

K-12 Education

Economics of Public and Social Issues

Jonathan Moreno-Medina

ECO3253, UTSA

9/13/23

Plan for today

- ▶ K-12 Education Review
- ▶ Education Production Function

K-12 Education

K-12 Education: Background

- ▶ U.S. spends nearly \$1 trillion per year on K-12 education
- ▶ Decentralized system with substantial variation across schools
 - ▶ Public schools funded by local property taxes sharp differences in funding across areas
 - ▶ Private schools and growing presence of charter schools

K-12 Education: Overview

- ▶ Main question: how can we maximize the effectiveness of this system to produce the best outcomes for students?
 - ▶ Traditional approach to study this question: qualitative work in schools
 - ▶ More recent approach: analyzing big data to evaluate impacts

Reference 1: Chetty, Friedman, Hilger, Saez, Schanzenbach, Yagan. "How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project STAR" QJE 2011.

Reference 2: Reardon, Kalogrides, Fahle, Shores. "The Geography of Racial/Ethnic Test Score Gaps." Stanford CEPA Working Paper 2016

Reference 3: Fredriksson, Ockert, Oosterbeek. "Long-Term Effects of Class Size." QJE 2012

Reference 4: Chetty, Friedman, Rockoff. "Measuring the Impacts of Teachers I and II" AER 2014

Using Test Score Data to Study K-12 Education

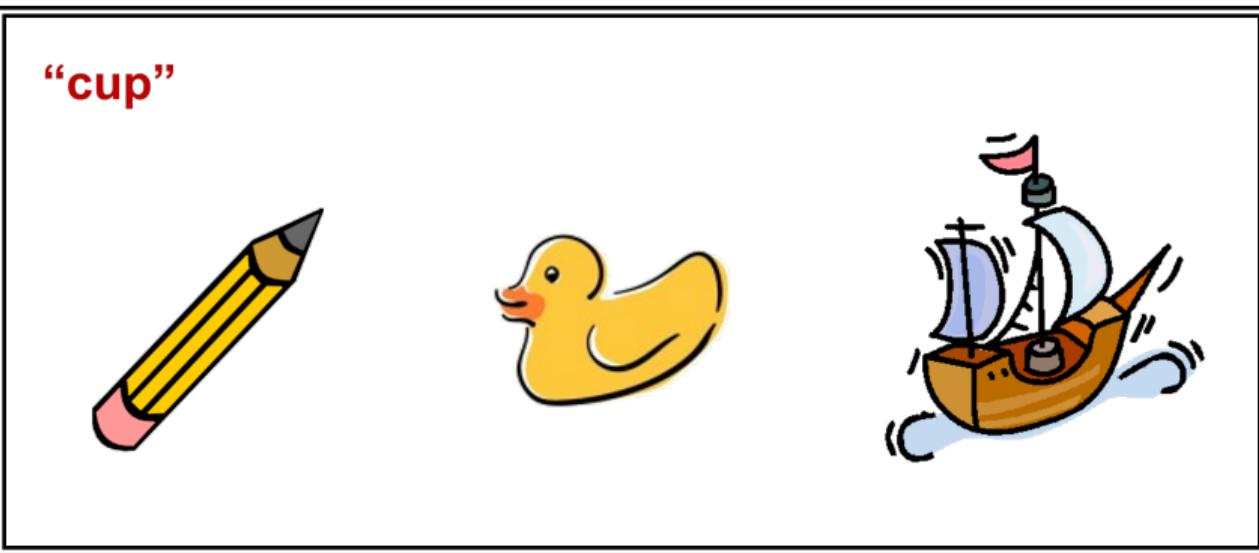
- ▶ Primary source of big data on education: standardized test scores obtained from school districts
 - ▶ Quantitative outcome recorded in existing administrative databases for virtually all students
 - ▶ Observed much more quickly than long-term outcomes like college attendance and earnings

Using Test Score Data to Evaluate Primary Education

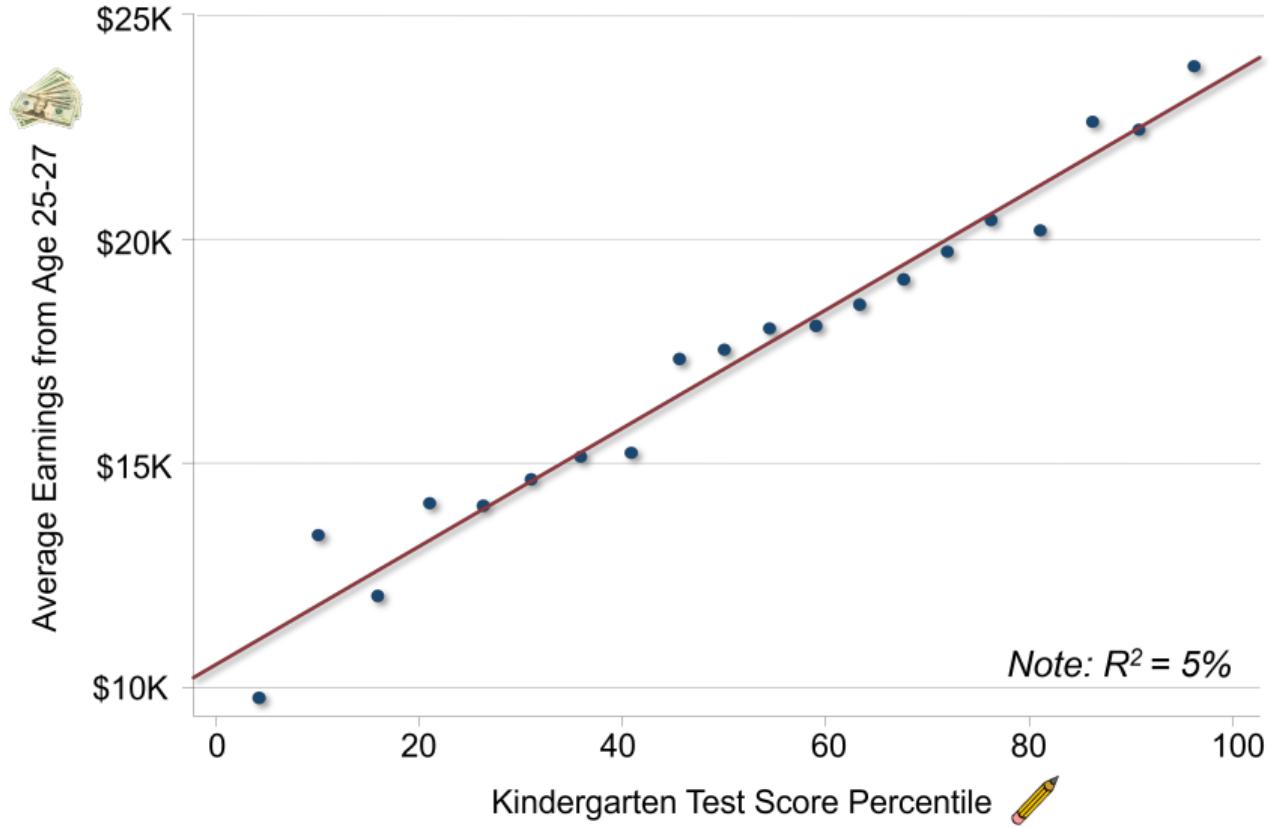
- ▶ Common concern: are test scores a good measure of learning?
 - ▶ Do improvements in test scores reflect better test-taking ability or acquisition of skills that have value later in life?
- ▶ Chetty et al. (2011) examine this issue using data on 12,000 children who were in Kindergarten in Tennessee in 1985
 - ▶ Link school district and test score data to tax records
 - ▶ Ask whether KG test score performance predicts later outcomes

A Kindergarten Test

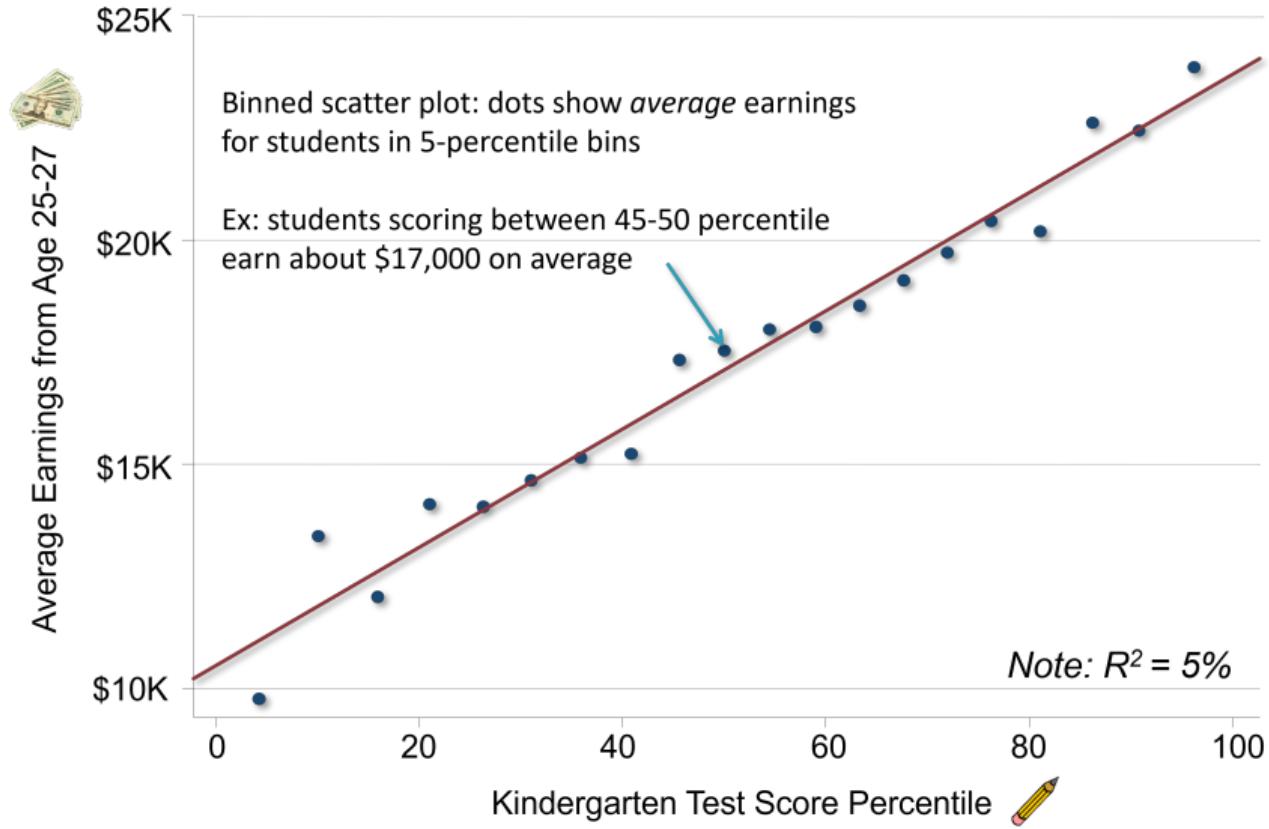
- ▶ I'll say a word to you. Listen for the ending sound.
- ▶ You circle the picture that starts with the same sound



