

# 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

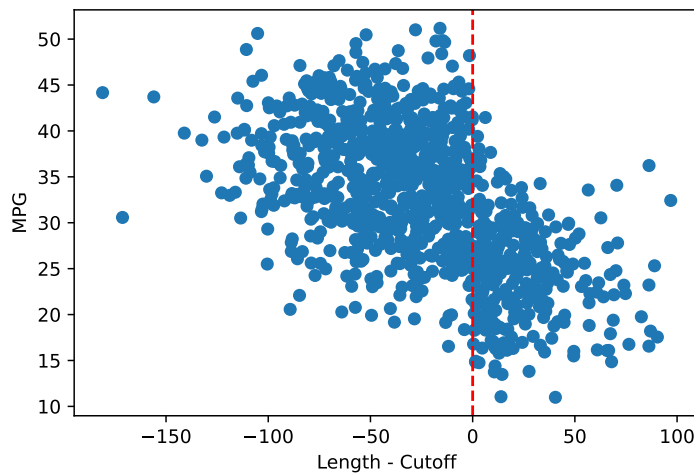


Figure 1: Scatter plot of  $mpg$  and  $length - cutoff$  with a line at the RD cutoff

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

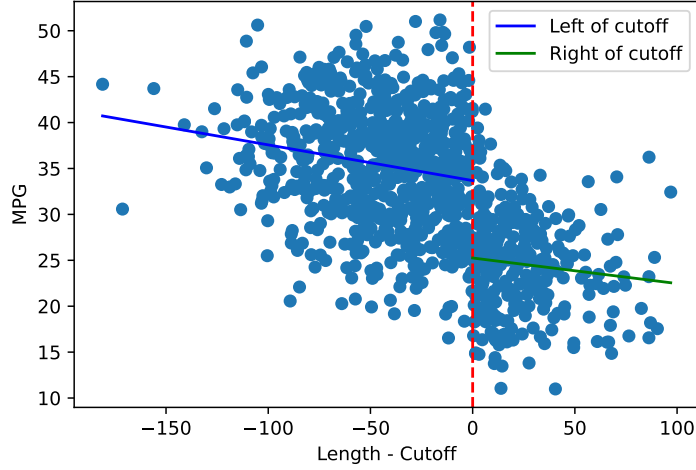


Figure 2: First-order Polynomial

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.

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

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

<b>Omnibus:</b>	8.531	<b>Durbin-Watson:</b>	1.393
<b>Prob(Omnibus):</b>	0.014	<b>Jarque-Bera (JB):</b>	6.180
<b>Skew:</b>	-0.102	<b>Prob(JB):</b>	0.0455
<b>Kurtosis:</b>	2.594	<b>Cond. No.</b>	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

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

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

<b>Omnibus:</b>	1.162	<b>Durbin-Watson:</b>	1.619
<b>Prob(Omnibus):</b>	0.559	<b>Jarque-Bera (JB):</b>	1.264
<b>Skew:</b>	0.124	<b>Prob(JB):</b>	0.532
<b>Kurtosis:</b>	2.785	<b>Cond. No.</b>	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

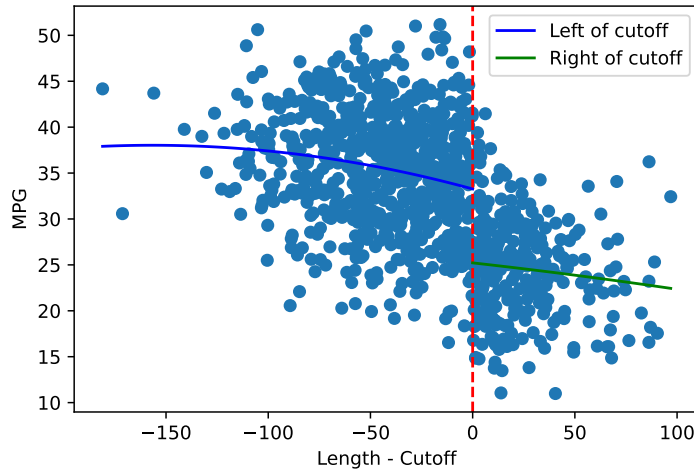


Figure 3: Second-order Polynomial

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

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

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

<b>Omnibus:</b>	7.964	<b>Durbin-Watson:</b>	1.393
<b>Prob(Omnibus):</b>	0.019	<b>Jarque-Bera (JB):</b>	5.920
<b>Skew:</b>	-0.105	<b>Prob(JB):</b>	0.0518
<b>Kurtosis:</b>	2.608	<b>Cond. No.</b>	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

