PubPol 713 Assignment 2

Nathan Mather

February 20, 2019

1 Question 1

I generated the variables for a b and c. This is shown in my code in the appendix

2 Question 2

2.1 A

I can't seem to figure out how to save tabulate output as a tex file. However, 230,653 individuals are pre-selected and 244,512 are not. Of those scoring below the 475 cutoff, 60.05 are pre-selected, 39.95 are not. Of those scoring above the cutoff, 43.92 are not pre-selected and 56.08 are pre-Selected.
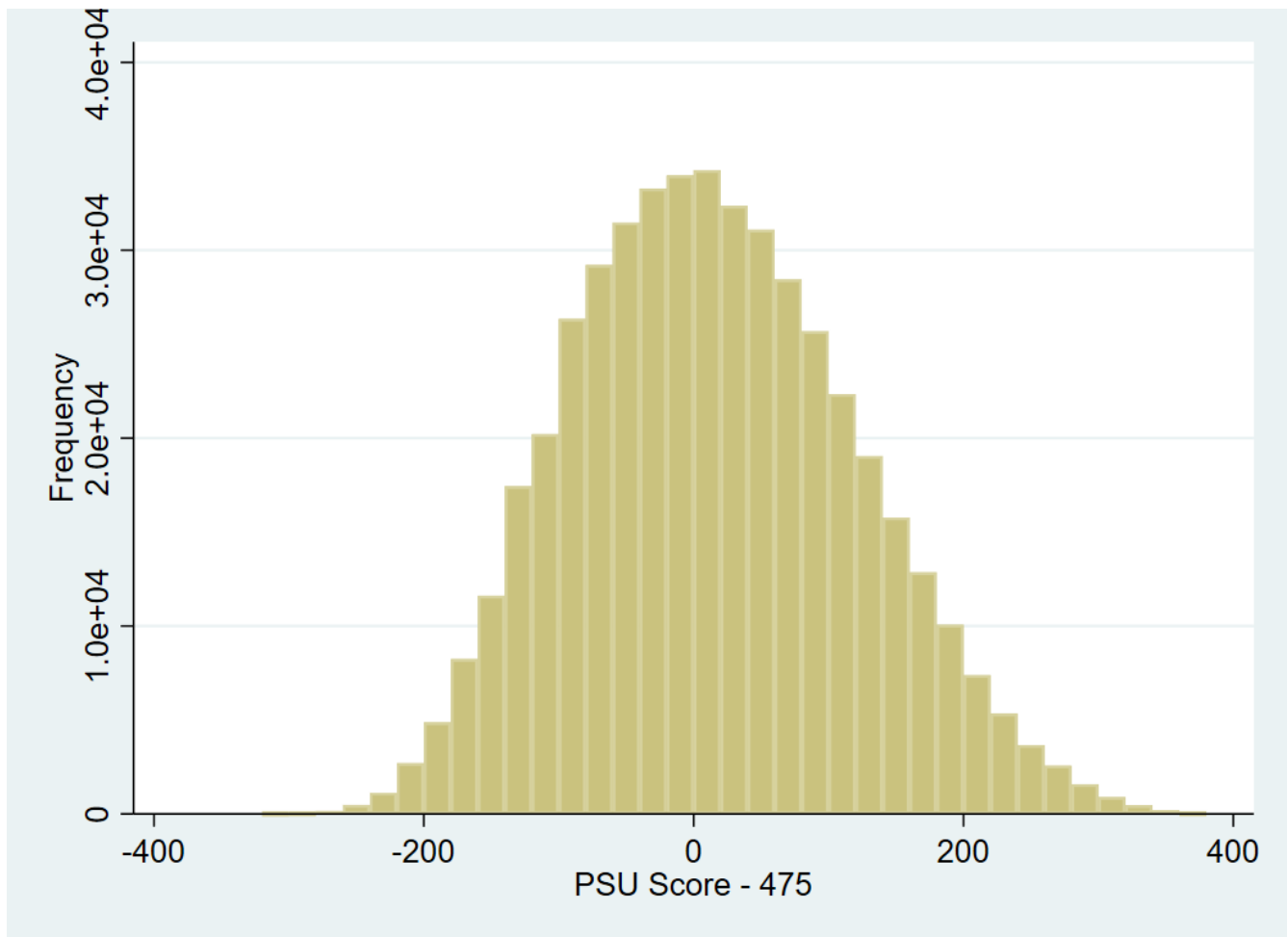The results are summarized in the tables below.

