

Problem Set #2

Danny Edgel
Econ 717: Applied Econometrics
Spring 2022

March 8, 2022

The attached file, `edgel_ps2_log.log`, includes all console output for this problem set. Table 1, below, displays the results of the regression that determines the average treatment on the treated (ATT) using the experimental group. Despite the fact that the treatment is, by design, uncorrelated with all covariates, we may still want to include the covariates in the regression in order to account for the selection of would-be treatment-receivers across the covariates, since the dependent variable still depends on the covariates. Excluding these covariates would result in a consistent but inefficient estimate of the ATT.

Table 1: Experimental impact regresison results

VARIABLES	(1) re78
treated	824.1 (485.9)
age2	0.391 (0.569)
educ	191.0 (167.3)
black	-1,468 (734.8)
hisp	134.9 (951.5)
married	73.44 (645.0)
nodegree	-369.1 (741.6)
re74	0.0861 (0.106)
re75	0.0808 (0.119)
Constant	3,781 (2,395)

Robust standard errors in parentheses

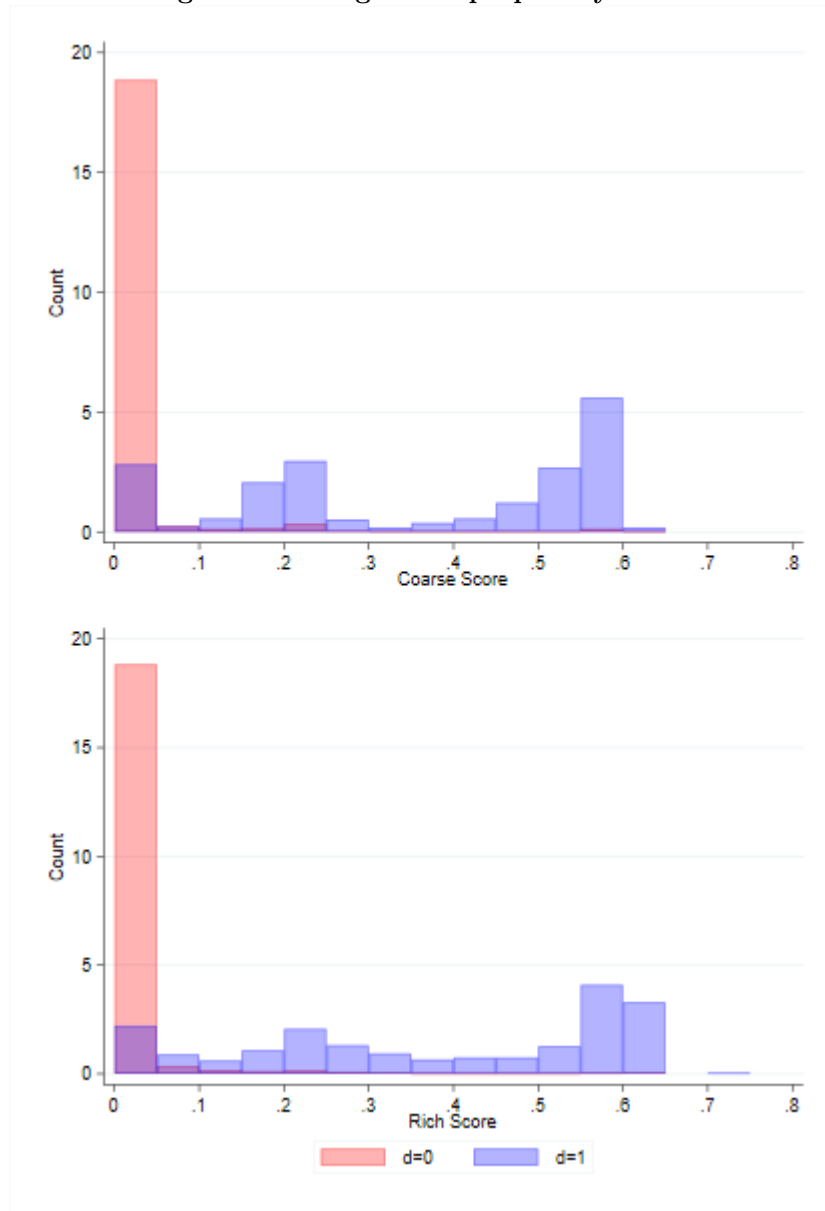
Table 2 displays the distributional statistics of the two propensity score estimates, and Figure 1 displays their histograms. The “completely determined” observations in the probit estimation are the result of one or more observations having no variation in a given set of covariate and outcome values. For example, if the regression were run on three dummy variables, and the third variable only equaled 1 for observations that were also equal in the other two dummies and were also treated, then the model would recognize them as “completely determined” at $p = 1$.

