Questions to answer for the January 27 class

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

*Note: We will not ask any questions concerning the Instrumental Variables estimation strategy. The reason is that 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?
   1. Does the offer of smaller class size improve student achievement in Israeli public schools?
   2. Do smaller classes improve student achievement in Israeli public schools?
2. Describe the conventional method for addressing this question?
   1. Conventional method is to regress performance on class size w/ variety of control predictors included
3. Provide an example to illustrate the problem with the conventional analysis method.
   1. Families w/ greater social capital may influence their school’s principal to place their child in a smaller class which would produce spurious correlations between class size and performance
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.)
   1. The authors conclude that being in a class of about 10 fewer students increases student achievement by 2.2 points (.2-.3 SDs) on average
   2. The estimate is for assignment to a small class in just that year. However, enrollment cohorts progress through school together and so small classes may follow cohorts.
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.)
   1. Weights on estimates primarily driven by those who attend smaller schools, with enrollment close to the discontinuity point (and have fewer poor students)
      1. This last point because use of PD funds makes fsc less determinative for those schools

B. Data

1. What data sources do the authors use for their study?
   1. They use class level data for 4th and 5th graders in 1991 from the ministry of education
   2. For 1992, they receive micro data for individual 3rd graders, compute average score for each class
   3. Link scores to school characteristics and class size w/ Central Bureau of Statistics Census of Schools
2. What are the characteristics of the samples? (i.e., who is included and who is excluded?)
   1. They have around 1,000 schools containing 2,000 classes (varies a little from year to year) with a total of 62,000 students
   2. Limited to Jewish system (religious and secular), but not ultra-orthodox
   3. Average school has 30 pupils/class, 78/grade
3. What is the range of class sizes in the data?
   1. Range of class sizes is 10th %ile: 21, 90th %ile: 38
4. How do the authors measure student achievement?
   1. Student achievement measures from national math and reading tests
      1. Composite of some of basic and all of advanced questions
      2. Scaled from 1-100 by dividing by # of questions
5. What is the unit of analysis in the research? Why does this matter?
   1. The unit of analysis is the class
   2. Permits them to use class level data from 1991, however, artificially compresses the standard errors b/c students’ scores have much more in-class variance than between-class variance, so taking class mean reduces overall variance.

C. Methodology

1. What is the authors’ identification strategy?
   1. Use the fsc-determined cutoff point to estimate the exogenous offer of a small class size to determine its effects on student achievement
   2. The other identification strategy to answer their 2nd research question is to predict actual class size from the function fsc that is the offer of a small class size, then use this to estimate the effect of smaller classes
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.)
   1. Key assumption in ITT identification is that parents do not exploit Maimonides rule to place their children into smaller classes
   2. (Key assumption to IV is that covariates partial out any other reason for differences in class size, and therefore the residual does not explain any of the variation in class size assignment)
3. What are the “sizes of the enrollment windows” that the authors use in their analysis?
   1. They use enrollments that are +/- 5 students of enrollment on either side of the cutoff
      1. Also specify a model that uses +/- 3 cutoff
4. What is the nature of the tradeoff in choosing the sizes of the windows?
   1. The smaller the window, the easier it is to specify the appropriate functional form because over small intervals, everything is linear, but the lower statistical power to detect an effect and the harder it is to generalize away from the cut point
   2. Bias v. variance tradeoff
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?
   1. They are concerned that middle of the year measures will include endogenous choices made by parents or principals to move children to smaller or larger classes
6. Why is the index of socioeconomic status such an important control variable?
   1. SES is hugely correlated w/ total enrollment b/c poorer families tend to live in the countryside where schools are smaller
   2. Therefore, until it is partialed out, class size is positively correlated w/ performance
7. Explain the method the authors use to control for enrollment effects in the models that they fit using the narrow “windows” (p. 559).
   1. They include dummies indicating whether total enrollment is around the 40, 80 or 120 window cutoff
   2. They then create other dummies that indicate whether the class falls to the left of the cutoff (i.e., gets the Treatment)
   3. That permits them to not need to control for enrollment effects b/c would assume that they’re equal in expectation
   4. Also, don’t have to specify a particular functional form b/c can vary at each cutoff and you just capture offer of small class or not
8. How do the authors test the hypothesis that the effects of class size on student achievement are greater for economically disadvantaged students?
   1. Interact fsc with PD
9. Explain the logic underlying fitting the equations that included a piecewise linear trend with slopes identical to the slopes of ***fsc***.
   1. If you can fit small linear trends, shouldn’t need to control for covariates b/c you’re estimating over a very small range of enrollment and over small ranges, all functions are linear
   2. You’re forcing linear trend identical to the slope fsc, you’re ensuring that the effect of enrollment on class size is exactly what function says it should be, so changes in class size are purely exogenous