pre_sel	No.
Not Pre-Selected	244,512.0
Pre-Selected	230,653.0
Total	475,165.0

	pre_sel		
PSU Score Above 475	**Not Pre-Selected**	**Pre-Selected**	**Total**
	%	%	%
Below 475	60.1	39.9	100.0
Above 475	43.9	56.1	100.0
Total	51.5	48.5	100.0

2.2 B

The normalized score (with a PSU score of 475 equal to 0) has a minimum of -314.5, a max of 375, and a mean of 14.6. The distribution does not seem to show any bunching. A histogram is shown below.

Normalized Test Score period 1

2.3 C

The results are summarized in the tables below

	Enrolled in college in t=1		
Group	No	Yes	Total
	%	%	%
Not Pre-Sel	74.4	25.6	100.0
Pre-Sel Below	88.9	11.1	100.0
Pre-Sel Above	36.3	63.7	100.0
Total	65.7	34.3	100.0

	Ever enrolled flag		
Group	No	Yes	Total
	%	%	%
Not Pre-Sel	62.2	37.8	100.0
Pre-Sel Below	80.1	19.9	100.0
Pre-Sel Above	25.5	74.5	100.0
Total	54.5	45.5	100.0

2.4 D

The results are summarized in the tables below

Income quintile for year 1	Enrolled in college in t=1		
	No %	Yes %	Total %
1	64.7	35.3	100.0
2	54.3	45.7	100.0
3	47.7	52.3	100.0
4	43.0	57.0	100.0
5	51.2	48.8	100.0
Total	55.8	44.2	100.0

Income quintile for year 1	Ever enrolled flag		
	No %	Yes %	Total %
1	56.4	43.6	100.0
2	44.4	55.6	100.0
3	35.9	64.1	100.0
4	29.7	70.3	100.0
5	31.3	68.7	100.0
Total	44.7	55.3	100.0