Table 2: Summary statistics of propensity scores by sample group

	CPS Untreated		Treated	
	<i>pscorea</i>	<i>pscoreb</i>	<i>pscorea</i>	<i>pscoreb</i>
Mean	0.017	0.017	0.355	0.378
Min	0.000	0.000	0.000	0.000
p5	0.000	0.000	0.017	0.009
p25	0.000	0.000	0.192	0.204
p50	0.000	0.000	0.426	0.420
p75	0.003	0.003	0.559	0.586
p95	0.060	0.069	0.587	0.624
Max	0.607	0.646	0.601	0.731

As you can see, the untreated propensity scores for the untreated observations are far more bunched near zero than the treated observations, over three quarters of which are above 0.10 for both the rich and coarse scores. Meanwhile, over 95% of the untreated observations have both rich and coarse propensity scores below 0.07. This stark difference in distributions is noticeable, as well, in the histogram of each propensity score, which are displayed in Figure 1.

Figure 1: Histogram of propensity scores



What this suggests is that a propensity score analysis would not be appropriate for this sample. Any such analysis will rely heavily on the small number of treated observations with extremely low propensity scores, and any treatment effect is likely to be statistically insignificant, due to most of the treated sample going unused for lack of a comparable control.

The non-experimental bias estimates using both estimated propensity scores and nearest-neighbor matching without replacement are not reported in any tables, but can be seen in the log file. Using `pscorea`, 0 observations, or 0.0%, are dropped, and the average treatment on the treated is -4179.6 . Using `pscoreb`, 1 observation is dropped, and the average treatment on the treated is -2261.4 . Using `pscorea` with replacement, 0 observations, or 0.00%, are dropped, and the average treatment on the treated is -3092.5 . Using `pscoreb` with replacement, 1 observation is dropped, and the average treatment on the treated is -1865.2 . Allowing comparison group observations to be reused substantially reduces bias in this context. This is due to the reality observed in the distribution statistics and plot, that there are only a handful of control observations with propensity scores above 0.10, where most of the treatment observations are. Thus, far fewer treatment observations are required to match with observations far from them on the propensity score axis in order to match when replacement is used.

The mean standardized bias reduction from matching is 98.6% and 98.8% for real earnings in 1974 and 1975, respectively. All remaining non-experimental bias estimates calculated for this problem set are displayed in Table 3. The nearest-neighbor result with replacement for the rich propensity score estimate is included at the top for comparison.

Table 3: Average treatment effect on the treated and estimated bias by estimation method, using the rich propensity score estimate

	ATT	Bias
NN w/ replacement	-1865.2	-326.3%
Gaussian Kernel		
<i>bw</i> = 0.02	-1865.2	-326.3%
<i>bw</i> = 0.2	-1865.2	-326.3%
<i>bw</i> = 2.0	-1865.2	-326.3%
Local Linear		
<i>bw</i> = 0.02	-2157.9	-361.8%
<i>bw</i> = 0.2	-2265.9	-374.9%
<i>bw</i> = 2.0	239.2	-71.0%
Linear Regression	-1936.9	-335.0%
Q12 Estimate	-1941.8	-335.6%
IPW		
<i>Scaled</i>	-1887.1	-329.0%
<i>Unscaled</i>	-1640.8	-299.1%

Using a Gaussian kernel with *any* bandwidth results in the same ATT estimate (to the first decimal place) as the nearest-neighbor result with replacement. This is also likely due to the fact that the common support condition results in some 90% of the treatment observations all being matched to a single control observation regardless of bandwidth, even when using a Gaussian kernel.

Using local linear matching, however, increasing the bandwidth meaningfully reduces the bias, ultimately changing the sign of the ATT at a bandwidth of 2.0, making it consistent with the true ATT, which does not occur with any other estimator we consider.

Estimating a simple linear regression on the control variables for the rich propensity score results in an ATT of -1936.9, which is less biased than the local linear estimate with a bandwidth at or below 0.2, but more biased than the nearest-neighbor estimate with replacement or the Gaussian kernel matching estimator. Calculating the ATT using the difference between the actual and predicted income in 1978 (based on a regression that uses only untreated observations on the rich score covariates) results in a slightly more biased ATT of -1941.8. These estimates differ due to correlation between the treatment selection and the covariates. However, the difference between the two is not statistically significant, since the standard error of the coefficient on the treatment variable is 341.3.

The ATT estimate using unscaled inverse propensity weighting (IPW) is -1887.1, which is more biased than the nearest-neighbor and Gaussian kernel

matching estimators, but the scaled IPW estimate is -1640.8, which is less biased than any other estimator from this problem set other than the local linear estimate with a bandwidth of 2.0. This is likely due to the large scale of the dependent variable, which should have been scaled or estimated in logs in the first place.