Foundations of Econometrics - Part II
Final Exam Solutions

Exam_2022.pdf

December 12, 2021

Question 1: Random Experiments

(10 + 5 points)

An internet and mobile provider randomly contacts half of its phone customers (who are not yet using their internet services) with an offer to purchase a high-speed internet package. You would like to evaluate the effectiveness of such promotional activities on sales of internet packages.

(a) (10 points)

Explain briefly how you would calculate the average treatment effect. Suppose you had additional information on individual customers, e.g. age, gender, race, would you use such information?

Answer: Because the provider randomly made the offer, we know treatment was randomly assigned. We could check that the distribution of the treatment and control matches across that information (age, gender, race, etc).
If we are not controling because of proper randomization, we can use the simple ATE, which is the difference in means: whats the difference in sign up rate for those who were contacted and those who were not.

(b) (Bonus: 5 points)

Could it be an issue if some phone customers live in the same household and share an internet connection? If yes, how could you address this in the randomization process?

Answer: Yes, this could be an issue. If someone lives in a household where another member pays the internet service, they will have no reason to accept the offer, so it may depress the treatment effect. The problem is that internet services is a household amenities. So we should focus our analysis on household level aggregation instead of individual. We could simply only look at phone customers who are heads of households (primary contact on the account).

Question 2: Matching

(20 points)

Answer the following questions about matching.

(a) (13 points)

State the formula of the propensity score estimator for the ATE. State one reason why it is suitable to use propensity score some cases rather than direct matching.

Answer: The formula for matching is:

$$\widehat{ATE}_{PS} = \frac{1}{N}\sum_{i=1}^{N}\frac{D_iY_i}{p(X_i)} - \frac{1}{N}\sum_{i=1}^{N}\frac{(1-D_i)Y_i}{1-p(X_i)}$$

We would use propensity score matching rather than direct matching when dealing with high dimensional data. That is, its easier to match based on a propensity score than attempt to match across many, often highly specific, other variables.

(b) (7 points)

Explain briefly how radius matching works.

Answer: Radius matching selects all options within a certain propensity score of the observation. This thus helps include a larger sample for control, instead of just selecting the best propensity score match. In addition, if we just seelct the best propensity score match, and one observation is the best match for multiple treatment observations, then it'll be repeated. Radius matching may utilize may of the treatment data in this case.

Question 3: Instrumental Variables

(25 points)

You would like to study the effect of staying in school on crime rates of highschool-aged children.

(a) (5 points)

Describe in a few sentences why an instrument might be needed in this case.

Answer: As is often the case in education, the choice to stay in school is not entirely random and there may be some unobserved factors that influence both stay in school rate and crime rates (socio-economic status, geographic location, etc). Those from worse backgrounds may be more likely to drop out and commit crime, which means stay in school rate alone does not describe crime rates of the children. Using an instrumental variable can ideally isolate these effects, as the assumptions tell us that a proper instrument would only impact the crime rates of children thru its impact on stay in school rate.

(b) (10 points)

You want to make use of an instrument, a law change that decreased the length of highschool by one year. Suppose you have data on two adjacent cohorts of children. For simplicity assume that all children begin studying in highschool at age 14 and that they cannot be retained (asked to repeat a grade) nor expelled. Hence, one cohort can stay in highschool until age 19 (cohort 0) and the other until age 18 (cohort 1). Some students might voluntarily drop out early. You further observe whether children of each cohort committed any crimes during ages 18-20.

Write down the formula for the IV estimator. What are Y, D and Z in this case?

Answer: In this case, Y is the crime rate (1/0 commited crime per student perhaps), D is the years of schooling and Z is the instrument to identify the cohort (stay until 19 or 18).
The formula is

$$\hat{\beta}_{IV} = \frac{E[Y|Z=1] - E[Y|Z=0]}{E[D|Z=1] - E[D|Z=0]}$$

(c) (10 points)

You are given the following information:

The cumulative distribution of students' school leaving age for each cohort is given by:

Leave School by Age/Cohort	**0**	**1**
18 or earlier	0.2	1.0
19	1.0	1.0

All children have left school by age 19 and it is apparent that cohort 1 is affected by the new law, with no children staying in school beyond age 18.

The crime rates for the two cohorts are given by:

Cohort	**0**	**1**
Crime Rate	0.02	0.05

Calculate the IV estimator. How general is the effect that you are estimating? Are there always-takers or never-takers?

Answer: E[Y—Z = 1] expected crime rate cohort 1 E[Y—Z = 0] expected crime rate cohort 2 E[D—Z = 1] expected years of schooling cohort 1 E[D—Z = 0] expected years of schooling cohort 0

So the first: 0.05 - 0.02 = 0.03 The second: 0.2(18) - 0.8(19) = 18.8 1.0(18) = 18.0 so: 18.8-18 - 0.8 so we get 0.03/-0.03 = -0.0375

This effect is only for the compilers. If monotonicity holds we dont have defier. There would be no always takers (you get forced out of school) but there may be never takers (you can still drop out).

Question 4: Regression Discontinuity

(15 points)

State the two conditions for Y and D around cut-off z_0 that are necessary for RD. Graphically illustrate the two conditions for a sharp design.

Answer: For regression discontinuity, we must have both continuity of the potential outcome around the cut off and a discontinuity in the treatment assignment.
Continuity of potential outcome:

$$\lim_{z \to z_0^-} E[Y_i|Z_i = z] = \lim_{z \to z_0^+} E[Y_i|Z_i = z]$$

This states that the only jump is from the treatment effect. Those that are just above and below the cutoff should be similar.
Discontinuous in treatment assignment: Sharp of fuzzy. Sharp means it jumps directly from 0 to 1, fuzzy is more gradual.

$$\lim_{z \to z_0^-} E[D_i|Z_i = z] \neq \lim_{z \to z_0^+} E[D_i|Z_i = z]$$

The probability of treating must jump at the cutoff. Sharp would jump directly from 0 to 1; fuzzy would mean that it still increases at the cutoff, but it may not be all the way from 0 to 1.

Question 5: Difference-in-Difference

(10 points)

You would like to assess the effect of Amazon's policy to offer same-day delivery on sales. There is data for locations A and B over time periods $T = \{0, 1\}$. Same-day delivery was only offered in location A in period 1. Assume you observe the following sales numbers in thousands of EUR:

Period/Location	**A**	**B**
0	100	150
1	200	200

Sketch the sales on the $Y - T$ plane for the two locations. Indicate the size of the treatment effect in your graph and also calculate it.

Answer: The treatment effect is the gap between where location A ended up and the counterfactual of where it would have ended up, given parallel trends.

$$\text{DiD} = (Y_{A,1} - Y_{A,0}) - (Y_{B,1} - Y_{B,0}) = (200 - 100) - (200 - 150) = 100 - 50 = 50$$