3 Question 3

Below is a table showing the distribution of PSU scores by income quantile. There is also a verticle line at zero indicating the point to look for discontinuities. The distributions all appear to change smoothly across zero.
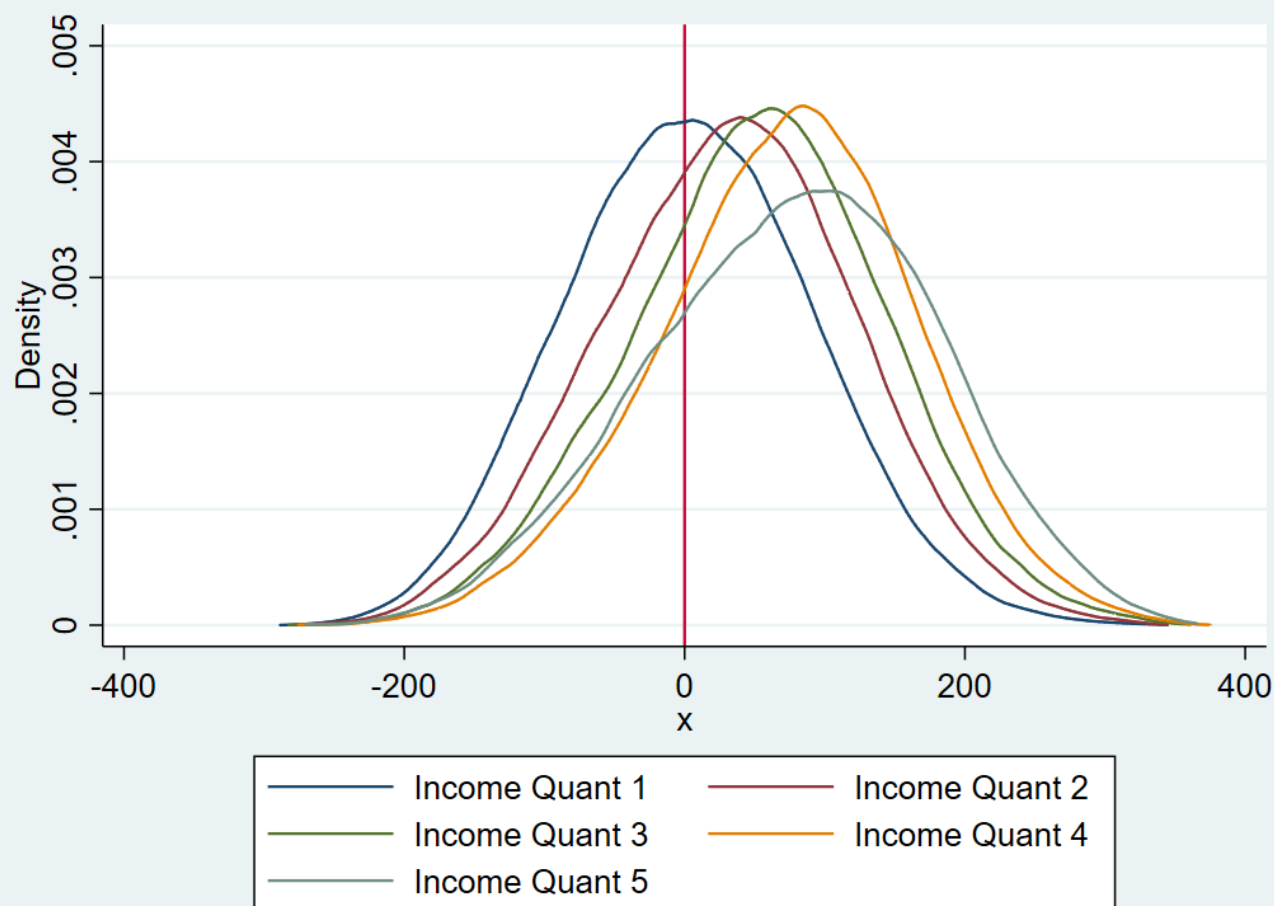
4 Question 4

Below I have the results from table three of the paper replicated. I used the same regression equation and bandwidth of 44. The main coefficient of interest here is the "PSU score Above 475". In column 1 this coefficient shows that, for pre-selected students, being above the cutoff implies an increase of 17.5 percentage points in the probability of enrolling in college immediately after the test. One potential threat to interpreting this relationship as causal is that passing the 475 threshold may provide some benefit other than loan eligibility. An example discussed in the paper is higher probability of acceptance to schools if students score above 475. Column two tests this possibility with a placebo test on non-selected students. Here be find no significant effect for being above the 475 mark. This is what we would expect since these students are not eligible for loans anyway.

Table 3 Replication

	(1)	(2)
PSU Score Above 475	0.175***	0.00273
	(0.00611)	(0.00556)
PSU Score - 475	0.00160***	0.00163***
	(0.000147)	(0.000133)
PSU Score if Above 475	0.00222***	0.000866***
	(0.000238)	(0.000223)
Constant	0.182***	0.159***
	(0.00387)	(0.00367)
Bandwidth	44	44

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

In addition to the straight replication, I also ran these regressions using an updated bandwidth selection technique and local linear regression (Calonico, Cattaneo, & Titiunik, 2014). The results are similar and can be found in the table below. "RD_Estimate" is comparable to "PSU Score Above 475" in the table above.

Table 3 with updated methods

	(1b)	(2b)
RD_Estimate	0.176***	0.00281
	(0.00668)	(0.00569)
Bandwidth	44.61	50.29

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

References

Calonico, S., Cattaneo, M. D., & Titiunik, R. (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, *82*(6), 2295-2326. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA11757