ECON3360 Causal Inference for Microeconometrics

Tutorial 9: Regression discontinuity design (fuzzy)

Instructor: Julie Moschion

Problem I: the impact of class size on students' outcomes

Background The dataset for this exercise comes from the paper by Angrist and Lavy (1999) "Using Maimonides' rule to estimate the effect of class size on scholastic achievement", published in Quarterly Journal of Economics, Vol 114, No 2. The data contains information on class size, average math and reading test scores for 2,024 5th grade classes in 1,003 public schools in Israel, as well as enrolment data for these schools. Use the data in maimonides.dta for the following questions.

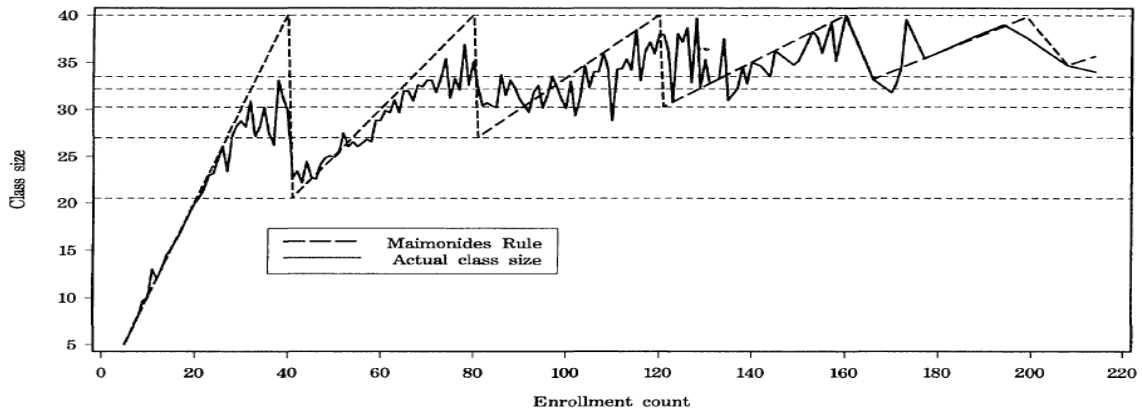
(1) Load and describe the data.

(2) Generate a table of summary statistics of class size, average math and reading test scores and the percentage of disadvantaged students (overall mean, standard deviation and mean for the following quantiles: 0.1,0.25,0.50,0.75,0.90.)

(3) Regenerate the same summary statistics by using a loop instead of summarising each variable separately.

(4) Plot the average class size as a function of school enrolment. Can you replicate the figure below?

(5) Estimate the linear regression for the relationship between a) class size and math scores and b) class size and verbal scores. Include controls for following variables progressively: the percentage of disadvantaged students in the class and the school enrolment. Use robust standard errors. Can you find similar numbers to the ones in the table below? (Hint: it is ok if your estimates are close but not exactly the same as this dataset is not the very last one they used in the final version of their paper.) How do you interpret the results?



	5th Grade					
	Reading comprehension			Math		
	(1)	(2)	(3)	(4)	(5)	(6)
Mean score		74.3			67.3	
(s.d.)		(8.1)			(9.9)	
Regressors						
Class size	.221	−.031	−.025	.322	.076	.019
	(.031)	(.026)	(.031)	(.039)	(.036)	(.044)
Percent disadvantaged		−.350	−.351		−.340	−.332
		(.012)	(.013)		(.018)	(.018)
Enrollment			−.002			.017
			(.006)			(.009)
Root MSE	7.54	6.10	6.10	9.36	8.32	8.30
R^2	.036	.369	.369	.048	.249	.252
N		2,019			2,018	

Now let's estimate the effect of class size on math scores using a RDD.

(6) Following Angrist and Lavy, generate a variable f_{sc} that predicts class sizes following Maimonides' rule as:

$$f_{sc} = \frac{enrollment}{int(\frac{1}{40}(enrollment - 1)) + 1}$$

What is the meaning of this variable?

(7) Estimate the effect of f_{sc} on avgverb and avgmath using OLS. Include controls for the percentage of disadvantaged students in the class and for school enrolment and enrolment^2/100. Use robust standard errors. What is this regression? Interpret the results.

(8) Estimate the effect of class size on avgverb and avgmath using f_{sc} as an IV. Use robust standard errors. Interpret the results.

(9) Compute the first stage regression. Interpret the result. How does the first stage relate to the estimates of questions (7) and (8).

(10) Summarise the findings of this exercise. As a policy marker would you advocate in favour of reducing the class size?