Using Big Data to Solve Economic and Social Problems

Professor Raj Chetty
Head Section Leader: Gregory Bruich, Ph.D.
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

- References:


  Fredriksson, Ockert, Oosterbeek. “Long-Term Effects of Class Size.” QJE 2012

  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
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

“cup”
Earnings vs. Kindergarten Test Score

Note: $R^2 = 5\%$
Earnings vs. Kindergarten Test Score

Binned scatter plot: dots show average earnings for students in 5-percentile bins.

Ex: students scoring between 45-50 percentile earn about $17,000 on average.

Note: $R^2 = 5\%$
Earnings vs. Kindergarten Test Score

Note: $R^2 = 5$

But lot of variation in students’ earnings around the average in each bin.
Earnings vs. Kindergarten Test Score

Test scores explain only 5% of the variation in earnings across students.

Note: $R^2 = 5%$
Earnings vs. Kindergarten Test Score

Lesson: KG Test scores are highly predictive of earnings...but they don’t determine your fate

Note: $R^2 = 5\%$
College Attendance Rates vs. KG Test Score

- Attended College before Age 27
- Kindergarten Test Score Percentile

The graph shows a positive correlation between kindergarten test score percentile and college attendance rates before age 27, with higher test scores correlating with higher attendance rates.
Marriage by Age 27 vs. KG Test Score

Kindergarten Test Score Percentile

Married by Age 27

25% 30% 35% 40% 45% 50% 55%
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
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 Socioeconomic Status
US School Districts, 2009-2013
Academic Achievement and Socioeconomic Status
California and Massachusetts School Districts, 2009-2013
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

Corr. = 0.1498

Mean PISA in 2012

Spending per pupil in primary education (World Bank 2009)
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

- Begin by analyzing effects of class size

- Cannot simply compare outcomes across students who are in small vs. large classes
  - Students in schools with small classes will generally be from higher-income backgrounds and have other advantages
  - Therefore, this simple comparison in observational data will overstate causal effect of class size

- Need to use experimental/quasi-experimental methods instead
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

<table>
<thead>
<tr>
<th>Dep Var:</th>
<th>Test Score</th>
<th>College Attendance</th>
<th>Earnings</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Outcome</strong></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>Small Class</td>
<td>4.81</td>
<td>2.02%</td>
<td>-$4</td>
</tr>
<tr>
<td></td>
<td>(1.05)</td>
<td>(1.10%)</td>
<td>($327)</td>
</tr>
<tr>
<td>Observations</td>
<td>9,939</td>
<td>10,992</td>
<td>10,992</td>
</tr>
<tr>
<td>Mean of Dep. Var.</td>
<td>48.67</td>
<td>26.4%</td>
<td>$15,912</td>
</tr>
</tbody>
</table>
### STAR Experiment: Impacts of Class Size

<table>
<thead>
<tr>
<th>Dep Var:</th>
<th>Test Score</th>
<th>College Attendance</th>
<th>Earnings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>Small Class</td>
<td>4.81</td>
<td>2.02%</td>
<td>-$4</td>
</tr>
<tr>
<td></td>
<td>(1.05)</td>
<td>(1.10%)</td>
<td>($327)</td>
</tr>
</tbody>
</table>

**Estimated Impact**

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

<table>
<thead>
<tr>
<th>Dep Var:</th>
<th>Test Score</th>
<th>College Attendance</th>
<th>Earnings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>Small Class</td>
<td>4.81</td>
<td>2.02%</td>
<td>-$4</td>
</tr>
<tr>
<td></td>
<td>(1.05)</td>
<td>(1.10%)</td>
<td>($327)</td>
</tr>
<tr>
<td>Observations</td>
<td>9,939</td>
<td>10,992</td>
<td>10,992</td>
</tr>
<tr>
<td>Mean of Dep. Var.</td>
<td>48.67</td>
<td>26.4%</td>
<td>$15,912</td>
</tr>
</tbody>
</table>

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 will lie between 2.71 and 6.91 percentiles.
## STAR Experiment: Impacts of Class Size

<table>
<thead>
<tr>
<th>Dep Var:</th>
<th>Test Score</th>
<th>College Attendance</th>
<th>Earnings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small Class</td>
<td>4.81 <em>(1.05)</em></td>
<td>2.02% <em>(1.10%)</em></td>
<td>-$4 <em>(−$327)</em></td>
</tr>
<tr>
<td>Observations</td>
<td>9,939</td>
<td>10,992</td>
<td>10,992</td>
</tr>
</tbody>
</table>

**Mean of Dep. Var.**

- **Test Score**: 48.67
- **College Attendance**: 26.4%
- **Earnings**: $15,912
## STAR Experiment: Impacts of Class Size

<table>
<thead>
<tr>
<th>Dep Var:</th>
<th>Test Score</th>
<th>College Attendance</th>
<th>Earnings</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td></td>
</tr>
<tr>
<td>Small Class</td>
<td>4.81</td>
<td>2.02%</td>
<td>-$4</td>
</tr>
<tr>
<td></td>
<td>(1.05)</td>
<td>(1.10%)</td>
<td>($327)</td>
</tr>
<tr>
<td>Observations</td>
<td>9,939</td>
<td>10,992</td>
<td>10,992</td>
</tr>
<tr>
<td>Mean of Dep. Var.</td>
<td>48.67</td>
<td>26.4%</td>
<td>$15,912</td>
</tr>
</tbody>
</table>
STAR Experiment: Impacts of Class Size

<table>
<thead>
<tr>
<th>Dep Var:</th>
<th>Test Score</th>
<th>College Attendance</th>
<th>Earnings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>Small Class</td>
<td>4.81</td>
<td>2.02%</td>
<td>-$4</td>
</tr>
<tr>
<td></td>
<td>(1.05)</td>
<td>(1.10%)</td>
<td>($327)</td>
</tr>
<tr>
<td>Observations</td>
<td>9,939</td>
<td>10,992</td>
<td>10,992</td>
</tr>
</tbody>
</table>

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
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

School 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

![Graph showing discontinuities in class size against normalized district enrollment in 4th grade.](image-url)
Discontinuities in Class Size Created by Maximum Class Size Rule

Maximum class size cutoff (25 students)
Discontinuities in Class Size Created by Maximum Class Size Rule

Class size falls by 5 students when school crosses threshold on avg.
Test scores jump by 0.2 standard deviations (8 percentiles) at cutoff
→ Reducing class size by 5 students causes 8 percentile increase in scores

Test Score Achievement: Regression Discontinuity Estimates
Earnings Impacts: Regression Discontinuity Estimates

Earnings jump by 0.04 log points (4 percent) at cutoff
→ Reducing class size by 5 students causes 4% increase in earnings
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