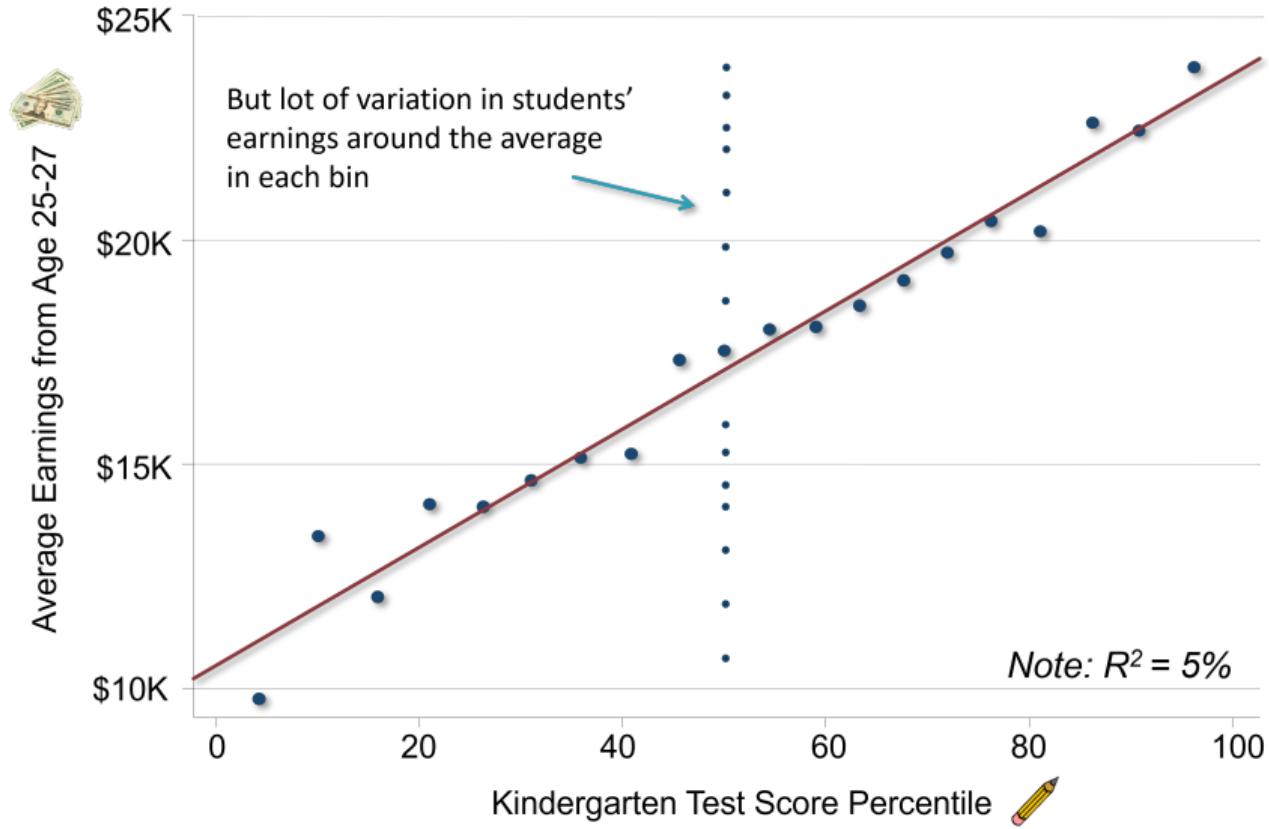
Earnings vs. Kindergarten Test Score



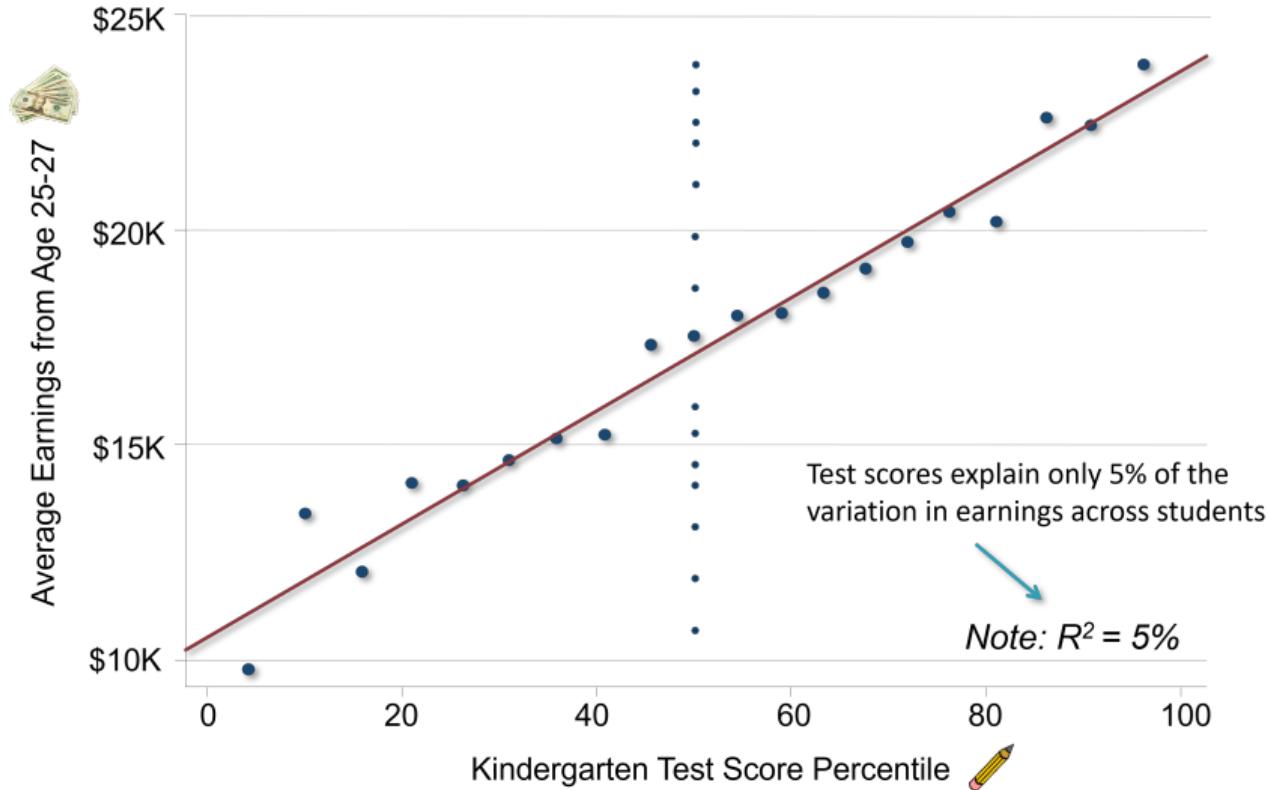
Earnings vs. Kindergarten Test Score



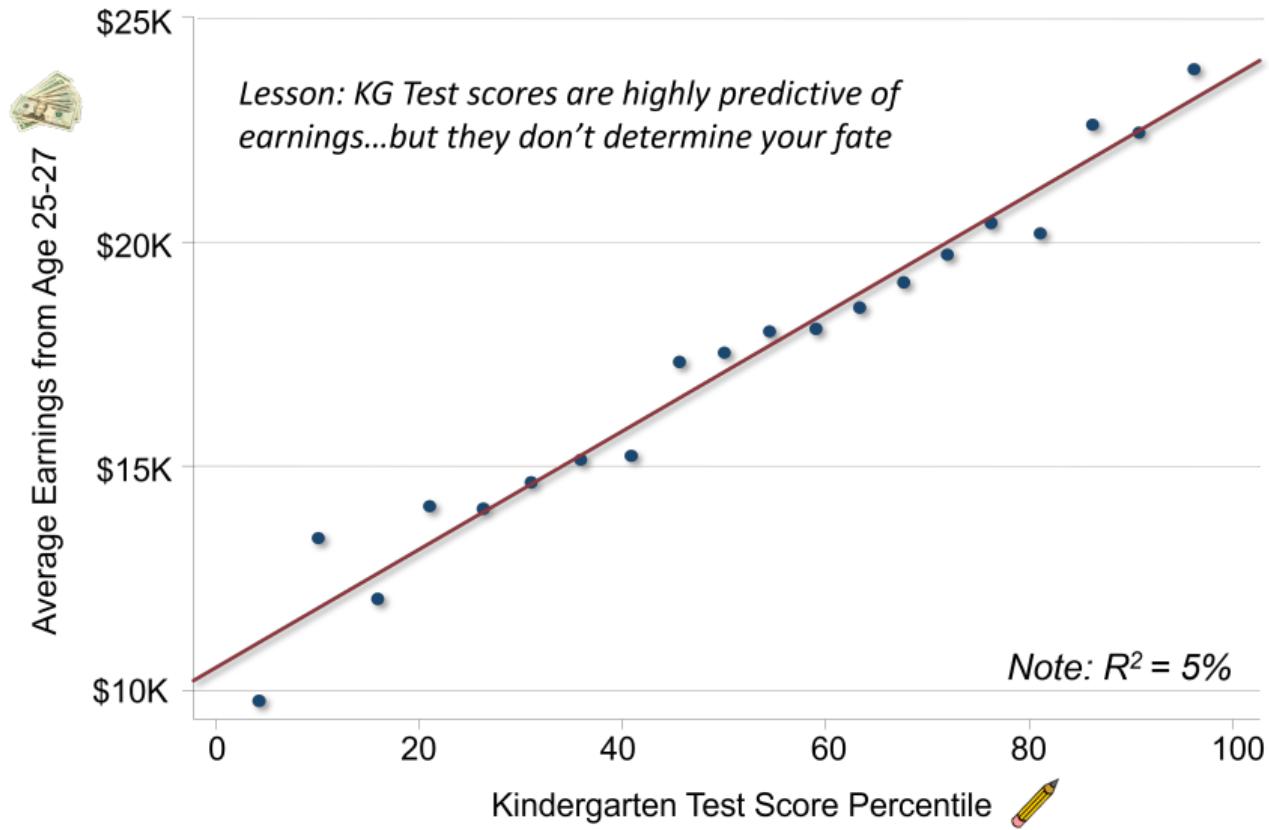
Earnings vs. Kindergarten Test Score



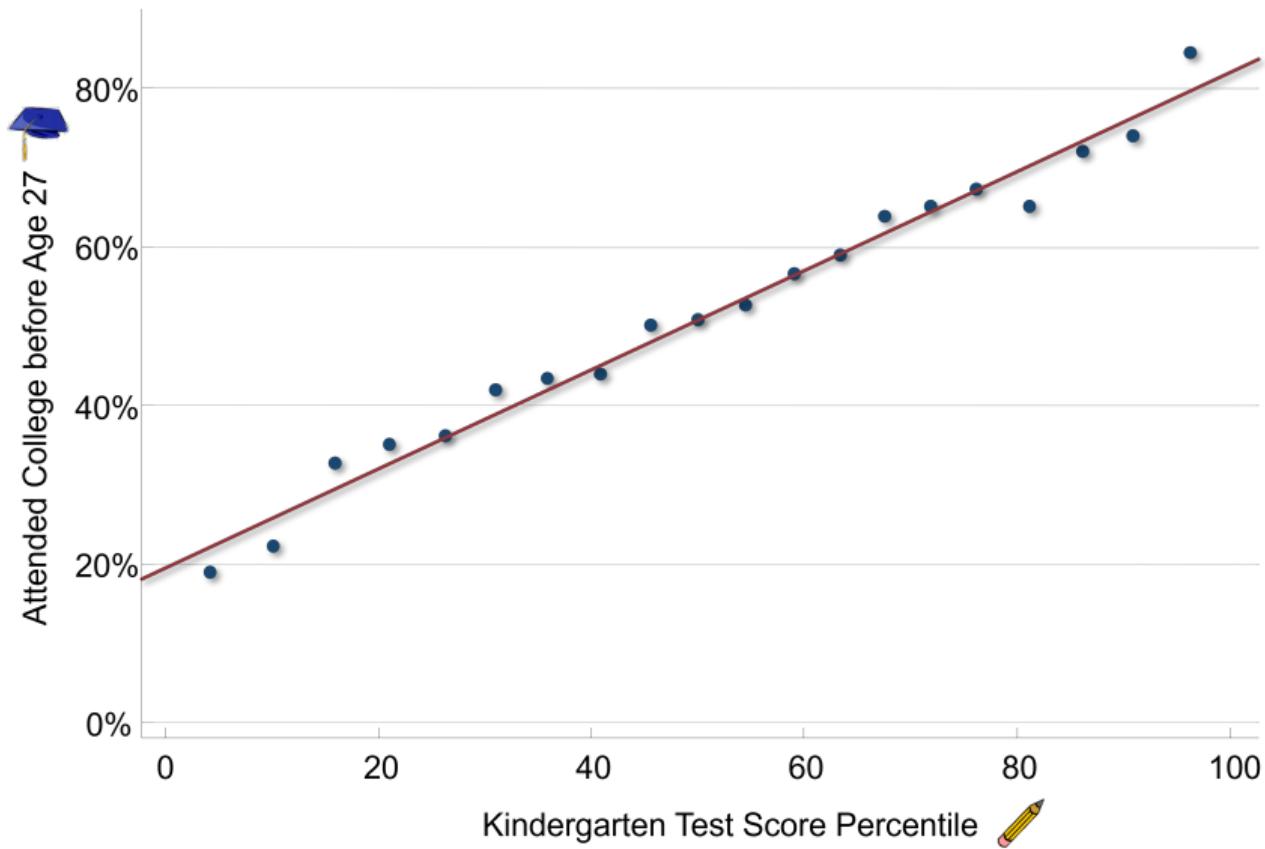
Earnings vs. Kindergarten Test Score



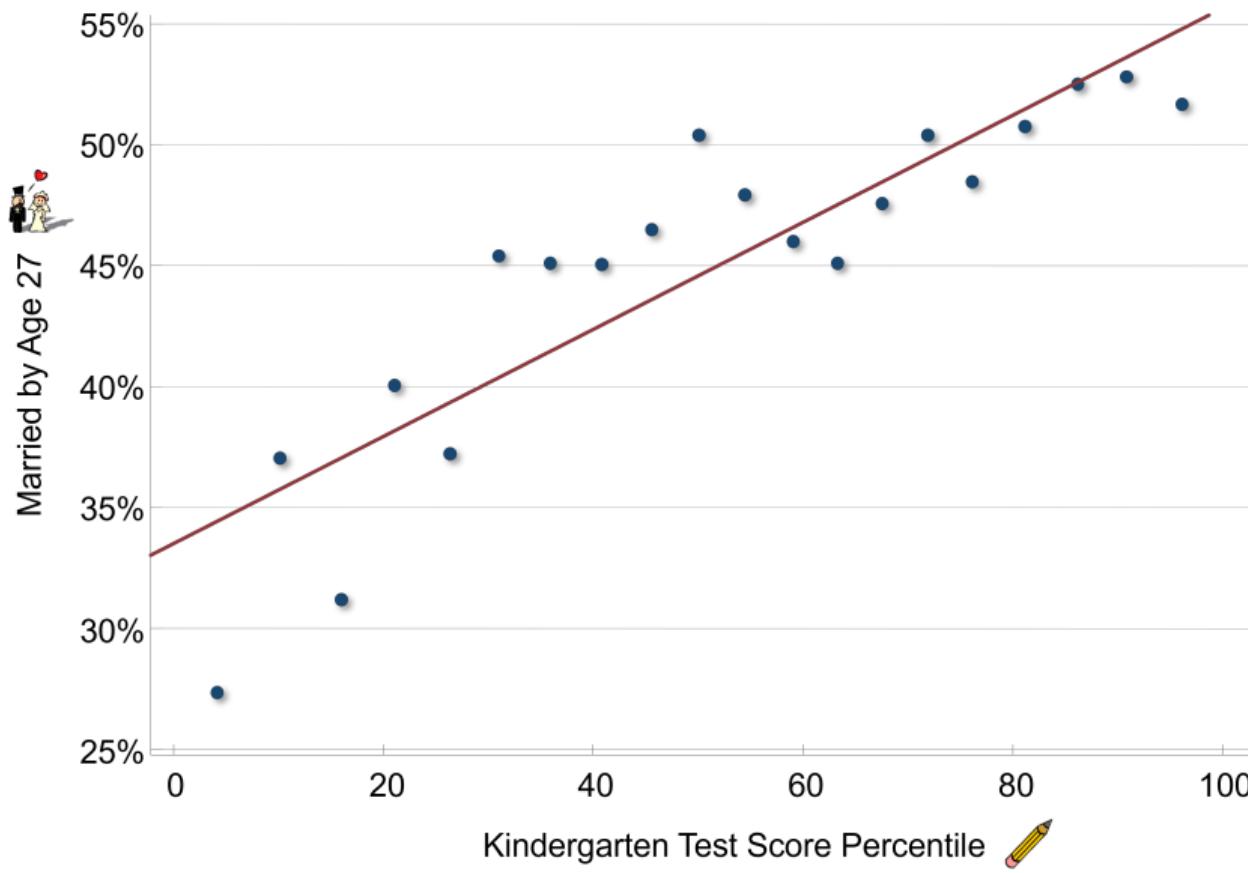
Earnings vs. Kindergarten Test Score



College Attendance Rates vs. KG Test Score



Marriage by Age 27 vs. KG Test Score



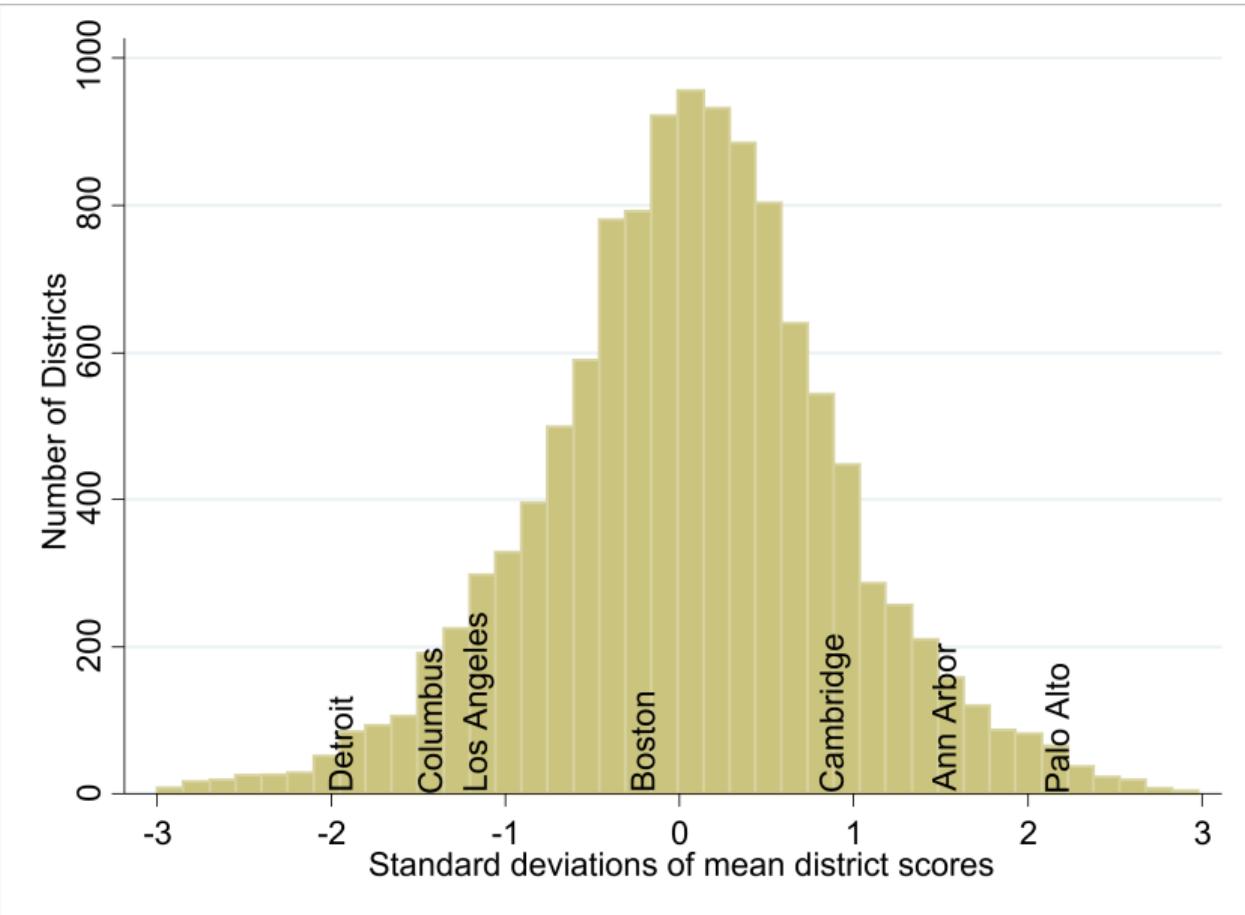
Studying Differences in Test Score Outcomes

