Homework Assignment 2

The assignment will be graded and solutions have to be turned in by email (ffossen@unr.edu) by **Tuesday**, **March 26th**. We will discuss the homework on Thursday that week. You may work in groups of up to 3 students. Turn in one properly formatted answer sheet per group and your program script (e.g., Stata do-file). Please include the names of all group members who have contributed to the solution.

Tennessee Student-Teacher Achievement Ratio (STAR) experiment

Problem description

In this problem set, you will replicate – as well as possible using publicly available data – key results on the Tennessee Student-Teacher Achievement Ratio (STAR) experiment presented in Krueger (1999). The STAR experiment was a pioneering randomized study from the field of education, designed to estimate the effects of smaller classes in primary school. The study was implemented for a cohort of kindergartners in 1985-86. It ran for four years, until the original cohort of kindergartners was in 3rd grade.

Economists have a long tradition of trying to establish causal links between class size and student learning. Many studies in the field of education using non-experimental data, however, suggest there is little or no link between class size and student learning.

In this problem set you are supposed to critically evaluate the experiment and to come up with estimates on the causal relationship between class size and learning.

Description of the data

The data base "Star13.dta" covers students who took part in the STAR experiment at some point between kindergarten and 3rd grade. It contains information on their birth year, race, treatment assigned, and on their learning outcomes. Further, information on their teachers' backgrounds is available:

Id	Fictitious student identification number
Sex	Sex (=1 if male, =2 if female)
Race	Race of the student (=1 if white, =2 if black, =3 if Asian, =4 if Hispanic, =5 if American Indian, =6 if other)
Mathk	Math score in Stanford Achievement Test (SAT) in kindergarten
Readk	Reading score in Stanford Achievement Test (SAT) in kindergarten
math1 (math2, math3)	SAT math scores in 1st (2nd, 3rd) grade
read1 (read2, read3)	SAT math scores in 1st (2nd, 3rd) grade
mathk_p (math1_p, math2_p, math3_p)	Percentile rank in math scores in kindergarten (1st, 2nd, 3rd grade) (see Krueger, 1999, pp. 507/508)

readk_p (read1_p, read2_p, read3_p)	Percentile rank in reading scores in kindergarten (1st, 2nd, 3rd grade) (see Krueger, 1999, pp. 507/508)
stark (star1, star2, star3)	Indicator whether student took part in the STAR experiment in kindergarten (1st, 2nd, 3rd grade) (=1 if yes, =2 if no)
ctypek (ctyp1, ctype2, ctype3)	Indicator of the treatment the student received in kindergarten (1st, 2nd, 3rd grade) (=1 if small class, =2 if regular class, =3 if regular class with aide)
csizek (csize1, csize2, csize3)	Class size in kindergarten (1st, 2nd, 3rd grade)
sesk (ses1, ses2, ses3)	Indicator of the student's social economic status in kindergarten (1st, 2nd, 3rd grade) (=1 if free lunch, =2 if non-free lunch)
yob	Year of birth
schidk (schid1, schid2, schid3)	School id

In total, the data contain 11,598 student observations. Your estimates may differ from the ones presented in Krueger (1999) as you are using a public use file instead of the original data. Qualitatively, however, your results should be consistent with the ones presented by Krueger.

Recoding missing variables and creating additional variables

a. In some variables, missing values are coded to 9, 99, 999 etc. You need to recode these values to missing (".").

b. Create dummy variables that indicate

(i.) in each grade (kindergarten, 1, 2, 3) students receiving a free lunch (freelunchk, freelunch1, freelunch2, freelunch3).

(ii.) in each grade (k, 1, 2) attrition (attritionk, attrition1, attrition2) as explained in note d. to Table 1.

(iii.) in each grade class type "small" (smallk, small1, small2, small3), class type "regular" (regulark, regular1, regular2, regular3), and class type "regular with aide" (regular_aidek, regular_aide1, regular_aide2, regular_aide3).

