ECON 7103 Homework 6

Yifan Liu (yliu3494)

Spring 2023

1 Python

---

1. RD

It should be a sharp RD.
The sharp RD ensures that the running variable completely determines the treatment, while the fuzzy RD appears when the threshold merely discontinuously increase the probability of treatment.
In this case, the policy requires all vehicles longer than 225 inches must be equipped with the specific safety technology.

2. Scatter plot

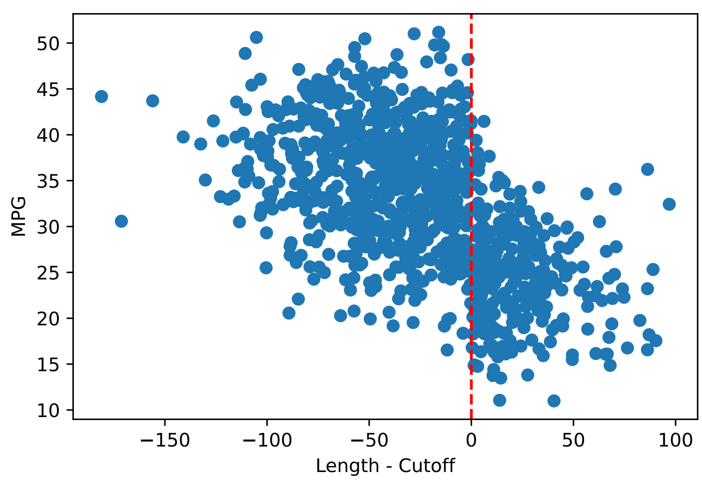

Figure 1: Scatter plot of mpg and $length - cutoff$ with a line at the RD cutoff

Figure 1 demonstrates a scatter plot with mpg on the y-axis and $length - cutoff$ on the x-axis with a line at the RD cutoff.
As Figure 1 shows, there is visual evidence of bunching. There is also visual evidence of an discontinuity above and below the cutoff.

3. First-order polynomial

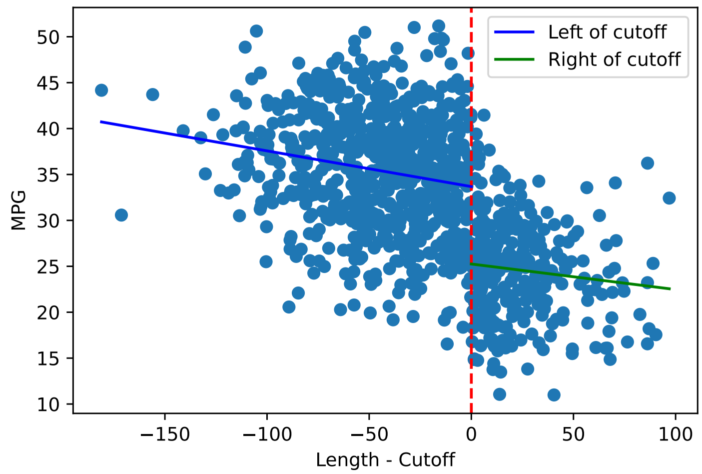

Figure 2: First-order Polynomial

Figure 2 above shows the resulting first-order polynomial over a scatter plot.
Table 1 and Table 2 below show the first-order polynomial regression results below and above the cutoff, respectively.
$effect = rightmodel.params[1] - leftmodel.params[1] = -0.0278 - (-0.0389) = 0.0111$
The first-stage treatment effect estimate is 0.0111.
$yright[0] - yleft[-1]$
The difference at the cutoff is -8.42.

Dep. Variable:	mpg	**R-squared:**	0.033
Model:	OLS	**Adj. R-squared:**	0.032
Method:	Least Squares	**F-statistic:**	24.88
Date:	Mon, 27 Feb 2023	**Prob (F-statistic):**	7.65e-07
Time:	17:34:01	**Log-Likelihood:**	-2355.2
No. Observations:	720	**AIC:**	4714.
Df Residuals:	718	**BIC:**	4724.
Df Model:	1		
Covariance Type:	nonrobust		

	coef	**std err**	**t**	**P>|t|**	**[0.025**	**0.975]**
const	33.6800	0.425	79.181	0.000	32.845	34.515
length-cutoff	-0.0389	0.008	-4.988	0.000	-0.054	-0.024