- ▶ Test scores can provide a powerful data source to compare performance across schools and subgroups (e.g., poor vs. rich)
- ▶ Problem: tests are not the same across school districts and grades
 - ▶ → makes comparisons very difficult
- ▶ Reardon et al. (2016) solve this problem and create a standardized measure of test score performance for all schools in America
 - ▶ Use 215 million test scores for students from 11,000 school districts across the U.S. from 2009-13 in grades 3-8

Making Test Score Scales Comparable Across the U.S.

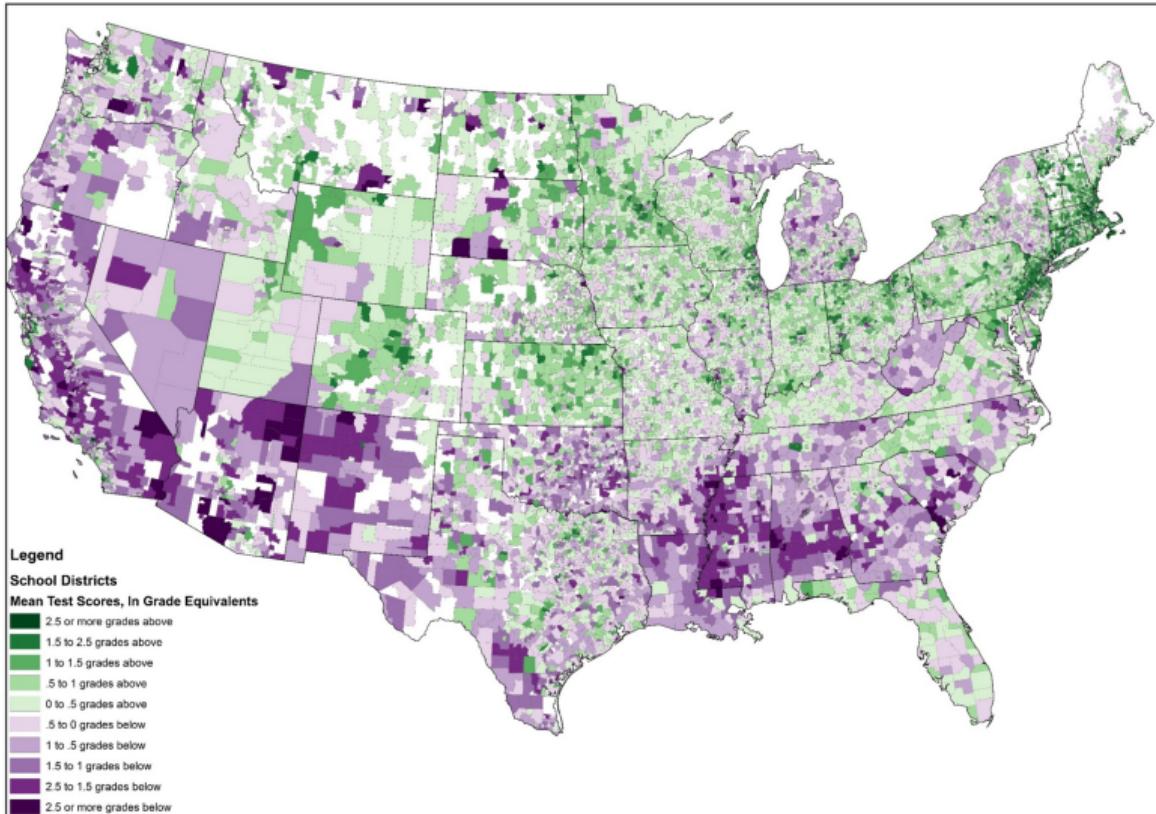
- ▶ Convert test scores to a single national scale in three steps:
 1. Rank each school district's average scores in the statewide distribution (for a given grade-year-subject)
 2. Use data from a national test administered to a sample of students by Dept. of Education to convert state-specific rankings to national scale
 - ▶ Ex: suppose CA students score 5 percentiles below national average
 - ▶ Then a CA school whose mean score is 10 percentiles below CA mean is 15 percentiles below national mean
 3. Convert mean test scores to “grade level” equivalents

