

# Homework 7

Economics 7103

March 5, 2024

## 1 Python

1. If I understood correctly, this policy is directed towards manufacturers. Since we have data on car sales in 2017, I believe it most likely includes already assembled cars that were made before the policy came into force. That's why I think there will still be cars longer than 225 inches that do not have specific safety technology. Additionally, it takes time to readjust to such manufacturing requirements, and governments usually provide some transition time before the policy is fully implemented. So, I expect fuzzy RD.

2-5. The scatter plot of fuel efficiency in miles per gallon versus length is presented in Figure 1. There is bunching around the cutoff on both sides. It looks like a fuzzy RD, but I am not fully sure. However, definitely it is not a sharp RD. The fuel efficiency of the cars above the cutoff is smaller, as was expected by the new policy, but I suspect that long cars might have less fuel efficiency even without new policy-required technology. I think we should conduct robustness checks and McCrary test to confirm that this is proper fuzzy RD. The first-order polynomial fitted line is added to the scatter plot in Figure 2, the second-order in Figure 3 and the fifth-order in Figure 4. The first-stage treatment effect estimates are presented in Table 1. In the first column, the first-stage treatment effect is presented from the regression including a first-level polynomial. In the second column, results from the model with a second-level polynomial are displayed, and in the third column, the model with fifth-level polynomials is presented. The results from the first model, including the first-level polynomial, are statistically significant. Therefore, it is more reasonable to use only first-level fitting. In other cases, we face an overfitting issue.

	<i>Dependent variable: mpg</i>		
	(1)	(2)	(3)
Treatment effect	-10.92** (4.50)	-8.25 (46.30)	0.38 (0.34)
Observations	1000	1000	1000
$R^2$	0.40	0.40	0.40
Adjusted $R^2$	0.40	0.40	0.40
Residual Std. Error	6.16	6.16	6.17
F Statistic	259.65***	162.43***	248.22***

