HW7_Giorgio_Crisi

Giorgio Crisi Grotti

17/11/2020

Question 1 Consider the model Yi = Diβ + ei The interest lies on β, which gauges the causal effect of the treatment on Yi . Di is defined as the treatment dummy, which is equal to 1 when the individual i is treated and zero, otherwise. Suppose that Di = {1 if xi x {0 if xi < x (a) Rewrite the model using potential outcomes framework and list all assumptions that are necessary for consistency, giving the intuition of each one of them.

The assumptions necessary for the consistency of the RDD estimator are: at the limiar (x) that decides the treatment, the potential results are independent of treatment

(Y1i, Y0i) Di|xi = x

E [Y1i| xi = x0] and E [Y0i| xi = x0] are continuous at xi = x

The intuitions behind the hypotheses are, respectively:

If the treatment is locally randomized (individuals not control xi in the vicinity of limiar x), then the rule of treatment works like a random experiment for individuals who are in the vicinity of the limiar.

If conditional expectations of potential outcomes are limiars, the only discontinuous thing is the allocation of the treatment, and we can isolate the effect of the treatment.

(b) Show that on a neighborhood around the discontinuity, the estimator of RDD is equal to the causal effect of treatment on Y , i.e, equal β.

The RDD estimator is the sample analog of

And E[Y1i| xi = x] - E [Y0i| xi = x]

note that

E[Y1i| xi = x] = E [αxi + β + ei| xi = x]= αx + β + 0

E[Y0i | xi = x] = E [αxi + ei| xi = x]= αx + 0

Then, E[Y1i| xi = x] - E [Y0i| xi = x] = αx + β - αx = β

(c) Does this propriety is valid for all values of x? What is the implication of this?

No. For the RDD estimator to be consistent, it is the hypothesis that the treatment is locally random. This means that the treatment is random in the neighborhood limiar (x), but not for any value of xi. This implies that we are estimating the average local effect of treatment (LATE), and we know nothing about the effect of treatment for individuals with xi values are very far from the x limiar.

Question 2 Consider the RDD approach where Yis is the score in a national math exam of individual i in the school s. The treatment is one robotics class per week. This is, all treated schools have a robotics class and the non-treated schools don't have it. However, the assignment for the treat was not random: All schools with less than 100 students are treated. Define Di as the treatment dummy, i.e. is equal 1 for the schools that has a robotics class and zero, otherwise.

(a) If we are interested in the causal eect of the robotics class on the student's math skills, what's the problem in comparing the mean of the treatment group and the non-treatment group?

Since the treatment was not random, there can be a lot of difference between the characteristics of treated and untreated schools, in so that the comparison of the average of the two groups is "contaminated" by these other characteristics. We can, for example, imagine that in larger schools teachers have a harder time keeping up with the performance of students and offer individual attention, which would lead these schools to perform worse not only because of the lack of robotics, but also due to the overload of teachers.

(b) We can give a better estimator for the causal effect of the robotic class on the student's math skills? What's assumptions are necessary? Is probable that all of these assumptions hold? Why?

Yes, we can use an RDD. It is necessary that the treatment is locally random and that the conditional expectations of potential outcomes are the number of students.

The first hypothesis probably holds because there is no reason to imagine that, for example, schools with 102 students are extremely different from schools with 96 students. As a slight variation in the number of students enrolling in a given year is beyond the control of the school, we can consider that treatment is locally random.

The second hypothesis probably also holds, for the same reason: there is no reason to believe that the results potential of a school changes a lot if it has 99 or 101 students, then the conditional expectations of the results potentials must be continuous.

Question 3 Still thinking about RDDs, explain the concept of "bandwidth." How does using a larger bandwidth impact your estimate of treatment effects?

The idea of the RDD is to explore variation close to the threshold that determines the random treatment. However, it remains to be defined what we consider "close", and this is the concept of window. In the example question 2, with treatment being decided by the number of students, we can think of a few different windows:

10 students: we compare schools with 89 to 99 students (treatment) with schools with 100 to 110 students (control).

5 students: we compare schools with 94 to 99 students (treatment) with schools with 100 to 105 students (control)

The larger the window, the more observations we can include, which tends to decrease the variance of the estimator. On the other hand, a bigger window keeps us away from the threshold so that we can be violating the hypothesis of local randomness of treatment, which induces bias. In general: the larger the window, the greater the bias and the lower the variance of the estimator.