D. Details of the findings

1. In what respects do the findings for the samples at different grade levels differ?
   1. Findings at grade 3 are non-significant, grade 4 are significant for reading, but not for reading, grade 5 are significant for both scores
2. Describe the authors’ conjectures about the reasons the findings differ by grade level.
   1. They argue that the effects could be cumulative and are higher for students who are in a cohort that consistently gets a small class
   2. They also argue that in 1992, there was a lot of teaching to the test which drastically reduced variability in scores, such that in 25% of classes the average scores was over 90%, this would make it hard to find any effect
3. Do you find the authors’ conjectures compelling? Explain your decision.
4. Do you agree with the authors that the grade 3 findings are not of interest? Explain your reasoning.
   1. I think they are of interest in that they suggest that findings such as Hanushek’s about class size’s importance in 1st year may not be accurate, but generally given the fact that variability shrunk so much, I would not place too much weight on them
   2. I would have been interested in seeing them analyze the 1992 results at the micro level since that would presumably capture more variance

E. Threats to validity

1. Describe the threats to the validity of the Angrist-Lavy identification strategy?
   1. Parents may choose into smaller classes, or principals may determine enrollment
2. How do the authors deal with these threats?
   1. They use IV estimation to determine what portion of variation in class size is a result of assignment to “small class” treatment
   2. They can claim that whatever portion of the variance due to assignment to one condition or another is the exogenous variation and therefore doesn’t have the selection problem
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?
   1. This does seem to be an issue b/c CDF in Figure V shows a lot more people piled up around 40
      1. See also class figure
   2. Manipulation of class size is limited b/c few attend private schools
   3. Use beginning of year measure that is less likely to have been influenced by parental/admin behavior