Omnibus:	8.531	**Durbin-Watson:**	1.393
Prob(Omnibus):	0.014	**Jarque-Bera (JB):**	6.180
Skew:	-0.102	**Prob(JB):**	0.0455
Kurtosis:	2.594	**Cond. No.**	97.6

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

Table 1: First-order polynomial to the left side of the cutoff in a RDD

Dep. Variable:	mpg	**R-squared:**	0.011
Model:	OLS	**Adj. R-squared:**	0.008
Method:	Least Squares	**F-statistic:**	3.184
Date:	Mon, 27 Feb 2023	**Prob (F-statistic):**	0.0755
Time:	17:34:01	**Log-Likelihood:**	-875.41
No. Observations:	280	**AIC:**	1755.
Df Residuals:	278	**BIC:**	1762.
Df Model:	1		
Covariance Type:	nonrobust		

	coef	**std err**	**t**	**P>|t|**	**[0.025**	**0.975]**
const	25.2510	0.514	49.101	0.000	24.239	26.263
length-cutoff	-0.0278	0.016	-1.784	0.075	-0.059	0.003

Omnibus:	1.162	**Durbin-Watson:**	1.619
Prob(Omnibus):	0.559	**Jarque-Bera (JB):**	1.264
Skew:	0.124	**Prob(JB):**	0.532
Kurtosis:	2.785	**Cond. No.**	51.3

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

Table 2: First-order polynomial to the right side of the cutoff in a RD

4. Second-order polynomial

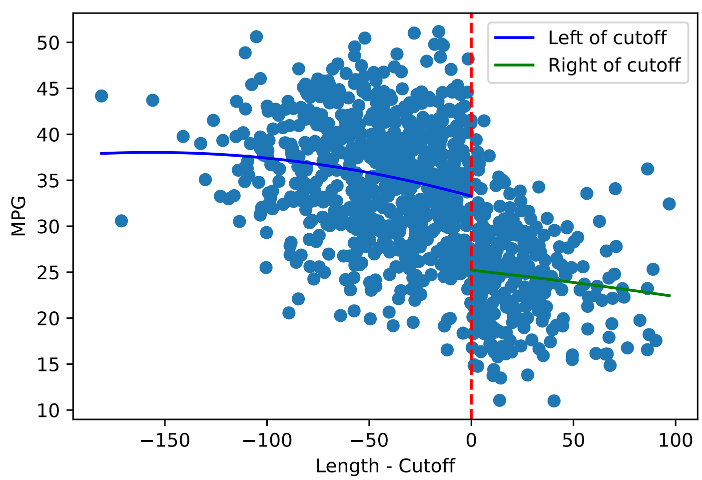

Figure 3: Second-order Polynomial

Figure 3 shows the resulting second-order polynomial over a scatter plot.

Dep. Variable:	mpg	**R-squared:**	0.035
Model:	OLS	**Adj. R-squared:**	0.032
Method:	Least Squares	**F-statistic:**	12.99
Date:	Mon, 27 Feb 2023	**Prob (F-statistic):**	2.86e-06
Time:	18:34:21	**Log-Likelihood:**	-2354.7
No. Observations:	720	**AIC:**	4715.
Df Residuals:	717	**BIC:**	4729.
Df Model:	2		
Covariance Type:	nonrobust		

	coef	**std err**	**t**	**P>|t|**	**[0.025**	**0.975]**
const	33.2671	0.579	57.444	0.000	32.130	34.404
x1	-0.0608	0.022	-2.731	0.006	-0.105	-0.017
x2	-0.0002	0.000	-1.051	0.294	-0.001	0.000

Omnibus:	7.964	**Durbin-Watson:**	1.393
Prob(Omnibus):	0.019	**Jarque-Bera (JB):**	5.920
Skew:	-0.105	**Prob(JB):**	0.0518
Kurtosis:	2.608	**Cond. No.**	1.15e+04

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 1.15e+04. This might indicate that there are strong multicollinearity or other numerical problems.