*Note:* \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

Table 1: RD regression output

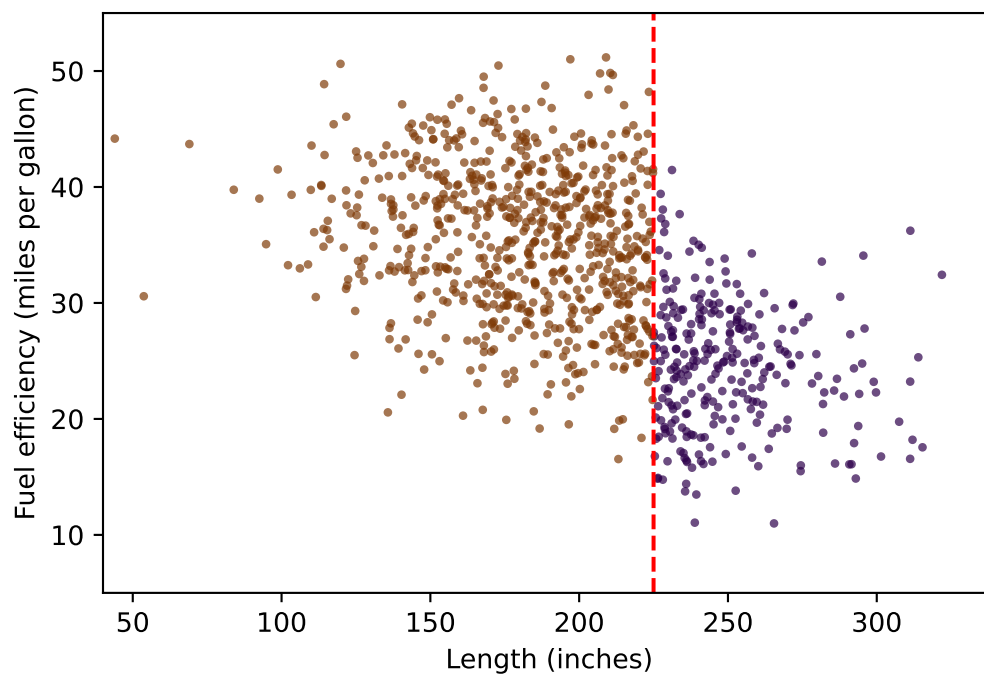


Figure 1: Scatter plot of Fuel efficiency in miles per gallon & length

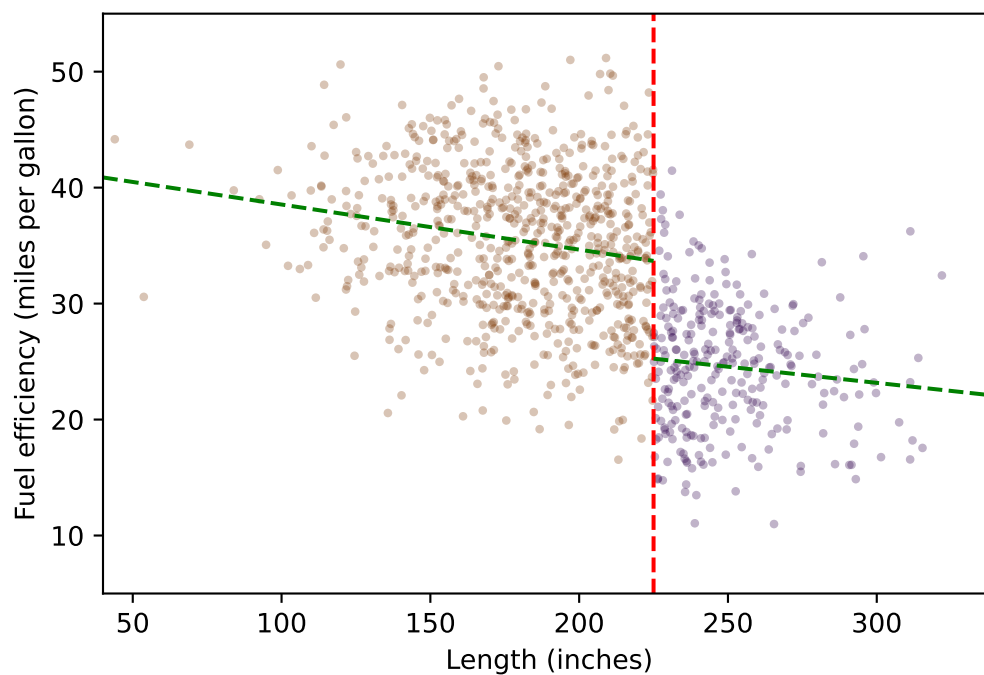


Figure 2: The first order polynomial.

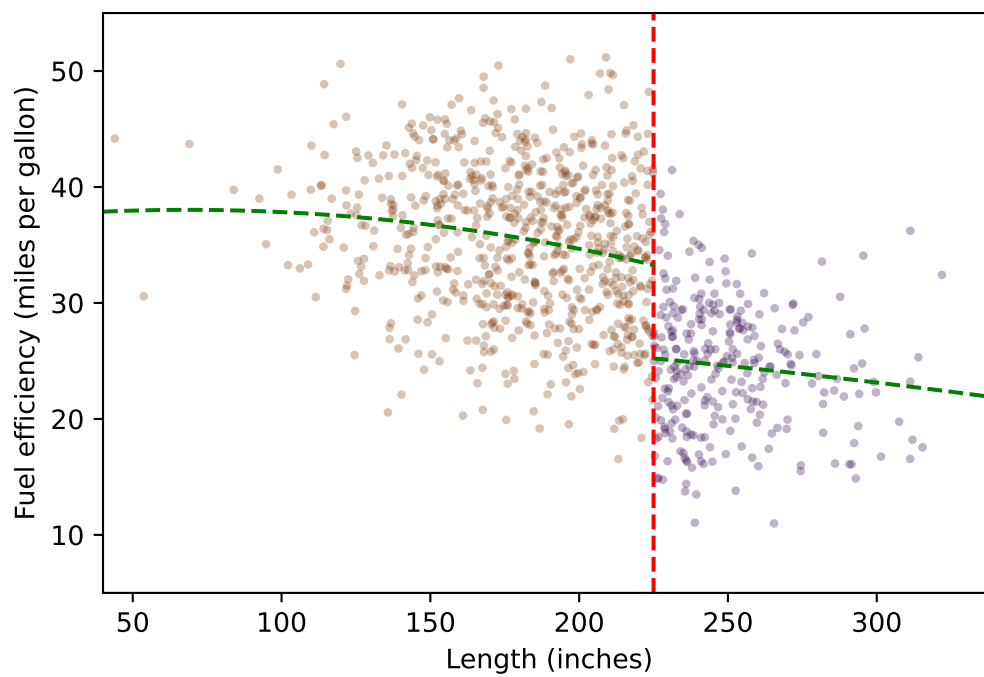


Figure 3: The second order polynomial.

