
How Does the AIM Program Affect Student Outcomes in the Davis Joint Unified School District?

Scott Carrell, Ph.D.

Michal Kurlaender, Ed.D.

Marianne Page, Ph.D.

K. A. Kramer, M.A.

School of Education & Department of Economics

University of California, Davis

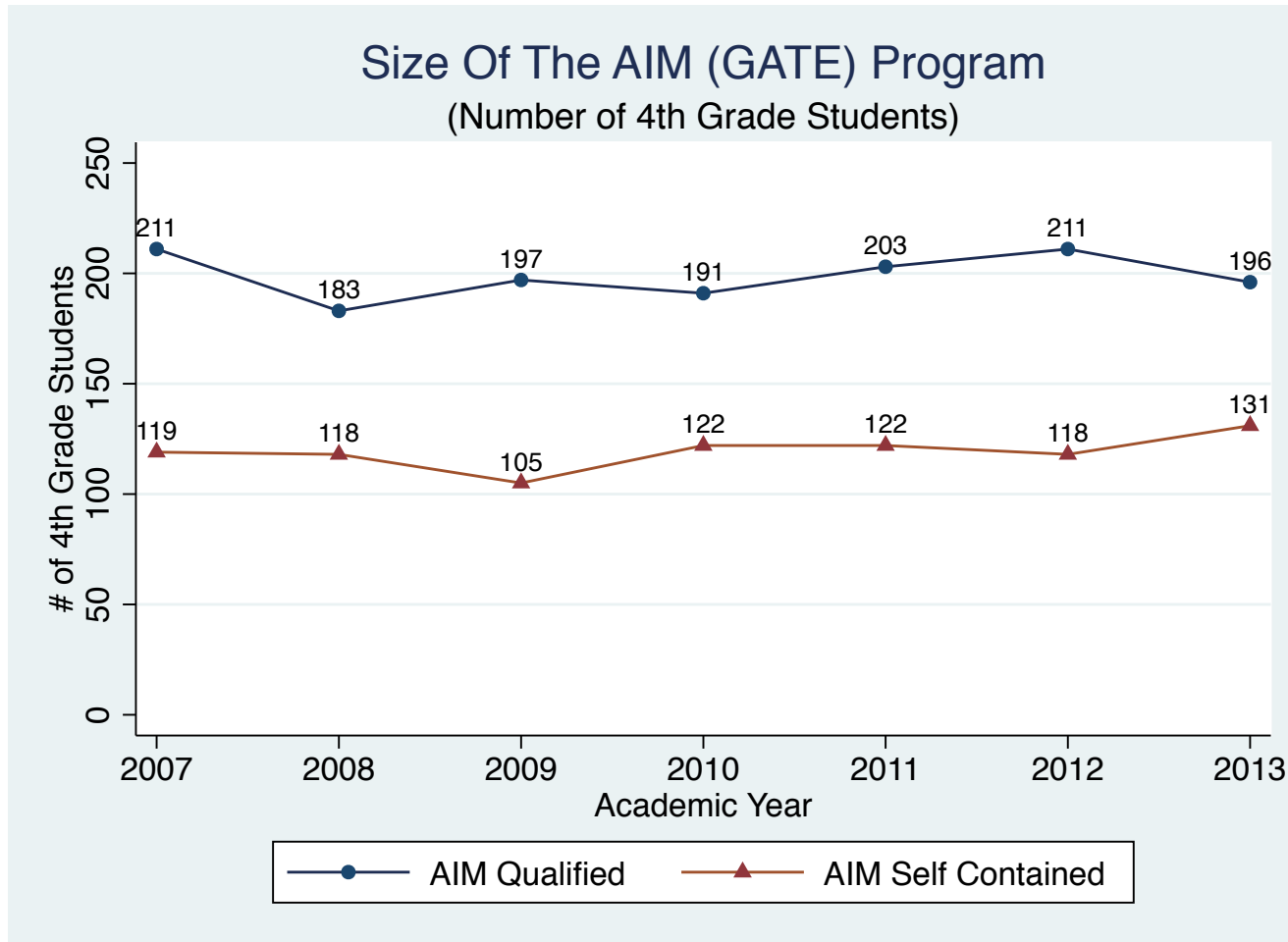
Research goals:

- Who is GATE / AIM identified in the DJUSD
 - How has this changed over time?
- What is the *causal* impact of AIM on student academic achievement for students in AIM self contained classrooms?
- What is the *causal* impact of AIM on student academic achievement for students **not** in AIM self contained classrooms?