Table 3: Second-order polynomial to the left side of the cutoff in a RDD

Dep. Variable:	mpg	**R-squared:**	0.011
Model:	OLS	**Adj. R-squared:**	0.004
Method:	Least Squares	**F-statistic:**	1.588
Date:	Mon, 27 Feb 2023	**Prob (F-statistic):**	0.206
Time:	18:34:21	**Log-Likelihood:**	-875.41
No. Observations:	280	**AIC:**	1757.
Df Residuals:	277	**BIC:**	1768.
Df Model:	2		
Covariance Type:	nonrobust		

	coef	**std err**	**t**	**P>|t|**	**[0.025**	**0.975]**
const	25.2194	0.713	35.390	0.000	23.817	26.622
x1	-0.0249	0.048	-0.519	0.604	-0.119	0.070
x2	-3.869e-05	0.001	-0.064	0.949	-0.001	0.001

Omnibus:	1.160	**Durbin-Watson:**	1.620
Prob(Omnibus):	0.560	**Jarque-Bera (JB):**	1.264
Skew:	0.127	**Prob(JB):**	0.532
Kurtosis:	2.792	**Cond. No.**	4.33e+03

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 4.33e+03. This might indicate that there are strong multicollinearity or other numerical problems.

Table 4: Second-order polynomial to the right side of the cutoff in a RDD

Table 3 and Table 4 show the second-order polynomial regression results below and above the cutoff, respectively.
$effect = yright[0] - yleft[-1]$
The difference at the cutoff (not sure if it is the treatment effect estimate) is -8.05.

5. Fifth-order polynomial

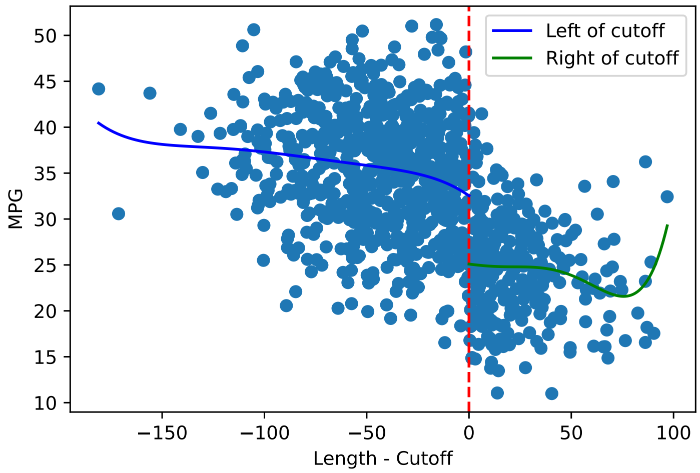

Figure 4: Fifth-order Polynomial

Figure 4 shows the resulting fifth-order polynomial over a scatter plot.

Dep. Variable:	mpg	**R-squared:**	0.036
Model:	OLS	**Adj. R-squared:**	0.029
Method:	Least Squares	**F-statistic:**	5.350
Date:	Mon, 27 Feb 2023	**Prob (F-statistic):**	7.74e-05
Time:	18:34:21	**Log-Likelihood:**	-2354.2
No. Observations:	720	**AIC:**	4720.
Df Residuals:	714	**BIC:**	4748.
Df Model:	5		
Covariance Type:	nonrobust		

	coef	**std err**	**t**	**P> |t|**	**[0.025**	**0.975]**
const	32.5318	1.088	29.911	0.000	30.396	34.667
x1	-0.1577	0.147	-1.076	0.282	-0.445	0.130
x2	-0.0035	0.006	-0.576	0.565	-0.015	0.008
x3	-4.473e-05	9.99e-05	-0.448	0.655	-0.000	0.000
x4	-2.738e-07	7.1e-07	-0.386	0.700	-1.67e-06	1.12e-06
x5	-6.247e-10	1.77e-09	-0.352	0.725	-4.11e-09	2.86e-09

Omnibus:	7.687	**Durbin-Watson:**	1.393
Prob(Omnibus):	0.021	**Jarque-Bera (JB):**	5.769
Skew:	-0.104	**Prob(JB):**	0.0559
Kurtosis:	2.614	**Cond. No.**	5.04e+10

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 5.04e+10. This might indicate that there are strong multicollinearity or other numerical problems.

