Questions to answer for Week 4

**Questions on the Angrist and Lavy “Maimonides’ Rule” paper**

*Note: I will not ask any questions concerning the Instrumental Variables estimation strategy because we have not learned this method yet. It will be the focus of our next unit. Please save your questions on this part of the Angrist-Lavy paper for then.*

 A. Focus of the paper

1. What research questions do the authors ask?
2. Describe the conventional method for addressing this question
3. Provide an example to illustrate the problem with the conventional analysis method.
4. What conclusions do the authors reach about the answer to their research questions? (Do their findings provide an estimate of the impact of a single year’s impact of class size on student achievement or an estimate of the cumulative impact? Please explain your answer.)
5. To what groups do the authors’ findings pertain? (*In reading Section V.A., pay attention to this question and be less concerned with the notation*.)

 B. Data

1. What data sources do the authors use for their study?
2. What are the characteristics of the samples? (i.e., who is included and who is excluded?)
3. What is the range of class sizes in the data?
4. How do the authors measure student achievement?
5. What is the unit of analysis in the research? Why does this matter?

 C. Methodology

1. What is the authors’ identification strategy?
2. What assumptions must hold for this to be a legitimate way to generate an unbiased estimate of the causal impact of class size? (Hint: see page 549.)
3. What are the “sizes of the enrollment windows” that the authors use in their analysis?
4. What is the nature of the tradeoff in choosing the sizes of the windows?
5. Why do the authors decide to predict class sizes from beginning of the year measures of enrollment instead of from measures of enrollment, say, two months later when enrollments have stabilized?
6. Why is the index of socioeconomic status such an important control variable?
7. Explain the method the authors use to control for enrollment effects in the models that they fit using the narrow “windows” (p. 559).
8. How do the authors test the hypothesis that the effects of class size on student achievement are greater for economically disadvantaged students?
9. Explain the logic underlying fitting the equations that included a “piecewise linear trend” with slopes identical to the slopes of ***fsc***.

 D. Details of the findings

1. In what respects do the findings for the samples at different grade levels differ?
2. Describe the authors’ conjectures about the reasons the findings differ by grade level.
3. Do you find the authors’ conjectures compelling? Explain your thinking.
4. Do you agree with the authors that the grade 3 findings are not of interest? Explain your reasoning.

 E. Threats to validity

1. Describe the threats to the validity of the Angrist-Lavy identification strategy
2. How do the authors deal with these threats?
3. Do you think that Angrist and Levy needed to worry about any of the concerns that led Urquiola and Verhoogen to decide that the data from Chile were not suitable for estimating the impact of class size on student achievement (see *MM* Ch. 9)? On what evidence do you base your answer? Do Angrist and Lavy mention these concerns? If so, what were their responses?

**Questions on the Dee and Penner Ethnic Studies paper**

*Again, for now we will not focus on the paper’s discussion of its “fuzzy” regression discontinuity results obtained by instrumental variables estimation. We will return to this in the next unit.*

1. Focus of the Paper
	1. What research questions do the authors ask? In answering this question, try to characterize what the “treatment” being studied in this paper is.
	2. What are the central findings of their study?
2. Data
3. Describe which students, schools and years were included in the authors’ analytic sample and what choices they made to exclude certain schools/students.
4. What are the authors’ outcome measures and how are they operationalized? Why do they calculate students’ GPA exclusive of their grades in PE and social studies?
5. What proportion of their sample is “eligible” for treatment?
6. Methodology
	1. List three ways in which prior research on ethnic studies curriculums (such as those in Tucson) could have their internal validity compromised.
	2. How were students selected for participation in the SFUSD high-school Ethnic Studies course?
	3. Dee and Penner construct a “binary indicator for [their] ITT variable.” What is that variable and what does it mean to refer to it as an *intent-to-treat* variable?
		1. *A point of clarification: when the authors refer to their model on pg. 139 as a “reduced-form” equation, you can interpret this for now as identical to an intent-to-treat model. We will unpack what the reduced-form term means in our IV unit.*
7. Details of findings
	1. Figure 1 is a critical test of the authors research design. What does it show and why is it an important first step in establishing validity?
	2. Look at the number in the first cell of Table 4 (6.164). Explain what this number represents. In doing so, try to use the term *average treatment effect* and make sure to explain what is the counterfactual and to whom the population of inference is for this coefficient estimate.
	3. In both Figures 3-5 and Table 7, the authors present analyses conducted on the full sample as well as on sub-samples of different “bandwidths” around the discontinuity. How (if at all) do their results differ based on their sample choice and why do they present these alternative specifications?
	4. Oregon recently passed HB 2845, which integrated ethnic studies into the statewide social studies standards in 2021 (<https://www.oregon.gov/ode/educator-resources/standards/socialsciences/pages/ethnic-studies-hb2845.aspx>). Local school districts and teachers will need to adhere to these standards (including the use of culturally responsive and relevant practice) as of the 26/27 school year. How would you explain what the results of this study do (and don’t) tell us about the ethnic studies curricula to an uncertain Oregon school superintendent?
8. Threats to validity
	1. Two central (related) threats to the validity of all regression discontinuity designs are the problems of bunching (“heaping”) and manipulation around the threshold. What tests do Dee and Penner conduct to assess these concerns? What do they find?
	2. Dee and Penner discuss two other validity concerns specific to the implementation of the ethnic studies in SFUSD on the top of pg. 145. What are they and how do they test these?

**Questions on Chapter 9 of *Methods Matter***

1. Explain in words what each of the three parameters in the canonical regression discontinuity model presented in Equation 9.1 (pg. 179) represent.
2. The example from Ludwig and Miller’s (2007) Head Start study generalizes the functional form of the relationship between the forcing variable and the outcome. Extending your answer above, what do each of the parameters in Equation 9.5 (pg. 188) represent and how are these different from Eq. 9.1?
3. What is the benefit of centering the running variable at the cutoff? How does this make interpretation of $\hat{α}$ in Eq. 9.6 more straightforward?
4. Take a close look at Figures 9.2 and 9.3. What do the dashed and solid lines represent, respectively? What potential concerns do the differences between these two sets of lines raise for you?
5. The Ludwig and Miller study assembles various data sources to bring insights to their study. List at least four data sources from their paper, as summarized in Murnane and Willett (note that there are an additional two sources not mentioned in *MM*). How does this inform your thinking about your own research?
6. What conditions need to be satisfied for a regression discontinuity approach to provide an unbiased estimate of the answer to a causal question?
7. Modern researchers interested in defending the assumptions under which regression discontinuity strategies return unbiased causal estimates often refer to the “Big Three” graphs. They argue that all regression discontinuity studies should minimally include three graphs justifying their assumptions. What are these graphs? (*note: Murnane and Willett do not explicitly reference the “Big Three,” but in describing the conditions under which RD assumptions are met, they do so implicitly. Can you figure it out?*)
8. Was compliance with the class size maximum rule better in the data set from Israel that Angrist and Lavy analyzed or in the dataset from Chile that Urquiola and Verhoogen analyzed? Be prepared to explain the evidence supporting your answer.
9. Why did Urquiola and Verhoogen conclude that the application of the regression discontinuity (RD) methodology to the data from Chile would not provide an unbiased estimate of the impact of an offer of class size on student achievement?
10. What lessons do you take away from the Urquiola and Verhoogen paper about the steps you should take in examining data to determine whether the RD method is appropriate for providing an unbiased estimate of the impact of a particular policy?