Nationwide District Achievement Variation, 2009-2013

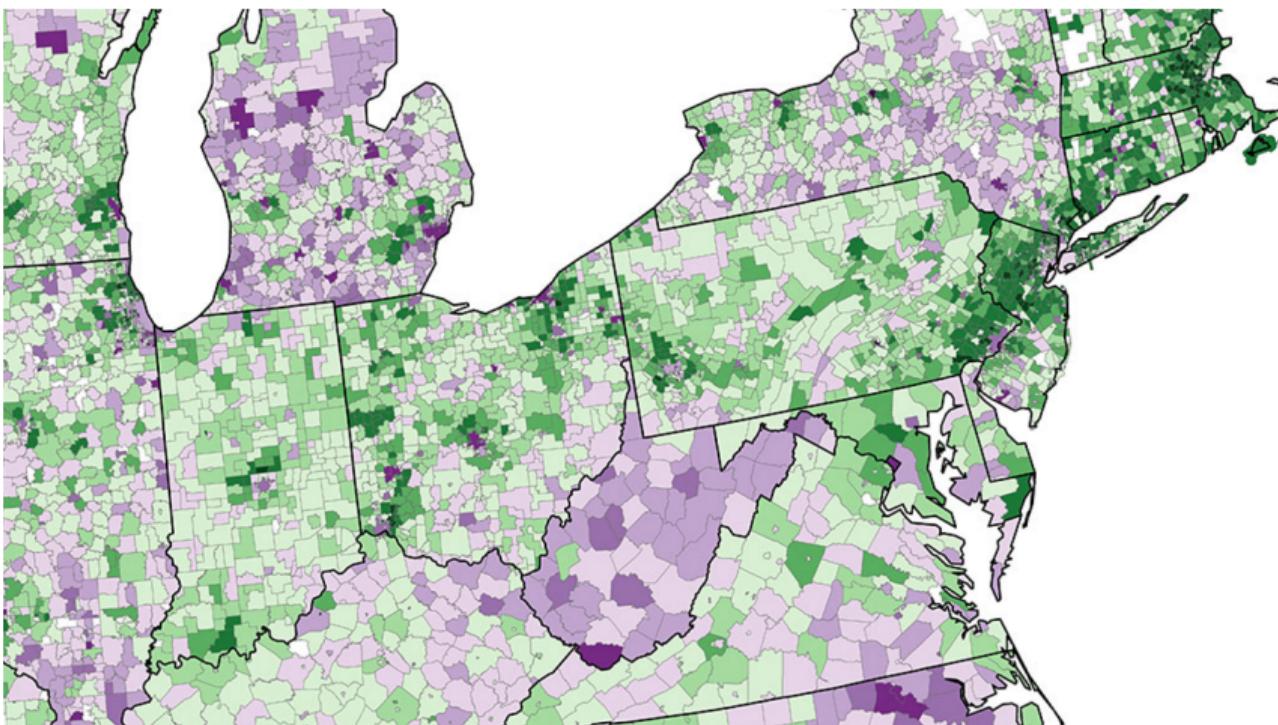


Geography of District Achievement Variation, 2009-2013

Average Test Scores, by School District, Grades 3-8, 2009-2013



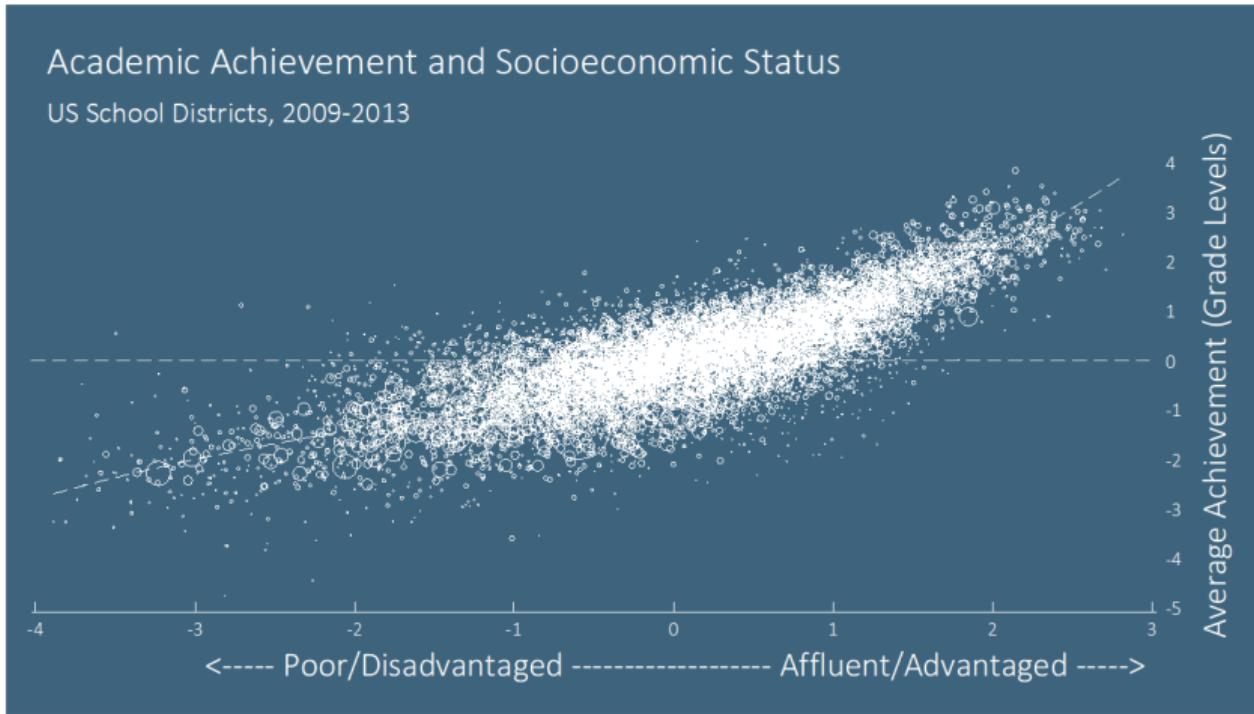
Geography of District Achievement Variation, 2009-2013



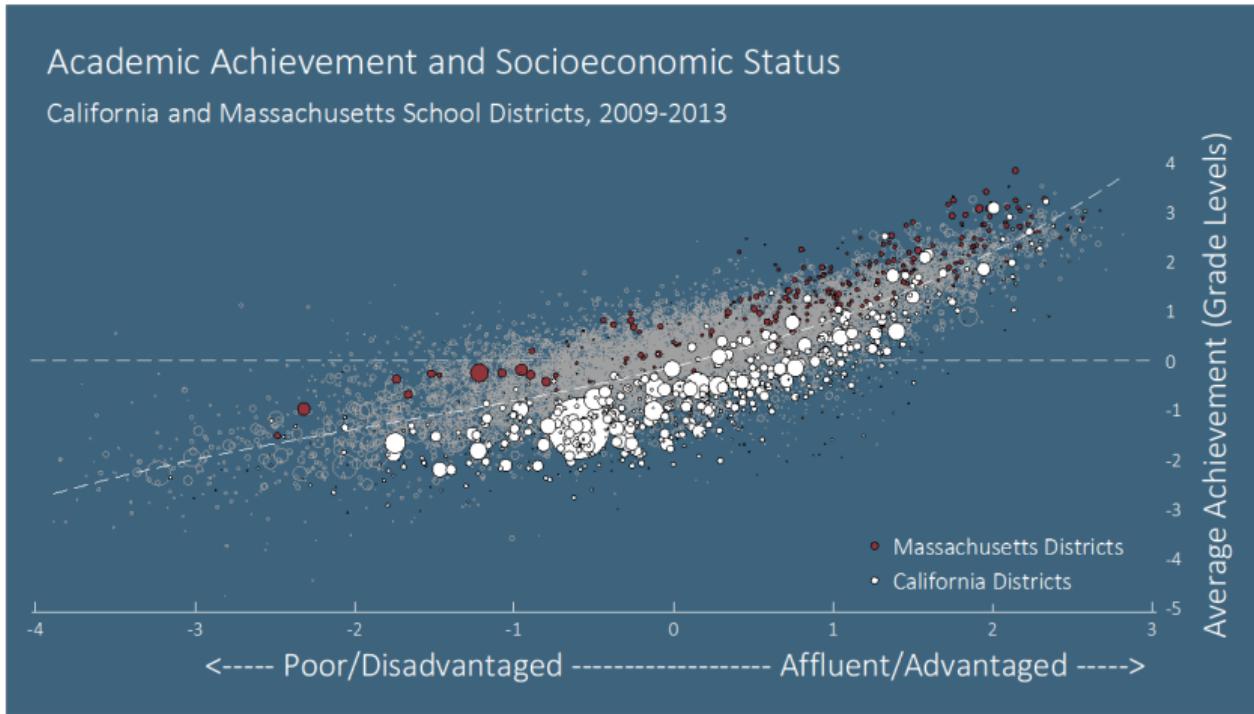
Achievement Gaps in Test Scores by Socioeconomic Status

- ▶ Next, use these data to examine how test scores vary across socioeconomic groups
 - ▶ Define an index of socioeconomic status (SES) using Census data on income, fraction of college graduates, single parent rates, etc.

