ECON 121 FINAL EXAM - FALL 2024

You will analyze two datasets, one from the US and one from Israel, to estimate the effect of class size on student achievement in elementary schools.

The dataset https://github.com/tomvogl/econ121/raw/main/data/project_star.rds contains data on kindergarten students in a sample of schools in Tennessee. An experiment called Project STAR randomly assigned students within each school to one of three types of classes: small class (13-17 students per teacher), regular class (22-25 students per teacher), or regular-with-aide class (22-25 students with a teacher's aide). Teachers were also randomly assigned to classes. Each observation is a student. Variables are:

`schoolid`	Unique school identifier
`srural`	= 1 if school is in rural area, 0 otherwise
`classid`	Unique class identifier
`cltype`	Class type (1 = small, 2 = regular, 3 = regular + aide)
`csize`	Number of students in class
`tmasters`	= 1 if teacher has a master's degree, 0 otherwise
`texp`	Teacher's years of experience
`twhite`	= 1 if teacher is white, 0 otherwise
`smale`	= 1 if student is male, 0 otherwise
`swhite`	= 1 if student is white, 0 otherwise
`sfree`	= 1 if student is eligible for free lunch (a proxy for being from a poor family), 0 otherwise
`stest`	Score on reading portion of the Stanford Achievement Test

The dataset https://github.com/tomvogl/econ121/raw/main/data/maimonides.rds contains data on 4^{th} grade classes in Israel. Each observation is a class, with potentially many classes per school. Variables are:

`school_id`	Unique school identifier
`grade_size`	Number of 4^{th}-grade students in the school
`class_size`	Number of students in the class
`disadv`	Share of students in the class from disadvantaged families
`avg_verb`	Average verbal exam score in the class

You may refer to books, internet resources (including AI), notes, and code from problem sets, solutions, and classroom examples. You may **not** communicate with any person but Prof. Vogl or Regina during the exam; evidence of such communication will result in failure. If you are taking the exam remotely, record yourself with Zoom and submit the recording by email; if you have questions during the exam, you may e-mail your question to Prof. Vogl. Write code and answers in the R Markdown template. When you finish,

knit to PDF (or knit to HTML and then print to PDF in a web browser), and submit the PDF on Gradescope. Refer to the exam information document for instructions.

Questions that require coding have a **C** next to them. (Full credit for **C** questions requires only correct code. We recommend explaining your code in comments, so you can receive partial credit even if your code is wrong.) Questions that require words have a **W** next to them. (**W** questions may require you to do some calculations "by hand," but no coding is necessary.) Every question is worth the same number of points.

Part A: Project STAR

1. Open the Project STAR dataset. The exam score has an weird scale, so standardize it by subtracting the mean and dividing by the standard deviation. Regress the standardized test score on the student's gender, race, and free lunch eligibility. Here and in subsequent questions, be careful to estimate the standard errors appropriately. **C**
2. Interpret the estimates in the above regression. Are they statistically significant? (2-3 sentences) **W**
3. Policymakers are interested in which types of students are below grade level. Classify a student as below grade level if the original test score is below 400. Run a logit regression of the student's grade level status on their gender, race, and free lunch eligibility. Compute odds ratios and marginal effects. **C**
4. Of the three variables included in the above model, what is the largest risk factor for being below grade level? Interpret the odds ratio and marginal effect for this variable. (2-3 sentences) **W**
5. Estimate the effect of class type on the standardized test score. Here and in subsequent questions, pay attention to the design of the experiment in how you specify the regression. **C**
6. Interpret the estimated effects in Question 5. How large are they relative to the test score gap between free lunch and non-free lunch students in Question 1? Do the estimates represent the causal effects of class type on students' test scores? (3-4 sentences) **W**
7. Test whether the effects of class type on the standardized test score differ between teachers with no experience and teachers with some experience. Estimate the ratio of the effect of "small class" when teachers have no experience to the effect of "small class" when teachers have some experience. Use the delta method to estimate the standard error. **C**
8. Interpret the results from Question 7. Can you reject that "small class" has the same effect when teachers have no experience and when teachers have some experience? Based on the ratio you estimated, how much larger or smaller is the effect for teachers with no experience? (2-4 sentences) **W**
9. Many education policymakers are not interested in the effect of having 13-17 versus 22-25 students per teacher. Instead, many want to know would be the effect of increasing class size by one student. Estimate the effect of increasing class size by one student on the standardized test score. **C**

10. Interpret the result from Question 9. What do you think is the most important assumption we need to convert the "small class" effect in Question 5 to a per student effect in Question 9? (2-3 sentences) **W**

Part B: Maimonides' Rule

11. Open the Israeli dataset. Regress the average test score on the share of students who are disadvantaged. Here and in subsequent questions, be careful to estimate the standard errors appropriately. **C**
12. Interpret the estimated slope in Question 11. Suppose a classmate told you that if you weighted the regression by the number of children in the class, the coefficient stay approximately the same, while the standard error would shrink. Under what conditions would this be true? (2-3 sentences) **W**
13. Israeli schools follow a rule set by the 12^{th}-century rabbi Maimonides: the number of students per class cannot exceed 40. This rule induces a discontuous decline in class size when the number of 4^{th}-graders rises above 40. At 41, the school must have two 4^{th} grade classes: one with 20 students, the other with 21. We will use this cutoff in a regression discontinuity (RD) design. Define the running variable for the RD design as `x = grade_size - 41`, so that the rule binds when `x` $\geq$ `0`. Next, draw binned scatterplots to visualize the RD design, restricting to observations that are within 25 students of the cutoff. To do so, first discretize `x` into bins that are 5 students wide, then estimate bin-specific means of the outcome variables, and then plot the binned data. Draw two binned scatterplots, one with `class_size` as the ourcome and one with `avg_verb` as the outcome. **C**
14. Interpret the two graphs. Is there evidence Maimonides' rule influences class size? In the language of instrumental variables, what is this relationship called? Is there evidence that the rule influences test scores? In the language of instrumental variables, what is this relationship called? (4-6 sentences) **W**
15. Estimate the discontinuities from Question 14 using local linear regression with a bandwidth of 25. **C**
16. Interpret the results from Question 15. What do we need to assume to interpret them as causal effects of Maimonides' rule? Describe one way to assess the validity of the assumption(s). (3-5 sentences) **W**
17. Implement the strategy you described to assess the validity of the assumption(s). **C**
18. Use your results from Question 15 to calculate ("by hand," without additional code) the effect of class size on the average verbal test score. If causal effects are heterogeneous, whose average causal effect does your result represent? Do we need any additional assumption(s) for this interpretation, beyond those you stated in Question 16? (4-6 sentences) **W**
19. Use two-stage least squares to obtain a standard error for the effect you calculated in Question 18. For comparison, also run an ordinary least squares (OLS) regression of the average verbal score on class size in the full sample. **C**
20. Is the TSLS result significantly different from zero? How does it differ from the OLS result? Provide one possible explanation for the difference. Be concrete. (4-6 sentences) **W**