Experimental Estimates of the Effect of Reduction: The Tennessee Class Size Experiment (SW Section 11.4)

Project STAR (Student-Teacher Achievement Ratio)

- 4-year study, $12 million, 11600 students affected, 80 schools participated
- Upon entering the school system, a student was randomly assigned to one of three groups:
    - regular class (22 – 25 students)
    - regular class + aide
    - small class (13 – 17 students)
- regular class students re-randomized after first year to regular or regular+aide
- Y = Stanford Achievement Test scores

Deviations from experimental design

- Partial compliance:
  - 10% of students switched treatment groups because of "incompatibility" and "behavior problems" – how much of this was because of parental pressure?
  - Newcomers: incomplete receipt of treatment for those who move into district after grade 1
- Attrition
  - students move out of district
  - students leave for private/religious schools

Regression analysis

- The "differences" regression model:

$$Y_i = \beta_0 + \beta_1 SmallClass_i + \beta_2 RegAide_i + u_i$$

where $SmallClass_i = 1$ if in a small class

$RegAide_i = 1$ if in regular class with aide

TABLE 11.1 **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** (2.45)	29.78** (2.83)	19.39** (2.71)	15.59** (2.40)
Regular size with aide	0.31 (2.27)	11.96** (2.65)	3.48 (2.54)	−0.29 (2.27)
Intercept	918.04** (1.63)	1,039.39** (1.78)	1,157.81** (1.82)	1,228.51** (1.68)
Number of observations	5,786	6,379	6,049	5,967

The regressions were estimated using the Project STAR Public Access Data Set described in Appendix 11.1. The dependent variable is the student's combined score on the math and reading portions of the Stanford Achievement Test. Standard errors are given in parentheses under the coefficients. **The individual coefficient is statistically significant at the 1% significance level using a two–sided test.

TABLE 11.2 Project STAR: Differences Estimates with Additional Regressors for Kindergarten

Regressor	(1)	(2)	(3)	(4)
Small class	13.90** (2.45)	14.00** (2.45)	15.93** (2.24)	15.89** (2.16)
Regular size with aide	0.31 (2.27)	−0.60 (2.25)	1.22 (2.04)	1.79 (1.96)
Teacher's years of experience		1.47** (0.17)	0.74** (0.17)	0.66** (0.17)
Boy				−12.09** (1.67)
Free lunch eligible				−34.70** (1.99)
Black				−25.43** (3.50)
Race other than black or white				−8.50 (12.52)
Intercept	918.04** (1.63)	904.72** (2.22)		
School indicator variables?	no	no	yes	yes
$\overline{R}^2$	0.01	0.02	0.22	0.28
Number of observations	5,786	5,766	5,766	5,748

The regressions were estimated using the Project STAR Public Access Data Set described in Appendix 11.1. The dependent variable is the combined test score on the math and reading portions of the Stanford Achievement Test. The number of observations differ in the different regressions because of some missing data. Standard errors are given in parentheses under coefficients. The individual coefficient is statistically significant at the *5% level or **1% significance level using a two-sided test.

How big are these estimated effects?

- Put on same basis by dividing by std. dev. of Y
- Units are now standard deviations of test scores

TABLE 11.3 Estimated Class Size Effects in Units of Standard Deviations of the Test Score Across Students

Treatment Group	Grade			
	K	1	2	3
Small class	0.19** (0.03)	0.33** (0.03)	0.23** (0.03)	0.21** (0.03)
Regular size with aide	0.00 (0.03)	0.13** (0.03)	0.04 (0.03)	0.00 (0.03)
Sample standard deviation of test scores (s_Y)	73.70	91.30	84.10	73.30

The estimates and standard errors in the first two rows are the estimated effects in Table 11.1, divided by the sample standard deviation of the Stanford Achievement Test for that grade (the final row in this table), computed using data on the students in the experiment. Standard errors are given in parentheses under coefficients. **The individual coefficient is statistically significant at the 1% significance level using a two-sided test.

How do these estimates compare to those from the California, Mass. observational studies? (Ch. 4 – 7)

TABLE 11.4 Estimated Effects of Reducing the Student-Teacher Ratio by 7.5 Based on the STAR Data and the California and Massachusetts Observational Data

Study	$\hat{\beta}_1$	Change in Student-Teacher Ratio	Standard Deviation of Test Scores Across Students	Estimated Effect	95% Confidence Interval
STAR (grade K)	−13.90** (2.45)	Small class vs. regular class	73.8	0.19** (0.03)	(0.13, 0.25)
California	−0.73** (0.26)	−7.5	38.0	0.14** (0.05)	(0.04, 0.24)
Massachusetts	−0.64* (0.27)	−7.5	39.0	0.12* (0.05)	(0.02, 0.22)

The estimated coefficient $\hat{\beta}_1$ for the STAR study is taken from column (1) of Table 11.2. The estimated coefficients for the California and Massachusetts studies are taken from the first column of Table 7.3. The estimated effect is the effect of being in a small class versus a regular class (for STAR) or the effect of reducing the student-teacher ratio by 7.5 (for the California and Massachusetts studies). The 95% confidence interval for the reduction in the student-teacher ratio is this estimated effect ±1.96 standard errors. Standard errors are given in parentheses under estimated effects. The estimated effects are statistically significantly different from zero at the *5% level or **1% significance level using a two-sided test.

Summary: The Tennessee Class Size Experiment

Remaining threats to internal validity

- partial compliance/incomplete treatment
    - can use TSLS with $Z =$ initial assignment
    - Turns out, TSLS and OLS estimates are similar (Krueger (1999)), so this bias seems not to be large

Main findings:

- The effects are small quantitatively (same size as gender difference)
- Effect is sustained but not cumulative or increasing biggest effect at the youngest grades