

Practice Midterm Exam

Answer all questions. If you spot typos or other errors, or need clarification, send me an e-mail. Cheers – Dr.D.

- 1) Explain briefly how and why simple models, optimization, equilibrium (particularly supply and demand), and an understanding of property rights together form an important part of the core of microeconomic analysis.
- 2) Suppose riding coaster A provides exactly the same total satisfaction as riding coaster B, and that each seats 50 riders per cycle, but that coaster A takes only $2/3$ as much time as coaster B to complete a cycle. Use the idea of equilibrium to estimate the length of the line for B relative to the line for A. Explain your approach.

Questions 3-7. One-hundred people fish in two lakes, East Lake and West Lake, on any given day. Each may choose to fish in either lake. Let n be the number that fish in East Lake and $100-n$ be the number that fish in West Lake. West Lake is very large and relatively plain with dispersed fish populations. The value someone receives fishing in W is 20. East Lake is spectacular with concentrated fish populations. The value someone gets from fishing in East Lake is $50-0.5n$.

- 3) If people individually choose which lake to fish in, how many will fish in East Lake and how many in West Lake, and what is the total value to everyone fishing? (Total value is just the number fishing East multiplied by the individual value of fishing in East, plus the number fishing in West multiplied by the individual value of fishing in West.)
- 4) If a benevolent dictator could limit the number fishing in East, how many people fishing in East and how many in West would maximize total value, and what would total value be?
- 5) Relate the difference between outcomes in #3 and #4 to property rights and the tragedy of the commons.
- 6) Suppose the government can charge a fee to fish in East Lake. How large would the fee need to be to induce the optimum number of people to fish there? How much revenue would the government collect?
- 7) Discuss the limitations of randomized control trials covered in the article by Deaton and Cartwright.
- 8) Discuss what gave rise to the credibility revolution.
- 9) List and discuss several factors that combined to increase the credibility of empirical economics over the course of the credibility revolution.
- 10) Discuss a potential downside of the credibility revolution.

Practice Midterm Exam

All remaining questions. You are interested in whether a tutoring program improves performance. You have a measure of initial ability, S_1 . You know who signed up for and received voluntary tutoring, T . You later observe a second measure of performance, S_2 . You are worried unobservables, U , influence the initial score, the probability of signing up for tutoring for a given initial score, and the second score.

11) Give some examples of unobservables relevant to this case.

12) Draw the DAG.

13) What is a backdoor path? Why do you have to block them? How do you block them? Based on the DAG, what backdoor path or paths, if any, can't be blocked with the data you have?

Questions 14-17. Suppose there are only four students and the data is as shown in the table to the right.

Subject	Y^1	Y^0	T
1	84	76	1
2	76	82	0
3	74	64	1
4	80	78	0

14) What are the observed sample difference, ATT, ATU, ATE, and selection bias?

15) Why, in reality, can't you actually calculate ATT, ATU, ATE, or selection bias? That is, what is the fundamental problem with causal inference?

16) Can big data or AI, on their own, possibly resolve this fundamental problem of causal inference? Why or why not?

17) How does random assignment help break selection bias?

Questions 18-21. Assume treatment was randomly assigned.

18) What exactly is Fisher's sharp null in this context? Why is Fisher's sharp null particularly useful for randomization inference?

19) Show there are 6 possible ways to assign treatment to 2 of 4 students.

20) With the small sample you can construct all possible assignments manually. What are the six sample differences under Fisher's sharp null?

21) Discuss how strong the evidence is for an actual effect based on this (small) sample. Basically, what is the relevant p-value and what do you make of it?

Question 22-23. Suppose you have a sample of 400 with 200 randomly treated.

22) Show the expression for the number of ways to randomly choose 200 from 400. Don't bother trying to calculate how many that is. Why not?

23) Practically, how can we approximate the Fisher exact test p-value with large samples?

Practice Midterm Exam

24) Now suppose treatment is not randomly assigned. Instead, everyone who scored under a 60 got extra tutoring and no one else did. Explain how a regression discontinuity design can get at the causal effect of tutoring. Draw a picture to illustrate the LATE.

25) What is a LATE? Why can we only estimate the LATE, not the ATE? Why do you often need a lot of data to make an RDD work?

Questions 26-27. Now suppose everyone with an initial score under 60 was assigned to tutoring, but some did not take it, and that a couple of students who scored a bit over 60 were able to get tutoring anyway.

26) How would you go about employing regression discontinuity to get at the causal effect in this case? What do you use as an instrument for treatment? Why do you need an instrument? What is an instrumental variable, anyway?

27) Draw a figure clearly illustrating the difference between sharp and fuzzy RDD. Hint: This is not about the outcome measure.

Questions 28-32. Data on 400 (hypothetical) observations for these questions is in the accompanying file “tutoring.csv”.

28) Calculate the simple sample difference.

29) Bin the data in 5-point increments, letting each bin be represented by its midpoint. For each bin, calculate the percent that were treated. Plot this against the bin midpoint. Plot a vertical line at the cut point. Interpret the figure.

30) For each bin from the previous question, calculate the average score on the second exam. Plot this against the bin midpoint. Plot a vertical line at the cut point. Interpret the figure.

31) Calculate the local effect of intention to treat. You will need to do a regression to accomplish this.

32) Calculate an estimate of the LATE. You will need to use an instrumental variables estimator.

NOTE: If you do not know R or Python, or like, know, or have access to Stata, you can use excel or Minitab for the regressions, both of which are available to students at Poly and have an easy learning curve (though they are not powerful for doing creative analytical things).