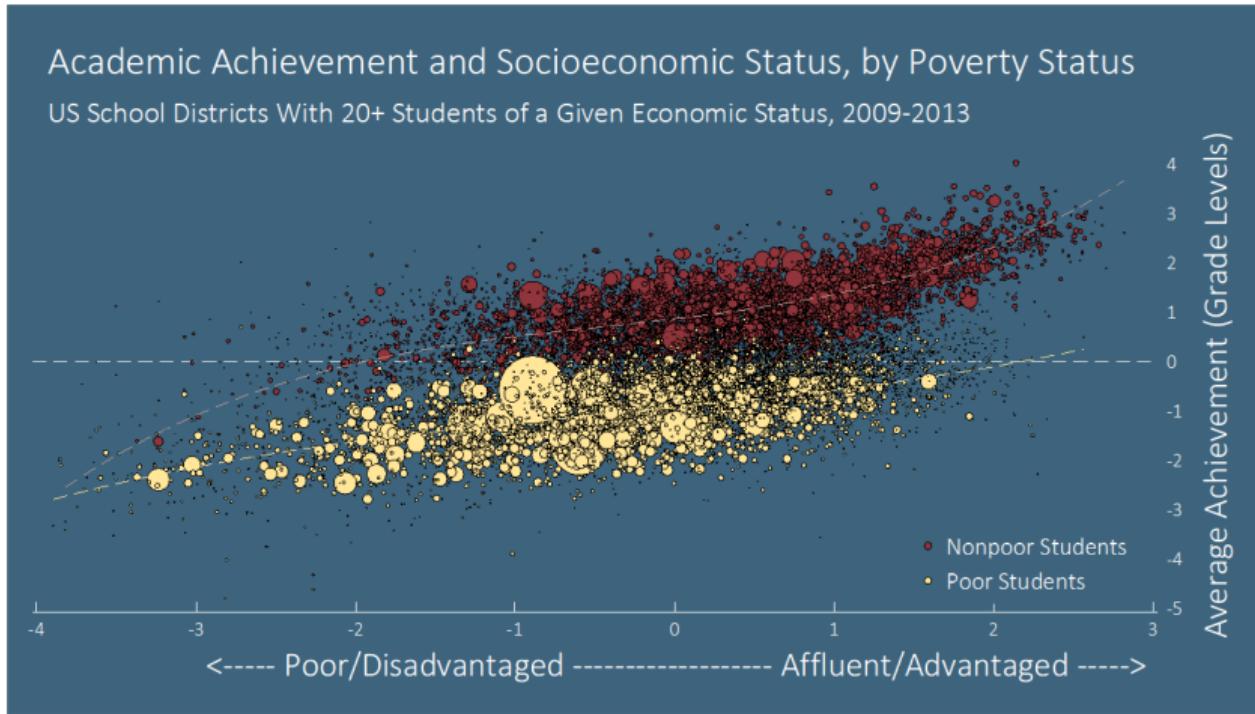
Academic Achievement and SES



Academic Achievement and SES



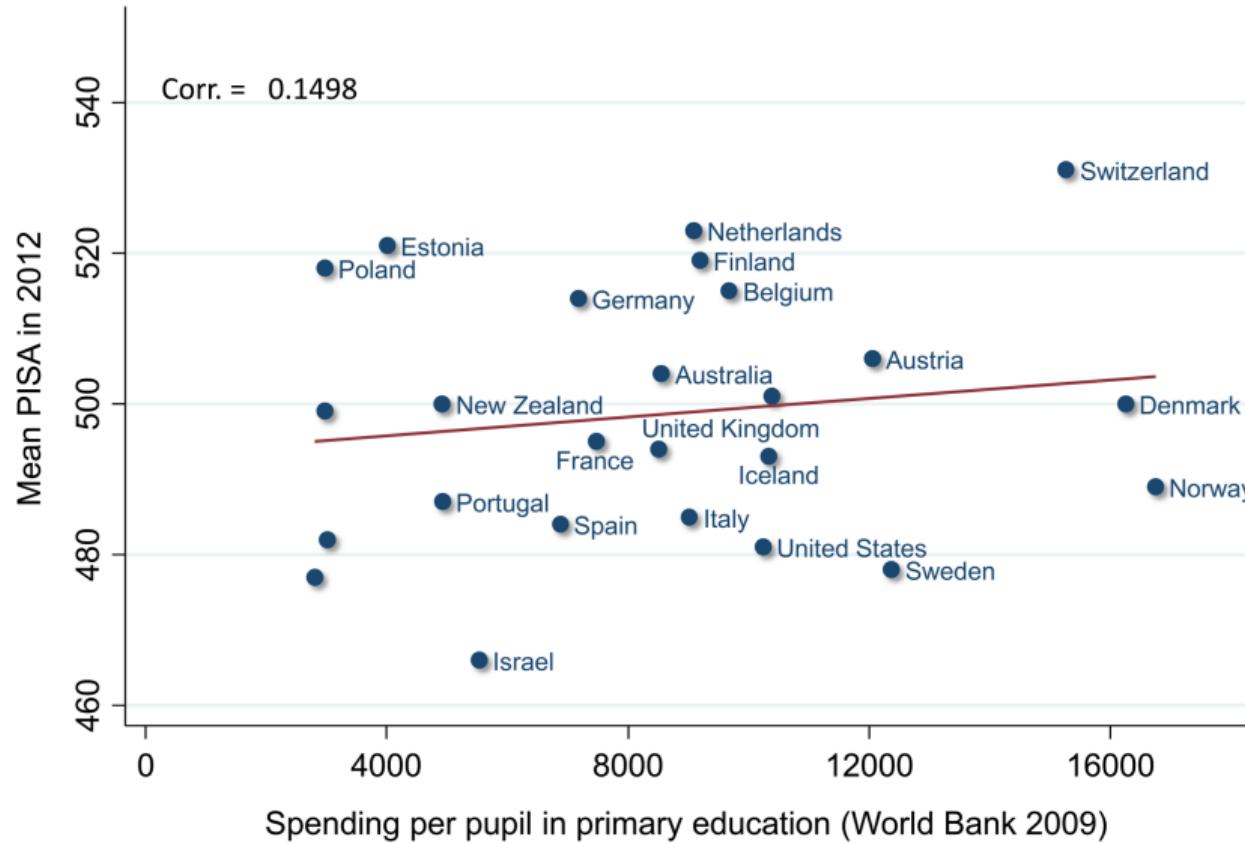
Academic Achievement and SES



How Can We Improve Poorly Performing Schools?

- ▶ There are many school districts in America where students are two grade levels behind national average, controlling for SES
- ▶ How can we improve performance in these schools?
 - ▶ Simply spending more money on schools is not necessarily the solution...

Test Scores vs. Expenditures on Primary Education Across Countries



Two Policy Paradigms to Improve Schools

- ▶ Two distinct policy paradigms to improve schools
 1. Government-based solutions: improve public schools by reducing class size, increasing teacher quality, etc.
 2. Market-based solutions: charter schools or vouchers for private schools
- ▶ Contentious policy debate between these two approaches
- ▶ We will consider each approach in turn

Government-Based Solutions: Improving Schools

Improving Schools: The Education Production Function

- ▶ Improving public schools requires understanding the education production function
- ▶ How should we change schools to produce better outcomes?
 - ▶ Better Teachers?
 - ▶ Smaller Classes?
 - ▶ Better Technology?

Effects of Class Size: Tennessee STAR Experiment

- ▶ Student/Teacher Achievement Ratio (STAR) experiment
 - ▶ Conducted from 1985 to 1989 in Tennessee
 - ▶ About 12,000 children in grades K-3 at 79 schools
- ▶ Students and teachers randomized into classrooms within schools
 - ▶ Class size differs: small (~15 students) or large (~22 students)
 - ▶ Classes also differ in teachers and peers

Effects of Class Size: Tennessee STAR Experiment

- ▶ Evaluate impacts of STAR experiment by comparing mean outcomes of students in small vs. large classes.
- ▶ Report impacts using regressions of outcomes on an indicator (0-1 variable) for being in a small class [Krueger 1999, Chetty et al. 2011]

STAR Experiment: Impacts of Class Size

Dep Var: Outcome	Test Score	College Attendance	Earnings
	(1)	(2)	(3)
Small Class	4.81 (1.05)	2.02% (1.10%)	-\$4 (\$327)
Observations	9,939	10,992	10,992
Mean of Dep. Var.	48.67	26.4%	\$15,912

STAR Experiment: Impacts of Class Size

Dep Var:	Test Score (1)	College Attendance (2)	Earnings (3)
Small Class	4.81 (1.05)	2.02% (1.10%)	-\$4 (\$327)
Observations	9,939	10,992	10,992
Mean of Dep. Var.	48.67	26.4%	\$15,912

*Estimated impact of being in a small KG class:
4.81 percentile gain in end-of-KG test score*

STAR Experiment: Impacts of Class Size

Dep Var:	Test Score (1)	College Attendance (2)	Earnings (3)
Small Class	4.81 (1.05)	2.02% (1.10%)	-\$4 (\$327)
Observations	Standard Error 9,939	10,992	10,992
Mean of Dep. Var.	48.67	26.4%	\$15,912

95% confidence interval = estimate +/- 1.96 times standard error
→ 95% CI for test score impact = 2.71 to 6.91 percentiles

If we repeat experiment 100 times, 95 of the 100 estimates

STAR Experiment: Impacts of Class Size

Dep Var:	Test Score (1)	College Attendance (2)	Earnings (3)
Small Class	4.81 (1.05)	2.02% (1.10%)	-\$4 (\$327)
Observations	9,939	10,992	10,992
Mean of Dep. Var.	48.67	26.4%	\$15,912
Mean Value of Outcome			

STAR Experiment: Impacts of Class Size

Dep Var:	Test Score (1)	College Attendance (2)	Earnings (3)
Small Class	4.81 (1.05)	2.02% (1.10%)	-\$4 (\$327)
Observations	9,939	10,992	10,992
Mean of Dep. Var.	48.67	26.4%	\$15,912

STAR Experiment: Impacts of Class Size

Dep Var:	Test Score (1)	College Attendance (2)	Earnings (3)
Small Class	4.81 (1.05)	2.02% (1.10%)	-\$4 (\$327)
Observations	9,939	10,992	10,992
Mean of Dep. Var.	48.67	26.4%	\$15,912

95% confidence interval: (-\$645, \$637)

→ Earnings impact could be as large as **\$637** (a 4% increase)

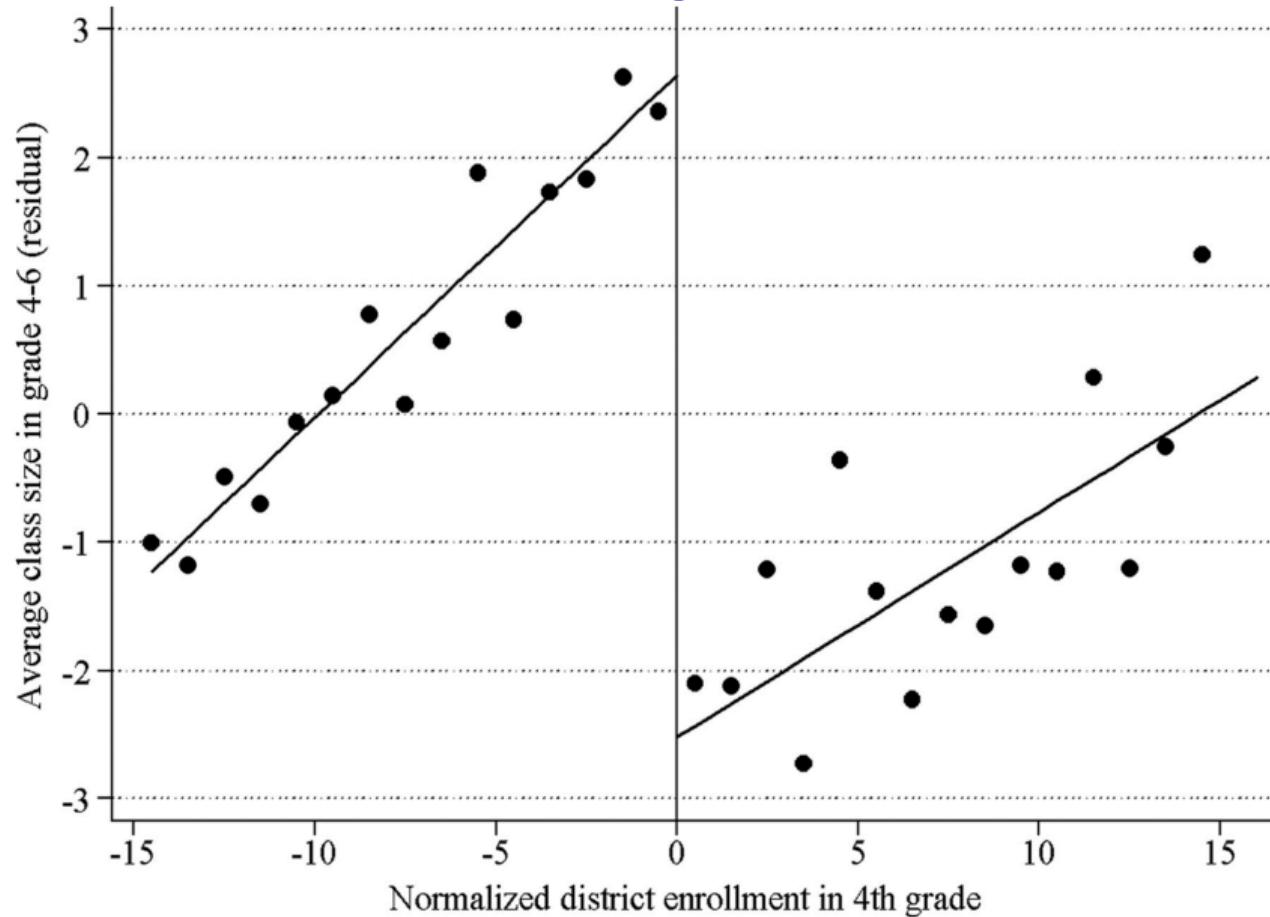
Effects of Class Size: Quasi-Experimental Evidence

- ▶ Limitation of STAR experiment: sample is too small to estimate impacts of class size on earnings precisely
- ▶ Fredriksson et al. (2013) use administrative data from Sweden to obtain more precise estimates
 - ▶ No experiment here; instead use a quasi-experimental method: regression discontinuity.