Table 5: Fifth-order polynomial to the left side of the cutoff in a RDD

Dep. Variable:	mpg	**R-squared:**	0.023
Model:	OLS	**Adj. R-squared:**	0.005
Method:	Least Squares	**F-statistic:**	1.265
Date:	Mon, 27 Feb 2023	**Prob (F-statistic):**	0.279
Time:	18:34:21	**Log-Likelihood:**	-873.81
No. Observations:	280	**AIC:**	1760.
Df Residuals:	274	**BIC:**	1781.
Df Model:	5		
Covariance Type:	nonrobust		

	coef	**std err**	**t**	**P>|t|**	**[0.025**	**0.975]**
const	25.0943	1.308	19.184	0.000	22.519	27.669
x1	-0.0267	0.323	-0.083	0.934	-0.663	0.609
x2	-2.09e-05	0.024	-0.001	0.999	-0.047	0.047
x3	6.545e-05	0.001	0.092	0.926	-0.001	0.001
x4	-2.013e-06	8.95e-06	-0.225	0.822	-1.96e-05	1.56e-05
x5	1.462e-08	4.01e-08	0.364	0.716	-6.44e-08	9.36e-08

Omnibus:	1.030	**Durbin-Watson:**	1.626
Prob(Omnibus):	0.597	**Jarque-Bera (JB):**	1.109
Skew:	0.140	**Prob(JB):**	0.574
Kurtosis:	2.871	**Cond. No.**	4.16e+09

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 4.16e+09. This might indicate that there are strong multicollinearity or other numerical problems.

Table 6: Fifth-order polynomial to the right side of the cutoff in a RDD

Table 5 and Table 6 show the fifth-order polynomial regression results below and above the cutoff, respectively.
$effect = yright[0] - yleft[-1]$
The difference at the cutoff (not sure if it is the treatment effect estimate) is -4.17.

6. 2SLS using the discontinuity as an instrument for *mpg*
I use the discontinuity as an instrument for *mpg* in two ways: a continuous variable $length - cutoff$ and a binary variable *policy* where 1 means the vehicle is longer than 225 and 0 for otherwise.

Here are the results respectively (see the code for details):
a continuous variable $length - cutoff$: the average treatment effect is 162.43; in other words, one unit increase in *mpg* is expected to increase the vehicle's sale price by 162.43 units, holding all other variables constant.
a binary variable *policy*: the average treatment effect is 158.28; in other words, one unit increase in *mpg* is expected to increase the vehicle's sale price by 158.28 units, holding all other variables constant.

2 Stata

---

1. (a)

	(1)
VARIABLES	
RD_Estimate	-7.78**
	(1.68)
Observations	1,000

Standard errors in parentheses
** p<0.01, * p<0.05

Table 7: First-stage regression results using the discontinuity as the instrument

I had difficulty in obtaining the predicted value from rdrobust command in Stata. If we can get the predicted y, we can put it to the second-stage regression.

1. (b)

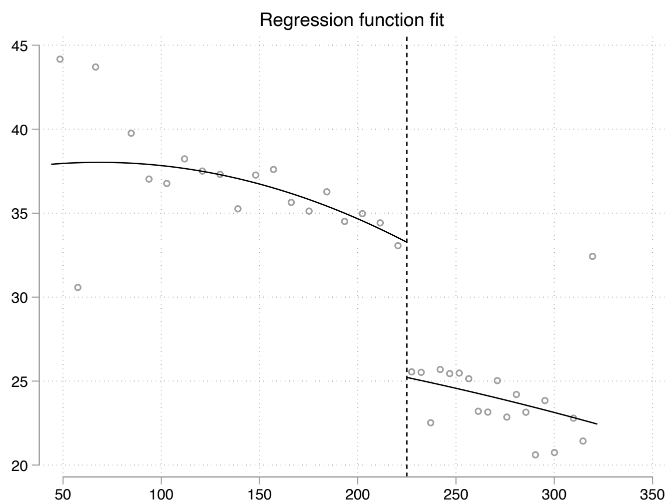

Figure 5: A plot of the results using rdplot

2.
To the best of my current understanding, I believe it is a valid instrument to some degree. 1) It causes variation in the treatment variable. With the strange policy, every vehicle equipped with the technology is significantly less fuel-efficient. 2) However, it might have a direct effect on the outcome variable *price* through other mechanisms than *mpg*. For example, a longer/bigger vehicle tend to be safer and prettier and cost more materials to produce. The vehicle is therefore more expensive.