University of Oregon
Department of Economics

Professor Jason Lindo
Fall Quarter 2010

Economics 607
PROBLEM SET 1

Due Monday, Oct. 11

In this problem set, you will replicate – as best as possible using publicly available data – key results presented in Krueger (1999). Krueger (1999) estimates the effect of class size on student achievement using data from the Tennessee Student-Teacher Achievement Ratio (STAR) experiment. Over a four-year period starting the fall of 1985, Project STAR randomly assigned elementary school students and teachers to classes of different sizes.

To do these exercises, you will use the Stata data set stardata.dta. I will also provide you with the raw ascii data (webstar.dat) and the program that reads in this raw data (stardata.do). You will have to create additional variables from stardata.dta to conduct the analysis. The codebook is given in file readme-star.txt.

You should do this problem set by writing a program (a *.do file in Stata). In your solution packet, include a well-annotated version of this program, as well as relevant Stata output (e.g., key regression output, key graphs, etc.). (Do NOT include pages and pages of "undigested" Stata log files.) Place the .do files in an appendix.

Data notes:

- The data set contains 11,598 observations. The unit of observation is an individual student. Each student in the data set participated in Project STAR at some point between kindergarten and 3rd grade.

- Missing values are coded to 9, 99, 999, etc. Before conducting the analysis, you will therefore need to recode these values to missing (with a "."). (See the codebook for more details. Use the recode or replace command in Stata.)

- Outcome variables:
  mathk, readk: Stanford Achievement Test (SAT) math and reading scores in kindergarten
  math1, read1: SAT math and reading scores in 1st grade
  math2, read2: SAT math and reading scores in 2nd grade
  math3, read3: SAT math and reading scores in 3rd grade

  Note: These scores are in units that are not easily interpreted. As discussed in Krueger's article (pages 507-08, footnote 11), these scores should be converted to percentile scores before conducting the analysis. (Use the pctile command in Stata to calculate percentile ranks; be careful about the subsample you use. Use the xtile command, with the cutpoints option, to calculate percentile scores for the entire sample of STAR participants.)

- Treatment variables:
  stark, ctypek: In a STAR classroom and classroom type (small, regular, or regular with aide) in K
  star1, ctype1: In a STAR classroom and classroom type in 1st grade
  star2, ctype2: In a STAR classroom and classroom type in 2nd grade
  star3, ctype3: In a STAR classroom and classroom type in 3rd grade

- To see variable labels, use the describe command in Stata. (If desired, add names to values of categorical variables using label define and label values.)

1. Descriptive Statistics

a. Create the following variables from the data given in stardata.dta: receipt of free/reduced price lunch in each grade (indicators), white or Asian student race (indicator), age in 1985, an attrition indicator (as described in note d to Table 1), and the average math and reading percentile test score (as described in footnote 11). Present summary statistics on these variables for the full sample. [Hint: Use the **generate** command to create the variables and the **summarize** command to generate summary statistics. Be sure that missing values are recoded.]

b. Calculate means of these five new variables and class size by treatment status (small class, regular class, and regular class with aide) for new STAR participants only in each of kindergarten, 1st grade, 2nd grade, and 3rd grade. Round your answers to three decimal places. [Hint: Use the **table** command or the **tabsum** command in Stata. Your answers should be consistent with those presented in Table 1 of Krueger (1999).]

c. Test the null hypothesis that there are not significant differences in the means of these variables by treatment status. Give both F-statistics and p-values in your answer, each rounded to three decimal places. [Hint: In Stata, you would use the command **regress** and possibly **test** or **testparm**.]

d. Create indicator for whether a student's teacher is white or Asian and for whether a student's teacher has a master's degree. Using these variables and teacher experience, repeat the exercise in parts (a) and (b). Once again, restrict attention to new STAR participants only in each of kindergarten, 1st grade, 2nd grade, and 3rd grade, and round your answers to three decimal places.

2. Random Assignment?

a. Some researcher claims that the tests conducted in (1) support the assertion that the treatment was randomly assigned to students and teachers. Do you agree? Explain.

b. In fact, the treatment was randomly assigned to students and teachers within schools. For each of the variables employed in (1), test the null hypothesis that, conditional on school of attendance, there are no significant differences across treatment status. Once again, give both F-statistics and p-values in your answer, each rounded to three decimal places. [Hint: In Stata, you would once again use the command **regress**. You will also use the command **test** or **testparm**. **xi** is useful in creating dummy variables. Your results should be consistent with those presented in Table 2 of Krueger.]

c. Are the results in (b) consistent with the fact that the treatment was randomly assigned conditional on school of attendance? Explain.

3. Regression Estimates

a. Using as a dependent variable the average percentile math/reading test score constructed in (1), produce regression results similar to those given in columns (1) to (4) of Table 5 in Krueger (1999). In your regressions, assume that the error terms are iid. [Use the **xi** and **regress** commands.]

b. Interpret the coefficients on the small class and regular/aide class indicators in the first specification.

c. How do the coefficients on the small class and regular/aide class indicators change as more covariates are added to the model? Would you say that there is strong evidence of selection on observables?

d. How do the standard errors on the small class and regular/aide class indicators change as more

covariates are added to the model? Is this to be expected? Explain.

e. Suppose that test scores are a noisy measure of true skill. How will that affect the coefficient estimates?

4. Error structure
In (1) and (2), you (should have) conducted hypothesis tests under the assumption that regression errors were independent and identically distributed.

a. Redo the regressions in (3a), but cluster the standard errors on teacher (tchidk, tchid1, etc.) [Use the **cluster** option to **regress**.] Are the standard errors higher or lower than before? What does this suggest about the sign of error correlation among children in the same class? Is this what you would have expected? Explain.

b. Collapse the data to class-level means. Redo the regressions in (3a) using the aggregated data and weighting by class size. Does this approach deal with the "clustering" problem? Explain.

5. Intent to Treat and Instrumental Variables
One problem with focusing on actual class type is that there were transitions between class types after the first year of the program.

a. Construct a variable that gives each student's *initial* class assignment (i.e., assignment in K if entered program in K, assignment in 1st grade if entered program in 1st etc.). Then construct an indicator for whether an individual was no longer in his initial class type (though is still participating in the program) in each later grade. Krueger argues that transitions were nonrandom, though doesn't present any direct evidence on the characteristics of the switchers. Do transitions appear to be random? Calculate any relevant or useful statistics in your answer.

b. Krueger argues initial class assignment is *both* random (conditional on school) *and* highly correlated with actual class assignment in later years. The former was addressed in question (2). For individuals who entered the program for the first time in kindergarten, show that initial class assignment is a good predictor of actual class assignment in each 1st, 2nd, and 3rd grade.

c. The conditions given in (b) are those necessary for initial class assignment to be a valid instrument for actual class assignment. Using the Stata command ivreg, do a two-stage least squares regression of average percentile math/reading test score on actual class type dummies, using initial class type dummies as instruments. Present results for the same four specifications as in (3a) (and (4b)); cluster the standard errors on teacher. How do the results compare to those in found in (4b))? Do they suggest that non-random transitions were a problem?

6. External Validity
The success of Project STAR has been used to motivate larger-scale class size interventions. Do you think it is reasonable for policymakers to expect similar effects in different settings? Explain.