**Questions on the Ludwig and Miller paper**

1. Focus of the paper
2. What questions do the authors ask? What answers do they provide?
   1. Did the provision of grant-writing assistance to the poorest counties cause a proportional increase in their receipt of Head Start services?
   2. Does the provision of grant-writing assistance for Head Start services improve child mortality rates on those indicators that Head Start screening and counseling might affect?
   3. Does the provision of grant-writing assistance for Head Start services improve educational outcomes for students residing in communities receiving greater Head Start funding?
3. Describe the causal mechanisms through which the authors envision Head   
   Start affecting health and education outcomes for participating children.
   1. Education, parental involvement, nutrition, social services, mental health services, and health services could all affect instances of anemia, meningitis, respiratory problems
      1. Health services directly and others through lower levels of chronic stress and influence on parental practices
      2. Low levels of inoculation could be primary mechanisms
   2. Education could have direct effect; others could have indirect
4. Data
5. Describe the data that the authors bring to bear to implement their identification strategy. (Make a list of the data sets used, and describe briefly the use of each.)
   1. National Archives and Records Administration: used to check county-level poverty
   2. 1960 Census: used to check which were 300 poorest counties
   3. Vital Statistics: used to determine cause of death and age
   4. County-level data: used to get level of schooling
   5. Restricted-use, geo-coded NELS: get individual level data for education outcomes, Head Start enrollment, total schooling
   6. Panel Study of Income Dynamics (PSID) to determine levels of mobility
6. What years of child mortality data do the authors use? What is the reason for this choice?
   1. 1973-1983 b/c all children aged 5-9 would have been aged 3-4 during a time where their counties would have been eligible for HS services
      1. They believed that ages 5-9 would capture effects of being in HS during 3-4
      2. Stop in 1983 b/c after then, all counties start to receive HS at equal levels irrespective of cut point
7. Methodology
8. What is the authors’ identification strategy?
   1. Authors exploit the discontinuity in grant-writing assistance at the 300th poorest county which had poverty rate of 59.2 percent in the 1960 decennial census to identify causal effect of grant-writing assistance to secure HS services in a regression discontinuity design
9. How do the authors make a judgment about the number of years that the discontinuity in the availability of Head Start remained in effect? What conclusion did they reach?
   1. By late 1990s, disparity between HS services across counties that had and hadn’t received grant had completely dissipated (based on ECLS-K data). Decided to only include up through those who had received services through late 1970s
   2. Starting w/ Community Partnership Act in 1974, effort to eliminate differences in HS funding
10. What functional form do the authors use in modeling *m(Pc)*?
    1. They decide to use both a quadratic and interaction parametric form as well as a non-parametric with local linear regressions
       1. Ultimately non-parametric fits data best, though they often present both
    2. Regression Discontinuity works best when you have a lot of data on either side of cut point
11. What is the “treatment” that the authors investigate? Is it a dichotomous variable?
    1. Treatment is receipt of grant-writing assistance
    2. It is a dichotomous variable
12. Details of the findings
13. Describe the findings in a way that a policymaker like Oregon Governor Kate Brown could understand. Your description should include any caveats about interpretation that you feel Governor Brown should know.
    1. Provisions of grant-writing technical assistance to poor communities improved the likelihood they would receive higher levels of HS funding/child
    2. In the 1970s and 80s this higher level of grant funding lowered child mortality rates by about 1-2 children/100,000 and may have slightly increased educational attainment, though problems with the data make us less confident about this second finding
    3. The study was completed in a very different context where alternatives to HS in the South in the 1970s and 80s tended to be worse than they are now, so conclusions about the value of moderate-quality early –childhood education should be made with caution
    4. Benefits we hope to see from early-childhood education and care tend to be in health outcomes. Any educational benefits (absent much better program quality) are likely to be moderate and long-term
14. Which of the findings do you find most interesting?
15. Threats to Validity
16. Explain the threat to validity caused by “selective migration”, including an explanation of why it would occur and the problem it would cause.
    1. Families might move to counties where HS services were being provided
    2. Would mean that those families were different in expectation from others, presumably with higher expected outcomes, and bias results upward
    3. Note: use NELS and Panel Study of Income Dynamics to test whether there was much migration, and looks like not
17. What are other threats to validity? How do the authors deal with them? What datasets do the authors use in responding to these threats?
    1. PMIs could have made HS better in counties they visited (External)
       1. Seems unlikely b/c there for short time and large training provided to everyone afterwards
       2. Using voter registration data, find no increased levels of community mobilization in places that receive tech assistance
    2. Other sources of government funding (Internal)
       1. Regress outcomes that couldn’t have been affected by OEO technical assistance on poverty and look at whether there’s a discontinuity at cutoff point, and don’t find anything
    3. Endogenous break-point (Internal)
       1. Look at other point where there might be breaks
       2. Look at whether socio-demographic variables are balanced at cut point

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

1. What conditions need to be satisfied for a regression discontinuity approach to provide an unbiased estimate of the answer to a causal question?
   1. Forcing variable predicts (some) change in probability of receipt of treatment
   2. No bunch at discontinuity
   3. No manipulation across discontinuity
   4. Treatment predicts change in outcome in some discontinuous fashion
2. 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?*)
   1. Density of observations in bandwidth around discontinuity
   2. Binned scatterplot of covariates in bandwidth around discontinuity
   3. Binned scatterplot of outcome against forcing variable
3. 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.
   1. Compliance was much better in the Urquila and Verhoogen paper.
      1. Figure 9.4 shows a nearly direct match between observed class size and the predicted size from the govt mandated function
      2. Figure 1 in the Angrist and Levy paper (as well as Table 9.1 in MM) show that there was much lower compliance in Israel than Chile
   2. While compliance with the rule was higher, this was still not a sufficient condition to justify the assumptions of the RD appraoch
4. 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?
   1. Schools made substantial efforts to not enroll children that would push them over the limit to require the hiring of another teacher, except for a few that charge much higher tuition, and therefore attract wealthier families
5. 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?
   1. Learn a lot about the context of data
   2. Examine it closely to see whether evidence exists that actions by participants led to violations in the equal in expectation assumption
      1. Should compare descriptive statistics for variable other than the one that you think will have been discontinuous
6. Explain how MDRC made use of the RD methodology to examine the impacts of the *Reading First* program that was the centerpiece of President George W. Bush’s No Child Left Behind legislation. How did a school district’s policy regarding the allocation of *Reading First* funds play a role in determining which districts would participate in the evaluation of the *Reading First* program?
   1. Districts created needs-based index that determined whether school would receive Reading First funds
   2. Those that assign a fixed cut score above which people receive RF services and below which they don’t were included in the MDRC/Abt evaluation