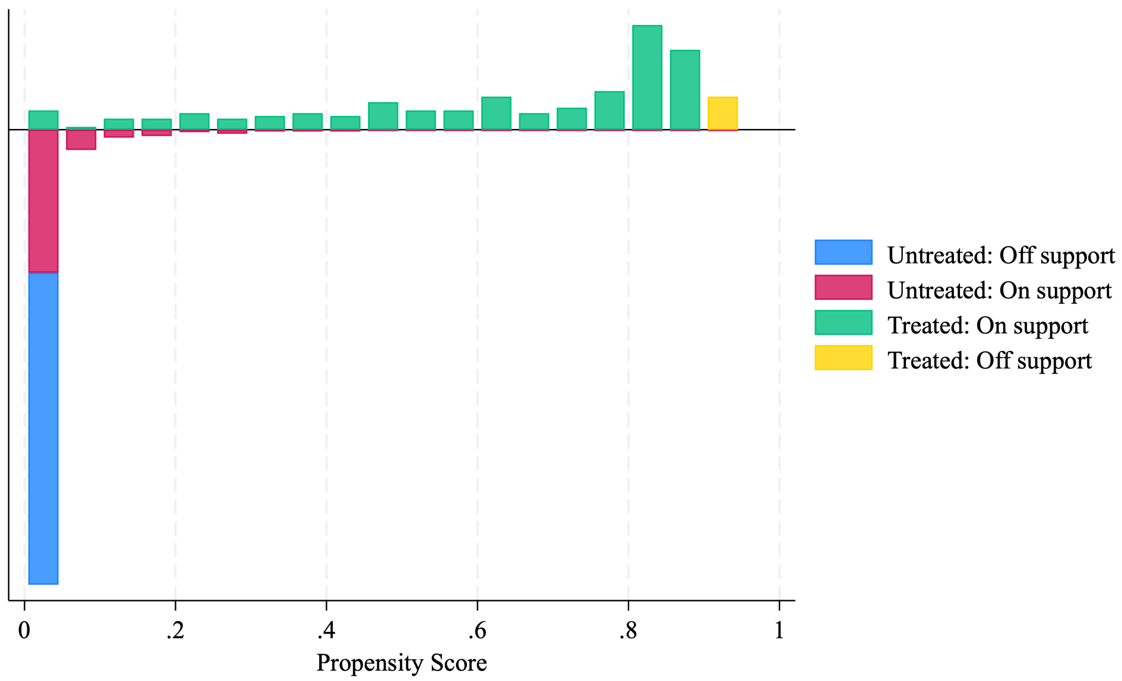

Figure A1: Distribution of common support status by propensity score levels