LPO 7870 Research Design and Data Analysis II
In-Class Exercise
Randomized Controlled Trials

The Tennessee class size reduction experiment, known as Project STAR (Student-Teacher Achievement Ratio) was a 4-year experiment designed to evaluate the effects on learning of small class sizes. Students in participating schools were randomly assigned to one of three conditions: a regular class, with 22 to 25 students per class; a regular class with a teacher's aide; or a small class, with 13-17 students per class. "Regular class" was the business-as-usual (control) condition. Students were randomly assigned in kindergarten and were expected to remain in their class size assignment through 3rd grade. The experiment began during the 1985-86 academic year, and cost approximately $12 million.

Use the command below to read an extract of the public-use Project STAR data into Stata:

```
use https://github.com/spcorcor18/LPO-7870/raw/main/data/STARexample.dta, clear
```

1) Get to know the data: how many students participated in the study in kindergarten (FLAGSGK)? In 1st grade (FLAGSG1)? How many students who participated in kindergarten were still participating in 3rd grade (FLAGSG3)?

2) In a randomized controlled trial, the treatment and control groups should be equivalent, on average. Test for baseline differences in these groups, using the following observed student characteristics: *male, white, black, other* (race other than white or Black), *age, freelunch* (eligibility for free lunch in kindergarten), *sped* (special education status in kindergarten), and *kglocale1-4* (dummy variables for locale type in kindergarten). The treatment assignment variable in kindergarten is *gkclasstype*.

    a. Since there are three randomly assigned groups (regular class, small class, teachers' aide), you could regress each of the above characteristics on indicators for treatment assignment. (It would make the most sense to use the control group as the omitted category—e.g., `ib2.gkclasstype`.) Then test for joint significance of the groups.

    b. Technically, random assignment was *within school*. When treatment assignment is conditional on some observed variable, it is important to hold that variable constant when doing balance checks. You can do this by including indicator variables for each kindergarten school (*gkschid*). The `areg` command can help with this.

3) Check to see if students complied with their treatment assignment. Create a cross-tab that shows treatment assignment *gkclasstype* against the number of years the student was <u>actually</u> in a small class (*yearssmall*). Note not every student was observed all four years. You could limit this analysis to students observed in all four years (FLAGSKG-FLAGSG3).

4) What was the effect of treatment *assignment* on actual class size in kindergarten? Regress the actual class size (*gkclasssize*) on the treatment assignment variables. This will quantify just how different the class sizes were across groups.

5) Create a combined math and reading score for each grade K-3 (e.g., *gkscore* = *gktreadss* + *gktmathss*). Then, estimate the effects of the two treatments—small classes and regular class with teachers' aide—on this combined score. You can do this for all four grades, and compare your results to those in Stock and Watson Table 13.1.

6) Repeat part 5 for kindergarten, but add covariates, following Stock and Watson Table 13.2. How does the inclusion of covariates like gender, free lunch eligibility, race, and teacher's years of experience affect the treatment effect estimates?

    a. Again, randomization was *within school*. Technically, you should include indicator variables for each kindergarten school (*gkschid*).

7) Finally, when analyzing outcomes in later years it is important to consider *selective attrition*. The first thing you can do is look at systematic differences in participation in later years (FLAGSG1-FLAGSG3) or systematic differences in data availability in later years (flagg1-flagg3) by initial treatment assignment. To the extent there is differential attrition, you can then look at whether students who exited in each group were systematically different by baseline characteristics. How might you do this?

TABLE 13.1 Project STAR: Differences Estimates of Effect on Standardized Test Scores of Class Size Treatment Group

	Grade			
Regressor	K	1	2	3
Small class	13.90 (4.23) [5.48, 22.32]	29.78 (4.79) [20.24, 39.32]	19.39 (5.12) [9.18, 29.61]	15.59 (4.21) [7.21, 23.97]
Regular-sized class with aide	0.31 (3.77) [−7.19, 7.82]	11.96 (4.87) [2.27, 21.65]	3.48 (4.91) [−6.31, 13.27]	−0.29 (4.04) [−8.35, 7.77]
Intercept	918.04 (4.82)	1039.39 (5.82)	1157.81 (5.29)	1228.51 (4.66)
Number of observations	5786	6379	6049	5967

The regressions were estimated using the Project STAR public access data set described in Appendix 13.1. The dependent variable is the student's combined score on the math and reading portions of the Stanford Achievement Test. Standard errors, clustered at the school level, appear in parentheses, and 95% confidence intervals appear in brackets.

TABLE 13.2 Project STAR: Differences Estimates with Additional Regressors for Kindergarten

Regressor	(1)	(2)	(3)	(4)
Small class	13.90 (4.23) [5.48, 22.32]	14.00 (4.25) [5.55, 22.46]	15.93 (4.08) [7.81, 24.06]	15.89 (3.95) [8.03, 23.74]
Regular-sized class with aide	0.31 (3.77) [−7.19, 7.82]	−0.60 (3.84) [−8.25, 7.05]	1.22 (3.64) [−6.04, 8.47]	1.79 (3.60) [−5.38, 8.95]
Teacher's years of experience		1.47 (0.44) [0.60, 2.34]	0.74 (0.35) [0.04, 1.45]	0.66 (0.36) [−0.05, 1.37]
Boy				−12.09 (1.54)
Free lunch eligible				−34.70 (2.47)
Black				−25.43 (4.52)
Race other than black or white				−8.50 (12.64)
School indicator variables?	no	no	yes	yes
$\overline{R}^2$	0.01	0.02	0.22	0.28
Number of observations	5786	5766	5766	5748

The regressions were estimated using the Project STAR public access data set described in Appendix 13.1. The dependent variable is the student's combined test score on the math and reading portions of the Stanford Achievement Test. All regressions include an intercept (not reported). The number of observations differs in the different regressions because of some missing data. Standard errors, clustered at the school level, appear in parentheses, and 95% confidence intervals appear in brackets.