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

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

<b>Omnibus:</b>	1.160	<b>Durbin-Watson:</b>	1.620
<b>Prob(Omnibus):</b>	0.560	<b>Jarque-Bera (JB):</b>	1.264
<b>Skew:</b>	0.127	<b>Prob(JB):</b>	0.532
<b>Kurtosis:</b>	2.792	<b>Cond. No.</b>	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.

$$effect = y_{right}[0] - y_{left}[-1]$$

The difference at the cutoff (not sure if it is the treatment effect estimate) is -8.05.

## 5. Fifth-order polynomial

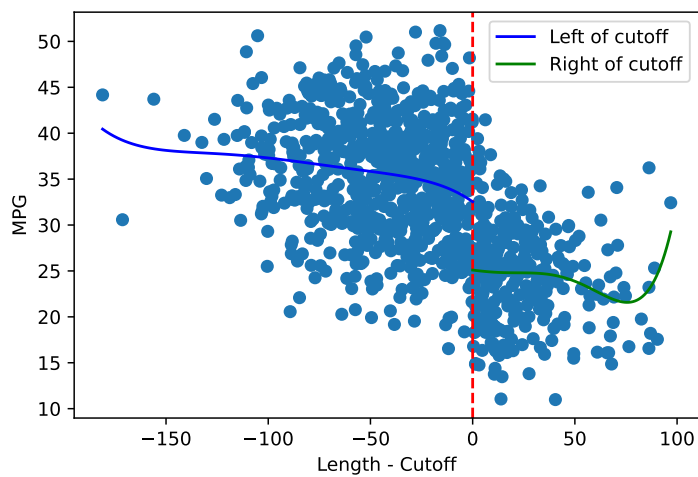


Figure 4: Fifth-order Polynomial

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

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

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

<b>Omnibus:</b>	7.687	<b>Durbin-Watson:</b>	1.393
<b>Prob(Omnibus):</b>	0.021	<b>Jarque-Bera (JB):</b>	5.769
<b>Skew:</b>	-0.104	<b>Prob(JB):</b>	0.0559
<b>Kurtosis:</b>	2.614	<b>Cond. No.</b>	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

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

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

<b>Omnibus:</b>	1.030	<b>Durbin-Watson:</b>	1.626
<b>Prob(Omnibus):</b>	0.597	<b>Jarque-Bera (JB):</b>	1.109
<b>Skew:</b>	0.140	<b>Prob(JB):</b>	0.574
<b>Kurtosis:</b>	2.871	<b>Cond. No.</b>	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.

$$effect = y_{right}[0] - y_{left}[-1]$$

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

- (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

	(1)
VARIABLES	
rdplot_hat_y	135.41** (22.21)
car	-3,693.27** (222.36)
Constant	17,622.91** (711.27)
Observations	1,000
R-squared	0.22
Standard errors in parentheses	
** p<0.01, * p<0.05	

Table 8: Second-stage regression results using the discontinuity as the instrument

1. (b)

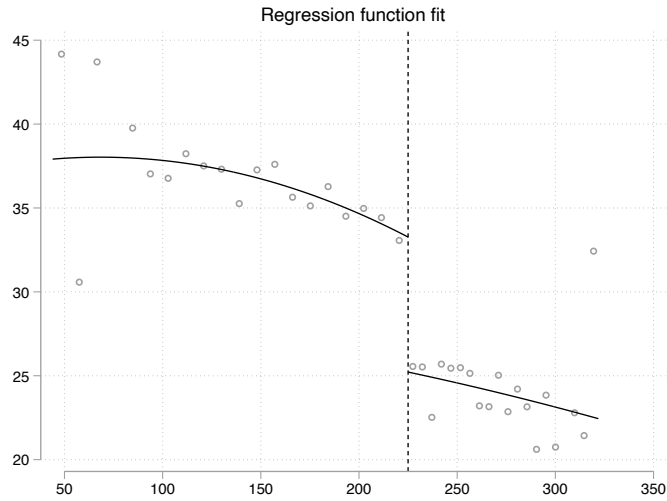


Figure 5: A plot of the results using rdplot



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.