Using Administrative Test-Score Data to Study Educational Opportunity in the US

Benjamin R. Shear

Center for Assessment, Design, Research & Evaluation University of Colorado Boulder School of Education

March 16, 2019 Keynote Address The 7th Annual Leadership Symposium St. John's University Queens, New York



What do you see?





And now?



And now?



And now?



What do you see?



Outline

What?

A new database for studying achievement and opportunity in US public school districts.

So what?

- Documenting patterns of educational opportunity.
- Studying variation across places.

Now what?

- What questions do these data raise?
- What steps might we take next?

What?

What are administrative test score data?

- Test score data stored after use that no longer serve instructional or accountability purposes.
- Common source: Accountability data.

What is educational opportunity?

- Schools attempt to help all students develop a range of knowledge, skills, and dispositions.
- Carefully designed standardized tests can measure **some** of these things.
- ▶ In the **aggregate**, test scores can serve as a proxy for educational opportunity.

Using Test Scores for System Monitoring (Fahle, Shear & Shores, forthcoming)

How can test scores be used to understand the education system?

- Inform stakeholders.
- Track changes and relative performance.
- Evaluate programs and policies.

What validity issues arise?

- What are intended interpretations and uses? Theory of action?
- What are potential unintended consequences?

New Developments

 Combining pre-existing administrative test score databases to study educational opportunity in the US.

The National Assessment of Educational Progress (NAEP)

- Reports results, by grade, subject, and demographic groups for:
 - The US overall.
 - Each state overall.
 - Small number of large urban districts (TUDA).
- Focus on mathematics and reading, but other subjects are also assessed.
- ▶ Focus on **4th** and **8th** grade, in **odd years** (2015, 2017, 2019).

How does opportunity vary at local levels?

► NAEP provides only limited information about local levels (schools, districts).

- Designed to track changes over time in a few grades at the **state** or **national** level.
- With **TUDA** we can compare a small number of **large urban** districts (only 1 year in some cases).
- States tests lack comparability.
 - States use different tests (some exceptions; PARCC, SBAC, or NECAP).
 - State tests change over time.
- Is there a way to learn more from the ~50 million tests that are given each year?

The Stanford Education Data Archive (SEDA)

Accountability achievement test data for all US public school districts in:

- 3rd-8th grade math and ELA.
- Reported separately for some student subgroups.
- Years 2008/09 2014/15.
- Based on approximately 300 million test scores during that span.
- On a **common national scale** based on the NAEP.
- Data Sources:
 - Test scores: EDFacts Proficiency Counts
 - District demographics: Census and Common Core of Data.
- Publicly available: seda.stanford.edu

- Sean Reardon, Demetra Kalogrides, Jenny Buontempo, Richard DiSalvo (Stanford University)
- Andrew Ho (Harvard University), Erin Fahle (St. John's University)
- Ross Santy, Michael Hawes, Marilyn Seastrom, Jennifer Davies (US Dept. of Education)
- Funding: IES, Spencer, WT Grant, Bill and Melinda Gates Foundation, Overdeck Family Foundation

So what?

Average 3rd Grade Achievement in US School Districts



Average 3rd Grade Achievement by District Socioeconomic Conditions



Average 3rd Grade Achievement by District Socioeconomic Conditions



- How much do students learn as they progress through school?
 - **Growth** in test scores across grades.
- > Have current students learned more than their same-age peers in the past?
 - Trends in test scores across years, but within grades.

Average Growth Across Grades by District Socioeconomic Conditions



From National to Local





Test Score Growth Among Chicago Public School Students, 2009-2014

AUTHORS	Executive Summary
Sean F. Reardon	A comparison of Chicago public school
Stanford University	those of public students across the U.S
Rebecca Hinze-Pifer Stanford University	scores improved dramatically more, on
	of the average student in the U.S. This i
	average Chicago student's test scores
	years from third to eighth grade. Secon
	Chicago students' scores improved mo
	students in the U.S. Test scores rose in
	2009 to 2014, compared to an increase
	equally true for black, Hispanic, and wh
	from increasingly test-aligned instruction

A comparison of Cheago public school student's transductive test scores in 2009-2014 with these of public students acrose the U.S. weaks test schrifting statume. First, Cheago students' scores improved dramatically more, on average, between third and eighth grade than those of the average student in the U.S. This is true for students of all incid-Vertime groups. The average facilitation that the score test use the proved by roughly grade-level squarkents in the grayears from third to eighth grades. Second: at each grade level in grades three through eight, students in the U.S. Test scores rose in Chicago by roughly two-thirds of a grade level from 2009 to 2014, compared to an increase of one-sixth of a grade level atomally Again. This was from increasingly lest-sligned instruction or from changing city demographics and envolution to result from increasingly lest-sligned instruction or from changing city demographics and envolutioned testers.