(iv.) students who entered star in kindergarten, in 1st, in 2nd, and in 3rd grade (enterk, enter1, enter2, enter3)

(v.) white or Asian students (combine both categories, whiteasian).

(vi.) girls (girl).

c. Create an age variable containing the student's age as of 31st December 1985 (age).

Question 1: Summary statistics

a. Calculate the share of students receiving free lunch, the share of white/Asian students, the average age in 1985, the attrition rate, average class size, and gender by the enter-variables

generated above and by treatment status (small class, regular class, and regular class with aide). Present your summary statistics for STAR participants in Table 1, structured as in Table I in Krueger (1999). You do not need to provide standard deviations.

b. Calculate the average of the percentile ranks of the math and reading tests for every individual in each year (name the variables testk etc.). (Hint: If one subtest score is missing, take the percentile score corresponding to the only available test as in Krueger (1999), fn.11.) Add the average values by the enter-variables and by treatment status to Table 1.

c. Comment on the characteristics of students assigned to the "small class" treatment, who entered STAR in kindergarten.

Question 2: Random assignment

The first question to ask about a randomized experiment is whether the randomization successfully balanced subjects' characteristics across the different treatment groups.

a. The STAR data does not include any pre-treatment test scores. Do you think that this is a problem? Explain briefly.

b. Compare the student characteristics collected in Table 1 across treatments. Formally test the null hypothesis of no difference across treatment groups using an *F*-test. Add both *F*-statistics (rounded to 2 digits after the decimal point) and *p*-values to Table 1. Do you think that randomization was successful, and why/why not? (Hint: Use the regress and test commands.)

c. In fact, the treatment was randomly assigned to students and teachers within schools. For each of the variables in Table 1, test the null hypothesis that, *conditional on school of attendance*, there are no significant differences across treatment groups. Once again, give both *F*-statistics and *p*-values and add them to Table 1. (Hint: Use school dummy variables.)

Are the results consistent with random assignment conditional on school attendance? Explain.

Question 3: OLS estimates of class size effects

a. Run OLS regressions with the average percentile test score (testk etc.) constructed in Question 1b as a dependent variable (assubming that the errors are iid). Produce regression results similar to those given in columns (1) to (3) of Table V in Krueger (1999) and present them in Table 2 A-D.

b. Interpret the coefficients on the small class and the regular/aide class indicators for kindergarten children in column (1) of Table 2.

c. Do the coefficients on the small class indicator change if additional explanatory variables are added to the model? What does this tell you about selection on observables?

d. Suppose that test scores are measures of true skills that are noisy (i.e., subject to measurement error). How do you expect this to affect your estimates of the coefficient? Explain briefly.

Question 4: Instrumental variable regression

So far, we have focused on actual class type. As described in Krueger (1999), there might be transitions between class types after the first year of the program. Students were moved between class types because of behavioral problems or parental complaints.

a. Reproduce Table IV in Krueger (1999) showing transitions between class types in adjacent grades. Display your results in Table 3.

b. Generate a variable that contains each student's initial assignment to a class type, i.e., assigned class type in kindergarten if a student entered STAR in kindergarten, assigned class type in 1st grade if a student entered the program in 1st grade etc. (name this variable ctype_assigned).

c. Krueger (1999) argues that initial class assignment is highly correlated with actual class assignment in later years. Show that initial class assignment is a good predictor of actual class treatment in each grade.

d. Consider using initial class assignment as an instrumental variable for actual class treatment. Name the two conditions that a valid instrument needs to fulfill. Do you think that the requirements are met by initial class assignment, and why/why not?

e. For each grade, run an instrumental variable regression of average math/reading test score on actual class type dummies (small class and regular with aide class), using initially assigned class type dummies as instruments (without further control variables). Report your results in Table 4. How do your results compare to those found in Table 2? Do they suggest that non-random transitions between class types were a problem?