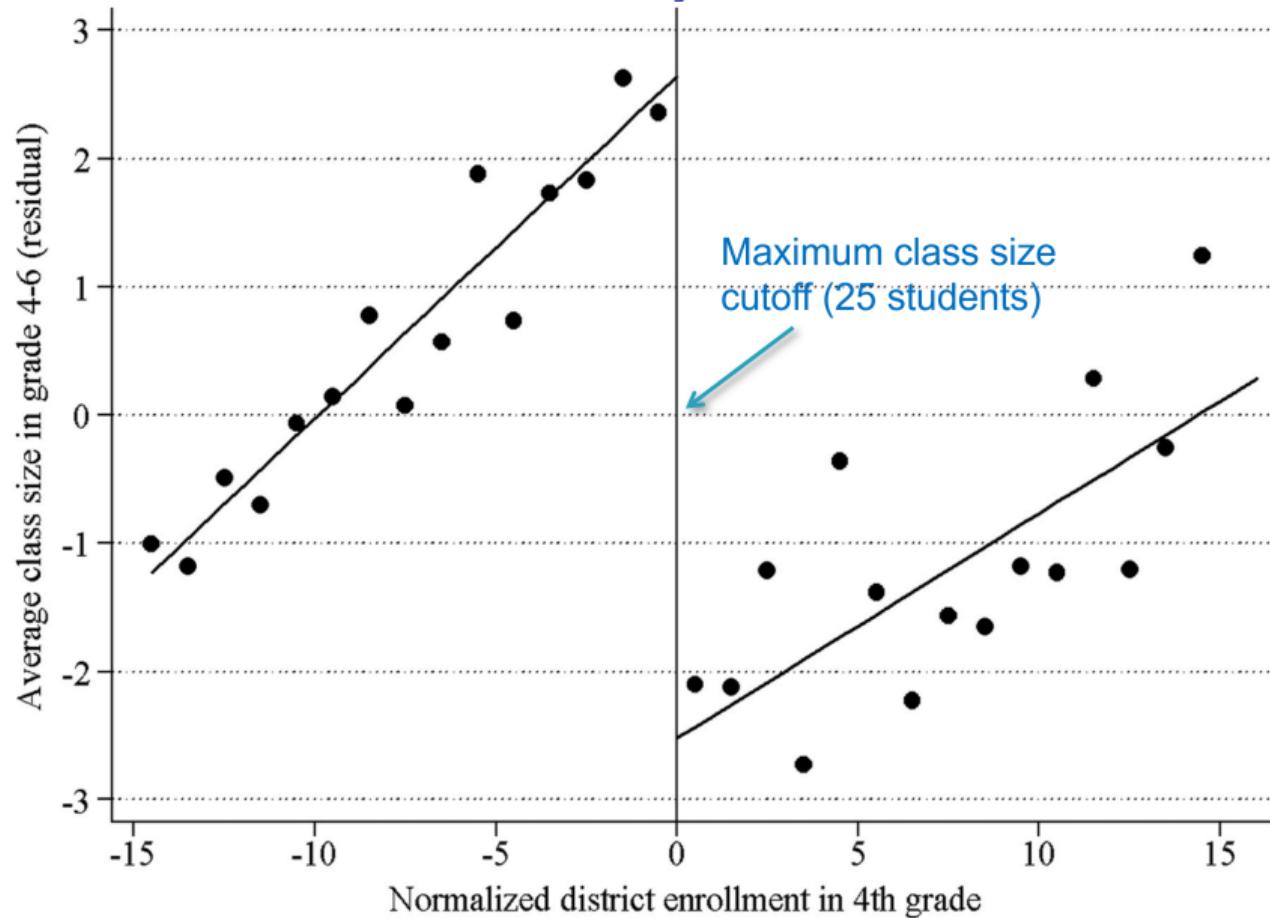
Regression Discontinuity Using Class Size Cutoffs

- ▶ Sweden imposes a maximum class size of 25 students
 - ▶ School that has 26 students in a given grade will therefore have two classes of 13 students each
 - ▶ School that has 25 students may have one class of 25 students
 - ▶ Schools that have 26 students in a grade are likely to be comparable to those that have 25 students
 - ▶ → Can identify causal effects of class size by comparing outcomes in schools with 26 vs. 25 students in a given grade

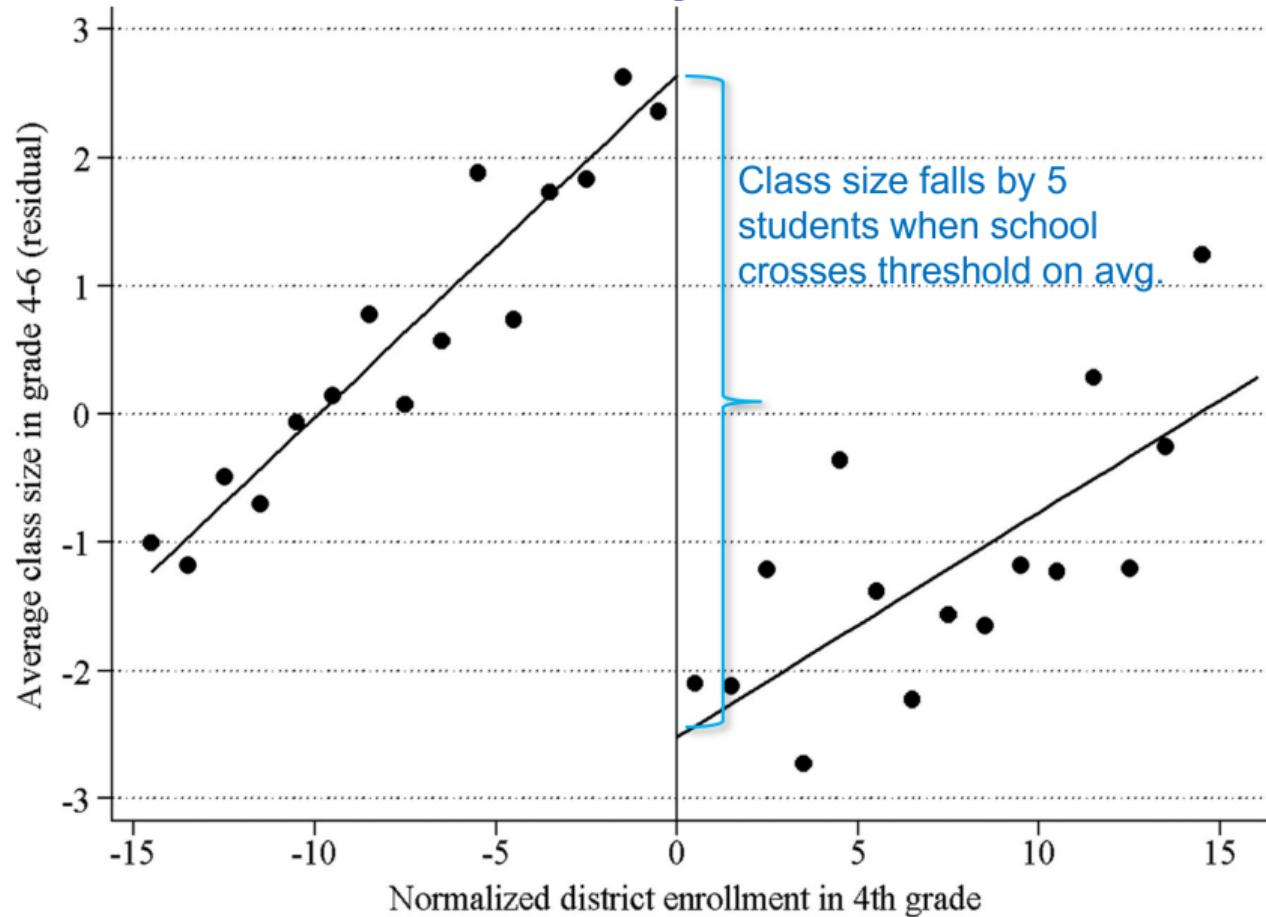
Discontinuities in Class Size Created by Maximum Class Size Rule