- Chicago students' scores increased about 1 additional grade from 3rd-8th grade relative to national average.
- Within grades, scores went up across years.

Source: Reardon & Hinze-Pifer (2017)

From Local...

STATE TEST RESULTS

With accelerated growth in literacy and math, Denver students close in on state averages

BY MELANIE ASMAR - AUGUST 16, 2018



Angel Trigueros-Martinez pokes his head from the back of the line as students wait to enter the building on the first day of school at McGione Academy on Wednesday. (Photo by Arcon Ontiverog/The Denver Post)

... to Nationa



... to Nationa



... to National



Socioeconomic Conditions

Understanding Relative Performance Between Boys and Girls



- Girls consistently outperform boys on ELA/Reading tests across districts.
- Boys outperform girls in more affluent districts, but not in less affluent districts.
- What causes these differences?

School Regulations of Boys' Rule-Breaking and the Gendered Social Construction of Exceptionalism in Early Adolescence

Michela Musto

FORTHCOMING AT THE AMERICAN SOCIOLOGICAL REVIEW

Abstract: from indegates through collega, sudests parevise loop as none statilizers than gait, yet free sociological adels how identified how school processes being indegates and this ballels. Therewise of 25 years of loopical devices how identified a racially device, public indefa exchool in Los Angeles, this orticle demonstrates how electronic differential regulation of boys' rule-bracking, course lead combined to gender-based differences in studenty proception of intelligence. This high lend accurace where differential is addresses to the student and anonpolated calcurates courses and anon provide the student base based boys' rule-bracking, such that boys, challenged gait's apprinted and anonpolated calcurates courses on the public gaits address the proceed based boys frage branching. The branching theorem, in lower-level causes - where non-offlaters, finds substates were coverspresented – electorem parallel substates (boys' rulebased). The state of the state based based

Source: Musto (forthcoming)

- Detailed ethnography of 1 LA middle school.
- Different patterns of behavior and discourse encouraged by educators based on students' gender, race, and course level.
- Partly consistent with larger systemic patterns seen in SEDA.

Now what?

- What can we learn from districts that are "beating the odds"?
- Why do districts with similar socio-economic conditions have such different levels on achievement tests?
- What leads to higher growth rates for some less affluent communities?
- What happens in schools that contribute to inequalities?
- What steps can schools take to improve educational opportunities for all students?

Limitations, Strengths, and Uses

Limitations

- Questions that SEDA can/cannot answer.
- Test scores are at best a partial indicator of educational opportunity.

Strengths

- Data are comprehensive in scope and scale.
- > Data are publicly available, and will be expanded.

Uses

- Describe broad patterns of educational opportunity.
- Identify patterns and places for further study.
- Provide national context for local phenomena or outcomes.

- "RPPs are long-term collaborations between practitioners and researchers that are organized to investigate problems of practice and solutions for improving schools and school districts" (Coburn & Penuel, 2016).
- Availability of data sources such as SEDA might:
 - Generate questions that can be investigated collaboratively between district or school partners and university researchers.
 - Be used as a source of outcome data in on-going partnership research.
- EdD. leaders may be ideally positioned to engage in this work.

- How can you use SEDA?
- What questions do these data or results raise?
- New website and tools.
- Feedback for us?

Contact: benjamin.shear@colorado.edu



- Coburn, C., Penuel, W., and Geil, K. (2013). *Research-practice partnerships: A strategy for leveraging research for educational improvements in school districts*. William T. Grant Foundation, New York, NY.
- Coburn, C. E. and Penuel, W. R. (2016). Research-practice partnerships in education: Outcomes, dynamics, and open questions. *Educational Researcher*, 45(1):48–54.
- Fahle, E. M., Shear, B. R., Kalogrides, D., Reardon, S. F., DiSalvo, R., and Ho, A. D. (2018). Stanford Education Data Archive: Technical documentation (Version 2.1).
- Fahle, E. M., Shear, B. R., and Shores, K. A. (2019). Assessment for monitoring of education systems: The US example. *The ANNALS of the American Academy of Political and Social Science*.
- Reardon, S. F. (2018). Educational opportunity in early and middle childhood: Variation by place and age. CEPA Working Paper No. 17-12, Stanford Center for Education Policy Analysis.

- Reardon, S. F. and Hinze-Pifer, R. (2017). Test score growth among chicago public school students, 2009-2014. Technical report, Stanford Center for Education Policy Analysis.
- Reardon, S. F., Ho, A. D., Shear, B. R., Fahle, E. M., Kalogrides, D., and DiSalvo, R. (2018a). Stanford Education Data Archive.
- Reardon, S. F., Kalogrides, D., Fahle, E. M., Podolsky, A., and Zárate, R. C. (2018b). The relationship between test item format and gender achievement gaps on math and ELA tests in fourth and eighth grades. *Educational Researcher*, pages 1–11.