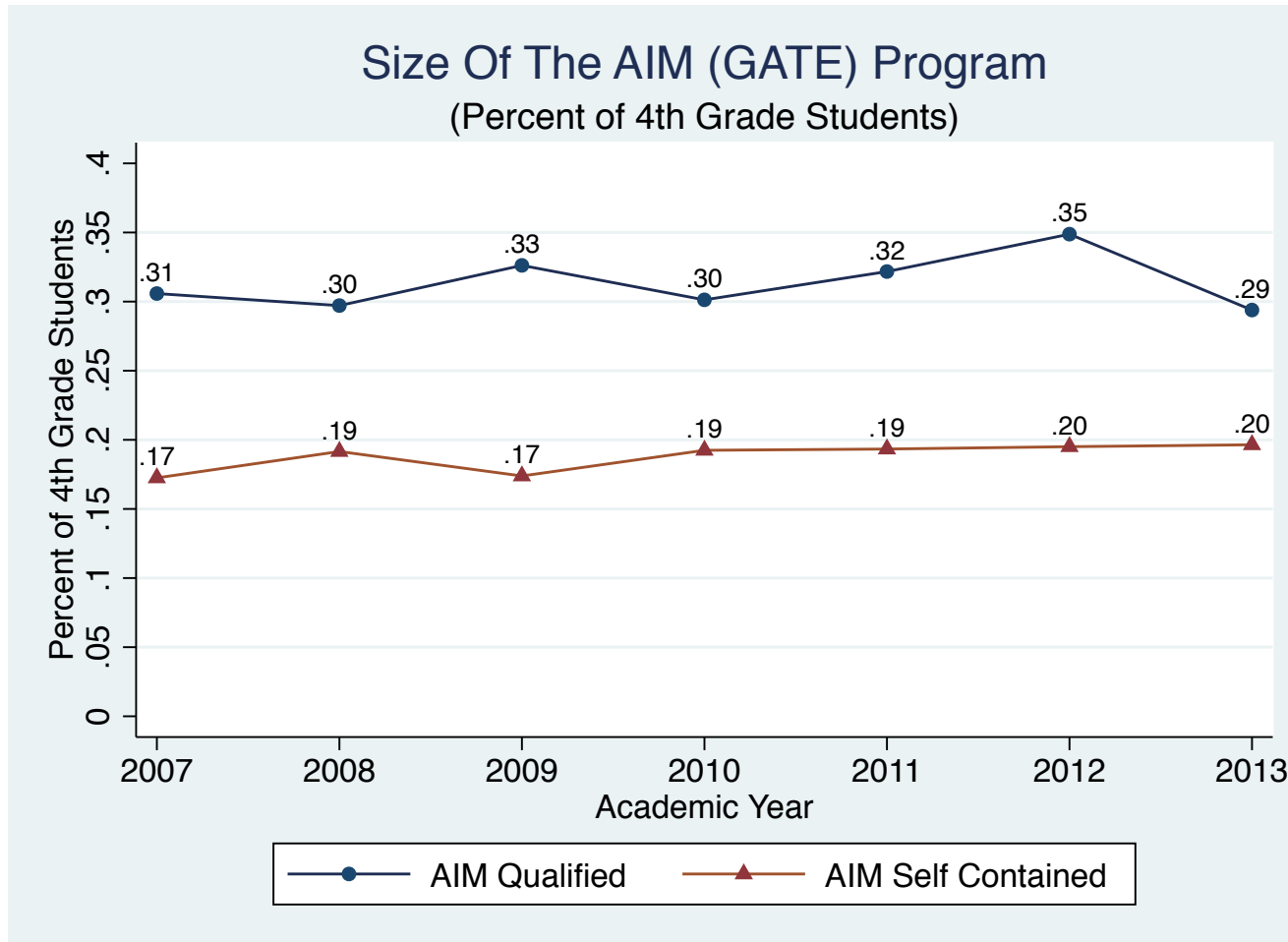
What is AIM?