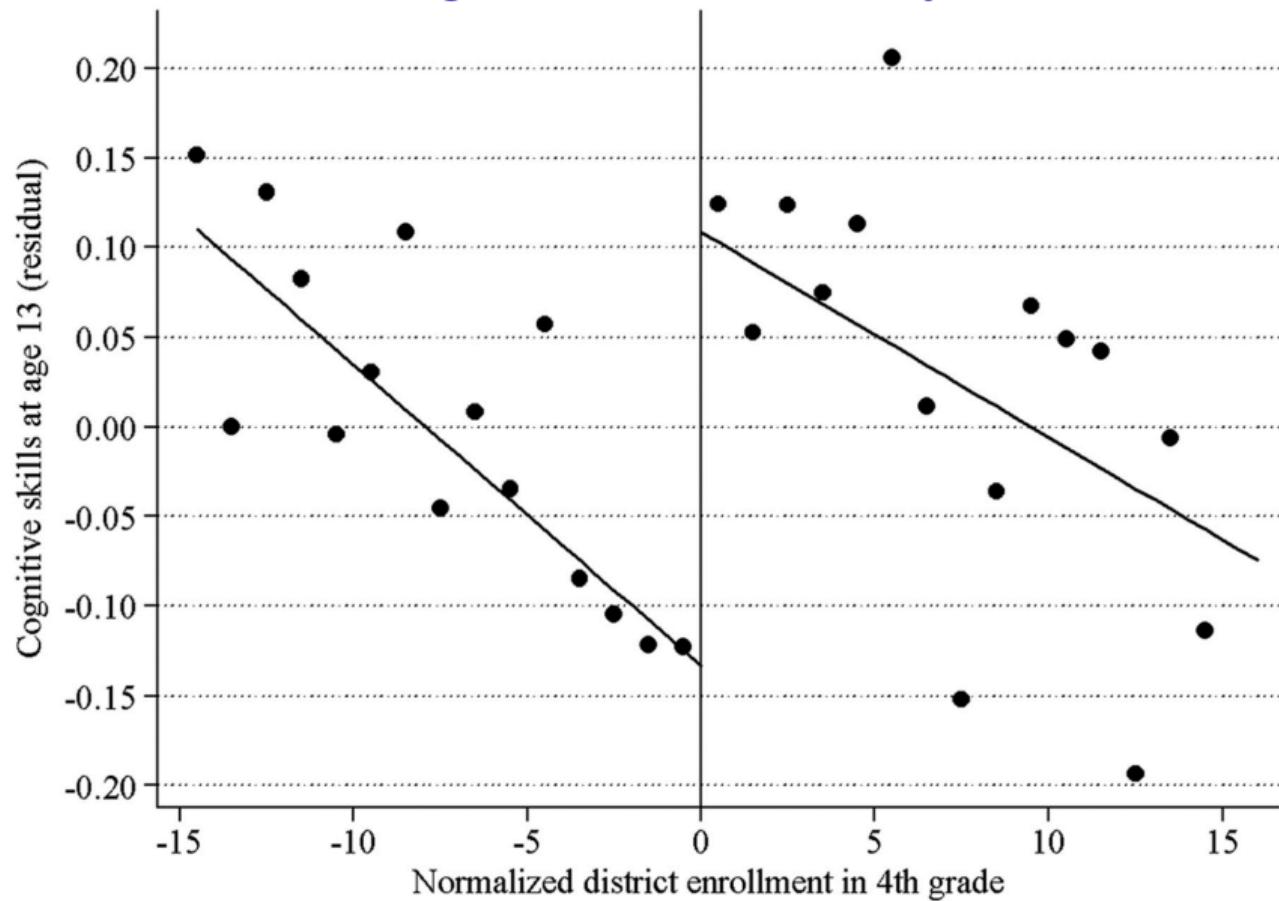
Discontinuities in Class Size Created by Maximum Class Size Rule



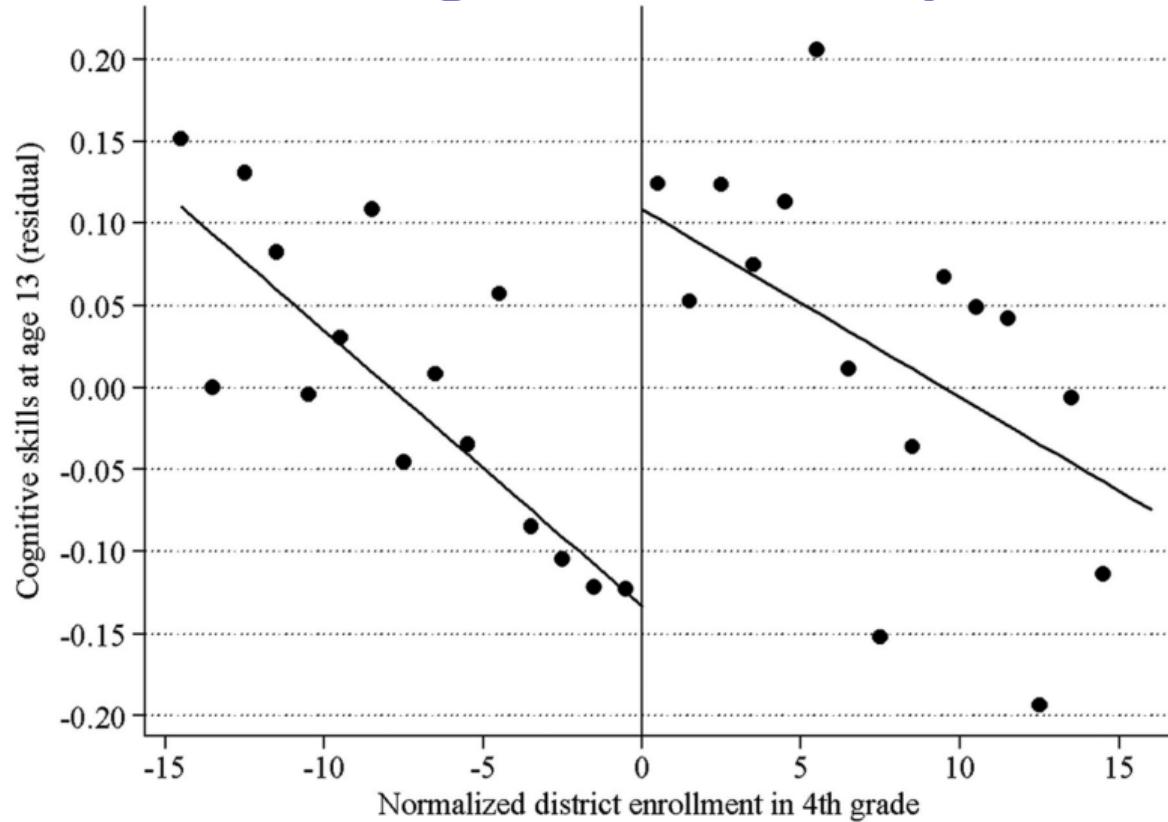
Discontinuities in Class Size Created by Maximum Class Size Rule



Test Score Achievement: Regression Discontinuity Estimates



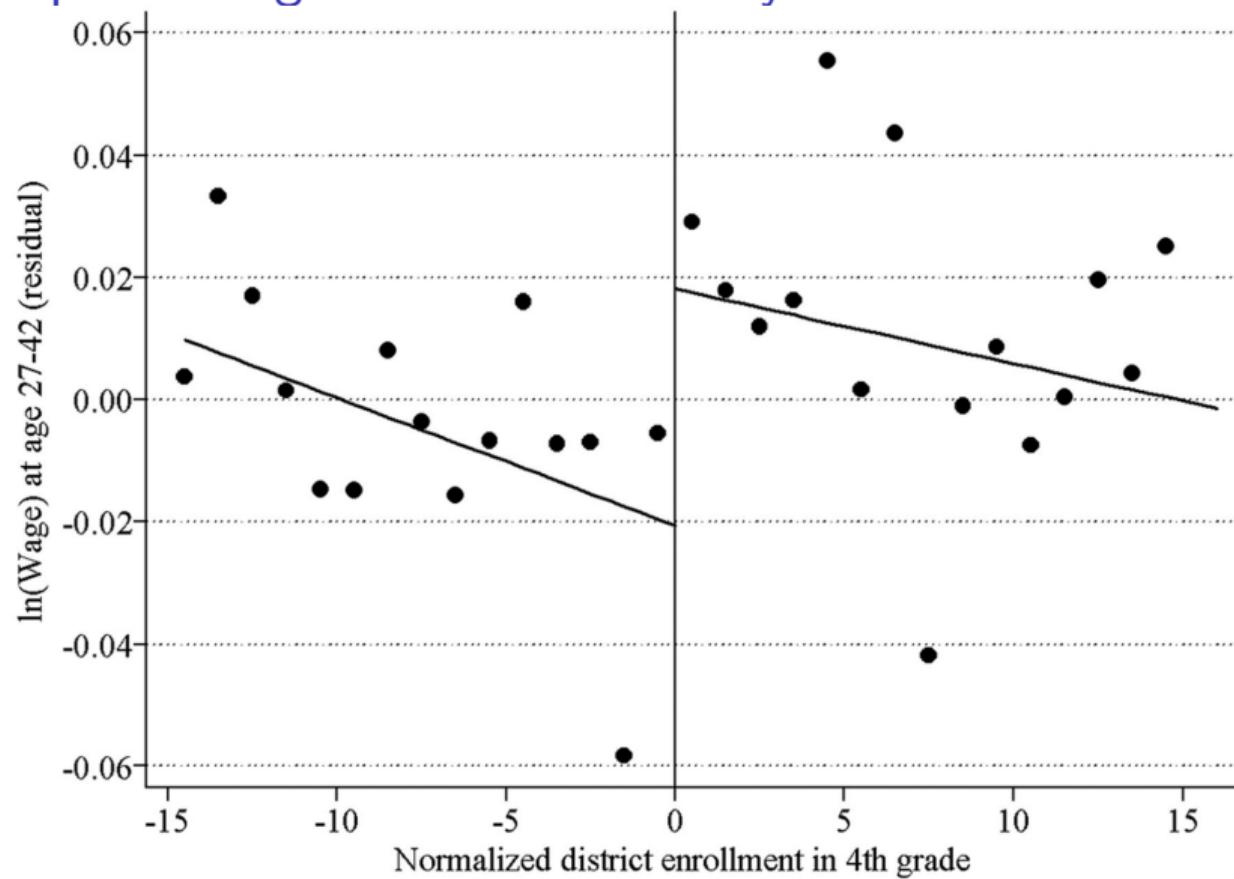
Test Score Achievement: Regression Discontinuity Estimates



Test scores jump by 0.2 standard deviations (8 percentiles) at cutoff

→ Reducing class size by 5 students causes 8 percentile increase in scores

Earnings Impacts: Regression Discontinuity Estimates



Earnings jump by 0.04 log points (4 percent) at cutoff

Lessons on Class Size

- ▶ Reducing class sizes in primary school by hiring more teachers can have large returns
 - ▶ Present value of lifetime earnings of a child growing up in a family at 25th percentile is about \$500,000 on average
 - ▶ 4% earnings gain from smaller class = \$20,000
 - ▶ Dividing a class of 30 students into two would increase total earnings of students by more than \$600,000
 - ▶ Costs (hiring another teacher and an additional room) likely to be well below \$600,000