ECO 348K, Applied Econometrics and Data Analysis   **Prof. Tom Vogl**

Spring 2019   **University of Texas at Austin**

PROBLEM SET 1: PROJECT STAR

DUE AT 4 PM ON FRIDAY 2/8

Education policymakers disagree over the extent to which school resources can improve educational outcomes. The Tennessee Student/Teacher Achievement Ratio experiment, known as Project STAR, aimed to provide evidence on this issue by randomizing students and their teachers into three types of classrooms: small classes (13-17 students per teacher), regular classes (22-25 students), and regular classes with full-time teachers' aides (also 22-25 students). The experiment took place in the mid-1980s in Tenessee public schools. The dataset contains data on the test scores, background characteristics, and classroom characteristics of kindergarteners involved in the experiment.

The questions below guide you through an analysis of the effects of classroom resources on student test scores. A major goal of the problem set is to highlight how far means, t-tests, and cross-tabulations can take you. Unless otherwise noted, *do not run regressions*.

Please place a hard copy of your write-up in the submission box at the economics department office in BRB. Write your name, Prof. Vogl's name, TA Kim's name, and the course number clearly at the top of the first page. And please have one group member e-mail your do-file to eco348k@gmail.com. Write your do file such that it can run if one places in the same folder as the dataset *on any computer*.

1. Summarize the data. What can you say about the characteristics of the population are we studying? How might the composition of the sample affect affect the policy implications of the experiment?

2. In the analysis of a randomized experiment, an important first step involves verifying that treatment assignment was in fact random. The designers of the study are convinced that the initial randomization of students to the three classroom types was successful. Unfortunately, they did not keep their records of that initial assignment; the dataset only has information on the classroom type in which students actually enrolled. Why might initial assignment differ from actual enrollment, and how would that difference affect the proper interpretation of the study's results? Test whether classroom type was independent of baseline student characteristics. Were classroom types evenly distributed across the three racial groups? Were boys more or less likely than girls to be assigned to any type of classroom? What about children from poor households?

3. Estimate the effects of class size on reading and math scores by comparing small classes to regular classes. As a first step, estimate the means and standard deviations of test scores for small and regular

classes. Using these means and standard deviations, calculate the t-statistic for the null hypothesis that the means are equal. You may find Stata's `display` command helpful for performing calculations. Now perform this calculation using Stata's `ttest` command. Are your results the same? Interpret the results. Are the effects on reading and math scores of comparable magnitude? (Think carefully about the scaling of the outcome variables.)

4. Estimate the effects of teachers' aides on reading and math scores by comparing regular classes to regular classes with aides. Interpret your results as in question (3).

5. Education policymakers have expressed interest in whether the effects of class size on test scores vary with student background characteristics. Estimate the effect of class size on the combined test score for blacks and whites separately. Is the effect on black students statistically different from the effect on white students? Also test whether class size has different effects on students from poor and non-poor families.

6. In the tests above, what are you assuming about the independence of students' test scores? Is that assumption reasonable in this context? Can you think of an approach that would impose more conservative independence assumptions? If so, recalculate the t-statistics for the small-vs.-regular reading and math score comparisons. If the results differ, explain why. (Hint: you may want to use the `collapse` command.)

7. Until now, we have relied on t-tests rather than regressions. Can regression be used to produce the same results? If so, use it to reproduce the results from question (3) and question (6). Now that we are using regression, can you also think of another way to relax the independence assumption? If so, do it.

8. Are the effects of class size on test scores large or small? Compare your estimated effects with other (non-experimental) test score gaps in the data. For instance, how large are the effects relative to the black-white test score gap? Provide a few comparisons along these lines.