- AIM: Alternative Instructional Model (formerly GATE).
- *It is the mission of the DJUSD GATE Program...to provide a quality educational program for gifted and talented students in order to develop their knowledge, skills, abilities, and values (2008 Master Plan).*
- Students who qualify for AIM are eligible to receive differentiated instruction in self contained classrooms.

How large is AIM?



How large is AIM?



How do students qualify for AIM?

There are three main channels:

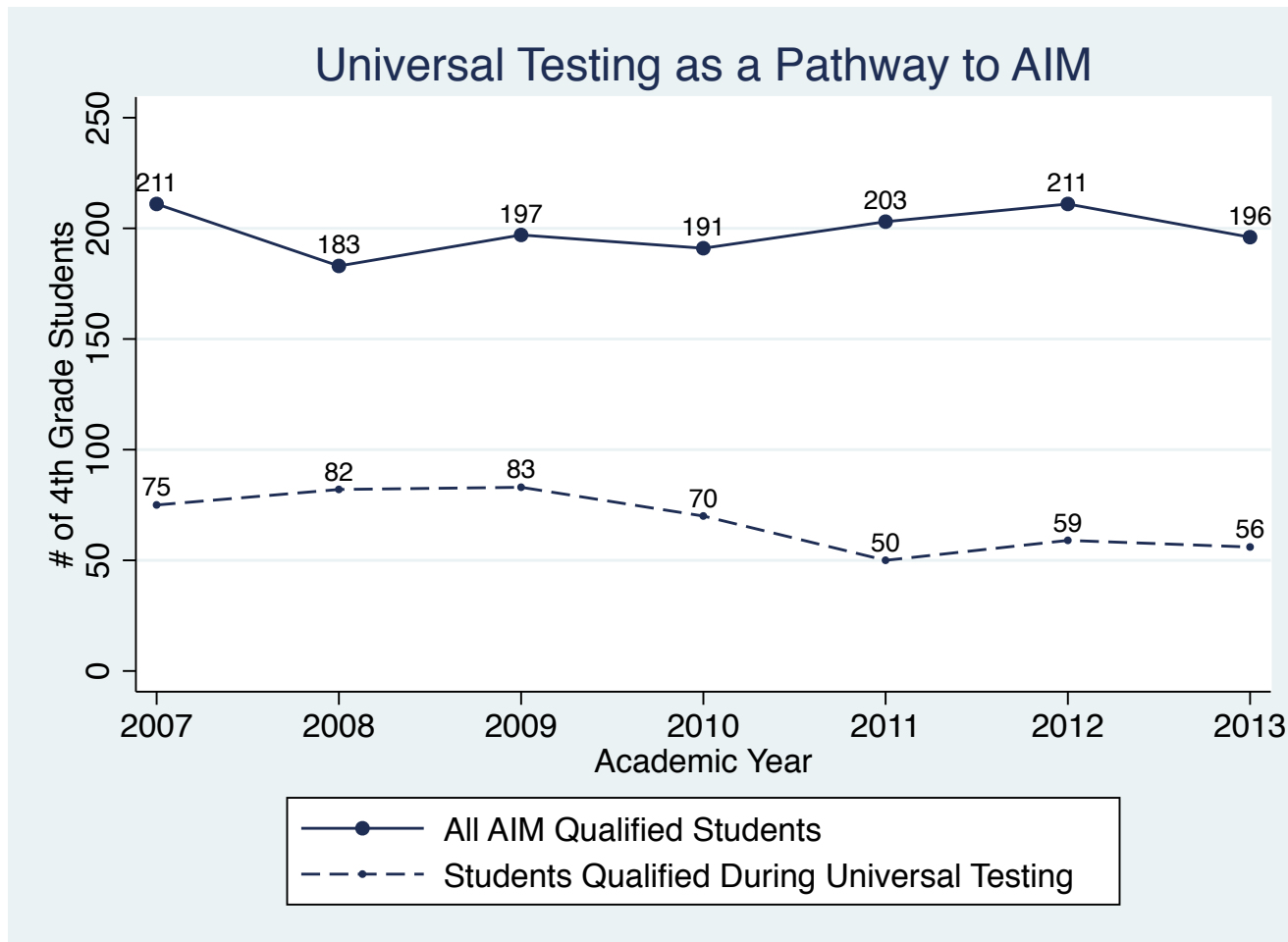
- 1. Universal Testing:** If a student scores in the 96th percentile on their total score *AND* on either their verbal or nonverbal score on the OLSAT the student qualifies for AIM.
- 2. Retesting:** If a student scores within +/- 5 standard errors of measurement of a 96th percentile on the OLSAT, the student automatically qualifies for retesting by the DJUSD. Students can also qualify for retesting through the search and serve process.

