The purpose of this exercise is to get you acquainted with regression discontinuity strategies. In this exercise you will be asked to estimate the impact of class size on student performance using the data from the paper by Angrist and Lavy (1999). This paper exploits the fact that the maximum class size in Israel is set at 40 students (known as Maimonides’ rule) to identify the causal effect of class size on test scores. Read Angrist and Lavy (1999) for additional details.

1. Read the 4th and 5th grade data in the class website “Data for Problem Set No. 2” named “final4.dta” and “final5.dta.” Estimate descriptive statistics (means, sd’s, 10th, 25th, 50th, 75th, and 90th quantiles) for class size (classize), enrollment (c_size), percent disadvantaged (tipuach), average verbal test scores (avgverb), and average match test scores (avgmath) for 4th and 5th graders for the entire sample and for the discontinuity sample (Hint: generate an indicator variable for the observations in the discontinuity sample). Eliminate observations with verbal and math test scores less than 1 and greater than 100.

2. Estimate OLS regressions of verbal and math test scores on average class size only, on average class size and percent disadvantaged, and on average class size, percent disadvantaged and enrollment for 4th and 5th graders separately. Explain the potential biases in these regressions.

3. Estimate “Maimonides’ Rule”:

   \[ f_{sc} = \left( \frac{e_s}{\text{int}\left(\left(\frac{(e_s - 1)}{40}\right) + 1\right)} \right), \]

   and generate a graph of Maimonides’ rule and the average class size by enrollment size for 4th and 5th grades. Explain and interpret these figures.

4. Estimate 2SLS regressions of average verbal and match test scores for 4th and 5th graders on class size using Maimonides’ rule as the instrument and controlling for percent disadvantaged only first and then controlling for percent disadvantaged and enrollment. Do this for the full sample and for the discontinuity sample only. Explain the difference between these results and those in part 2.