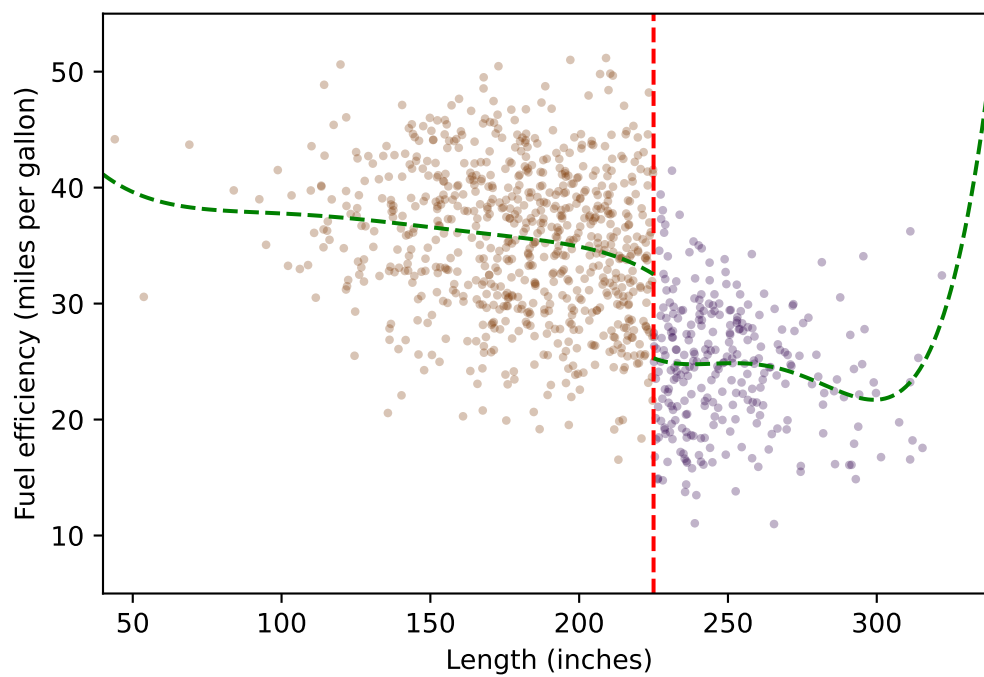


Figure 4: The fifth order polynomial.

6. It is more reasonable to use first-level polynomial in the first stage because as I mentioned in the previous section, treatment effect is statistically significant only in that case.

<i>Dependent variable: price</i>	
(1)	
Intercept	17392.94*** (776.88)
Car	-4747.32*** (323.69)
ATE	158.78*** (26.95)
Observations	1000
$R^2$	0.22
Adjusted $R^2$	0.22
Residual Std. Error	3427.23
F Statistic	135.63***
<i>Note:</i> *p<0.1; **p<0.05; ***p<0.01	

Table 2: 2SLS regression output

## 2 Stata

1. The average treatment effect from the second-stage regression results is presented in Tabel 3 and the scatter plot with fitted lines in Figure 5. ATE coefficients are very similar in section 6 and in the Stata output and both are statistically significant.

(1)	
VARIABLES	Average treatment effect from the second-stage regression results
ATE	157.5*** (26.73)
car	-3,671*** (224.1)
Constant	17,403*** (775.0)
Observations	1,000
R-squared	0.221
Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1	

Table 3: RD regression output

2. Yes it is a good IV.

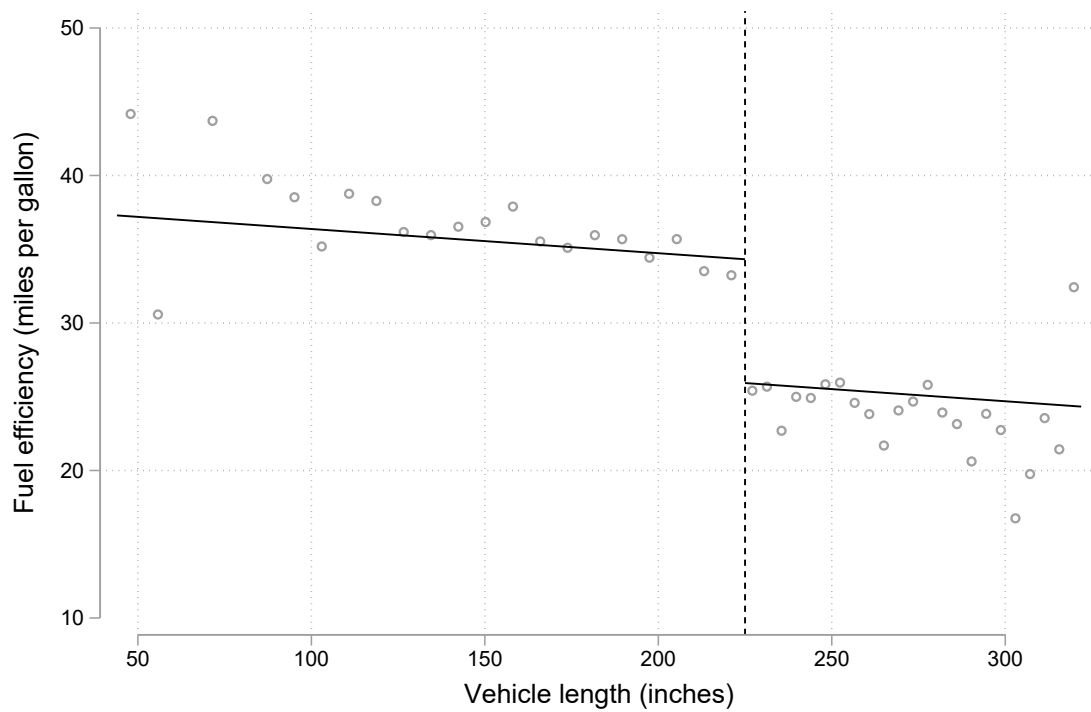


Figure 5: RD with the first order polynomial using 'rdrobust' command.