How do students qualify for AIM?

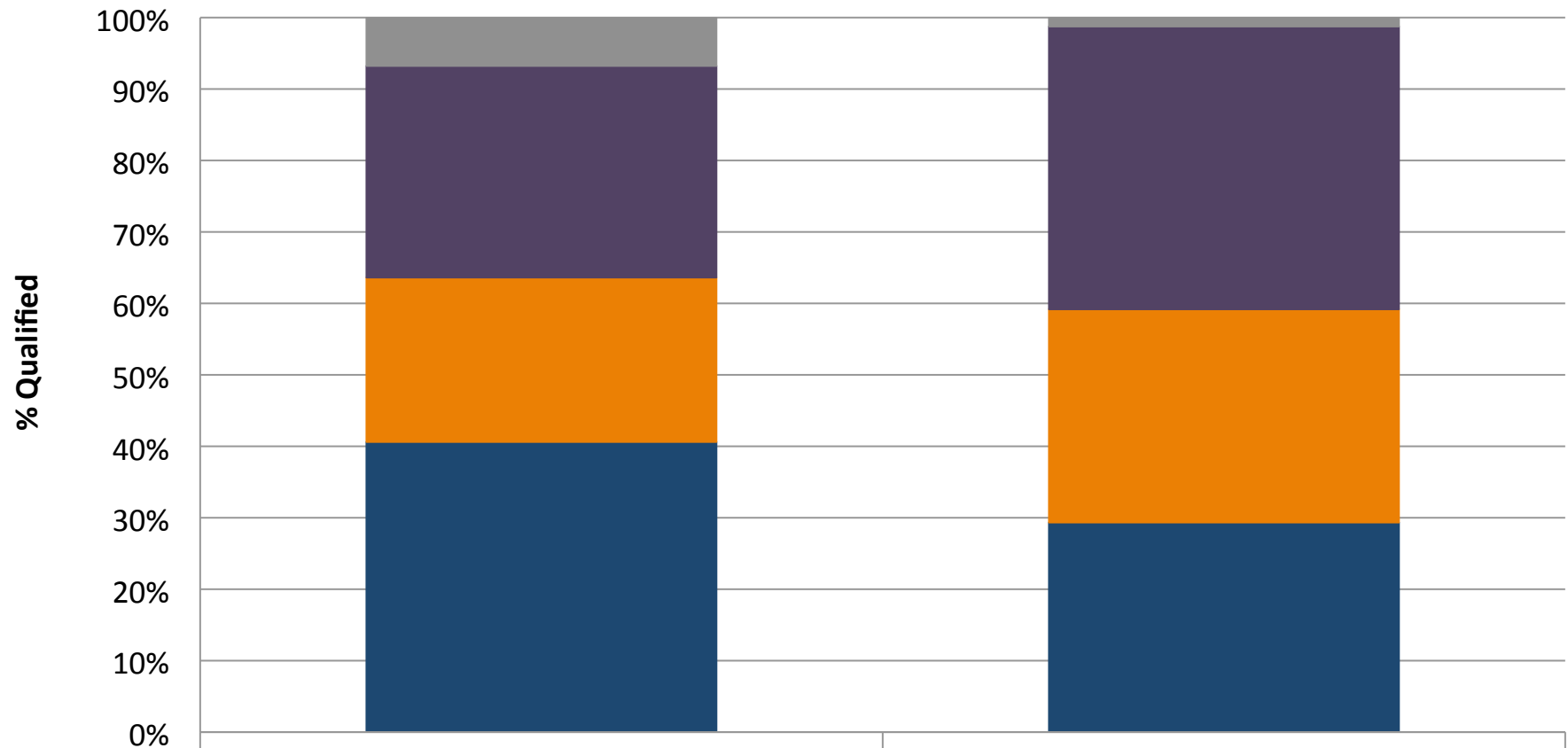
There are three main channels (cont.):

- 3. Private Testing:** A student can also qualify by taking a test of mental reasoning administered by a licensed psychologist.

How many students qualify for AIM during universal testing?

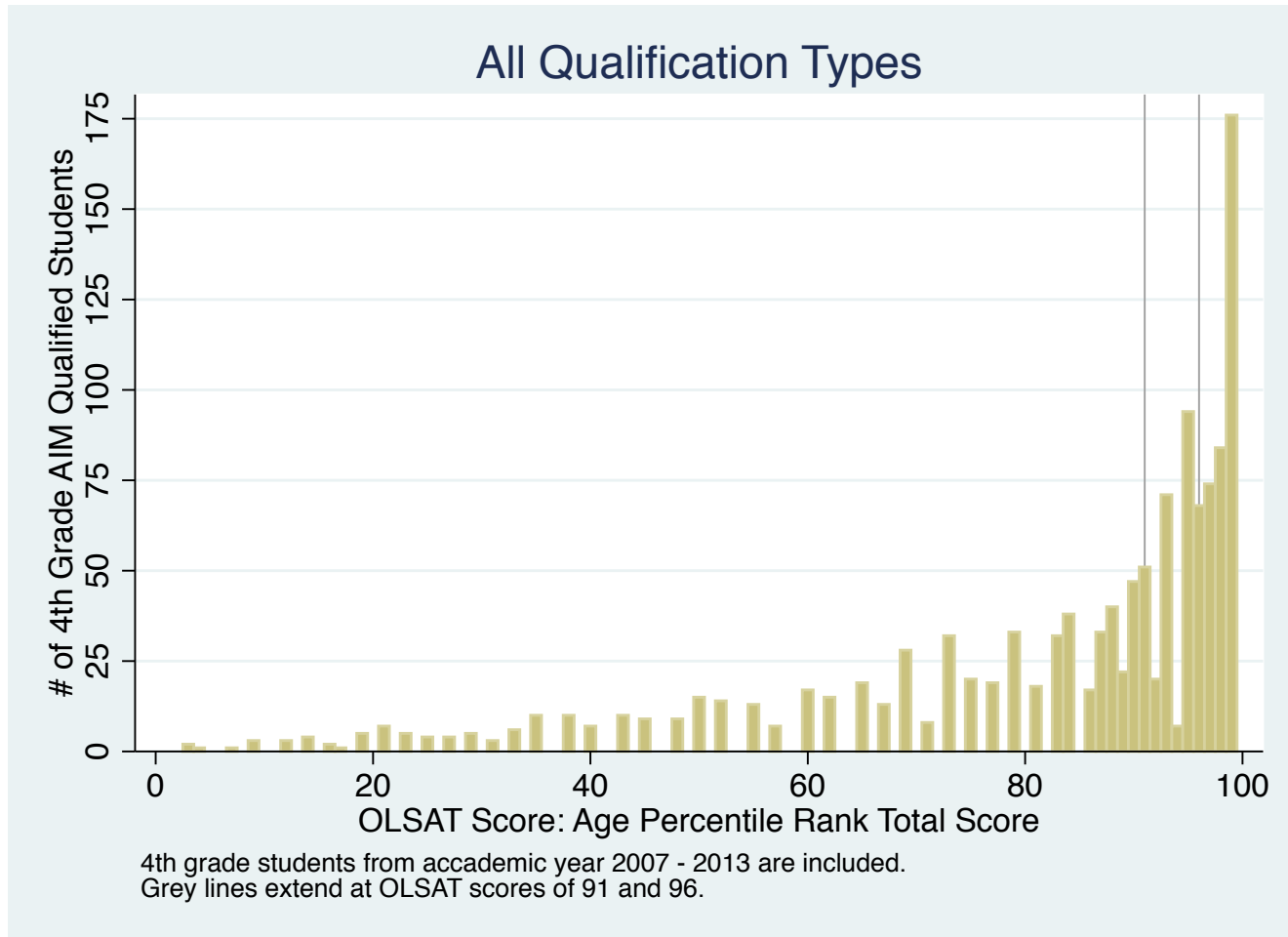


AIM (GATE) Qualification Type by Master Plan

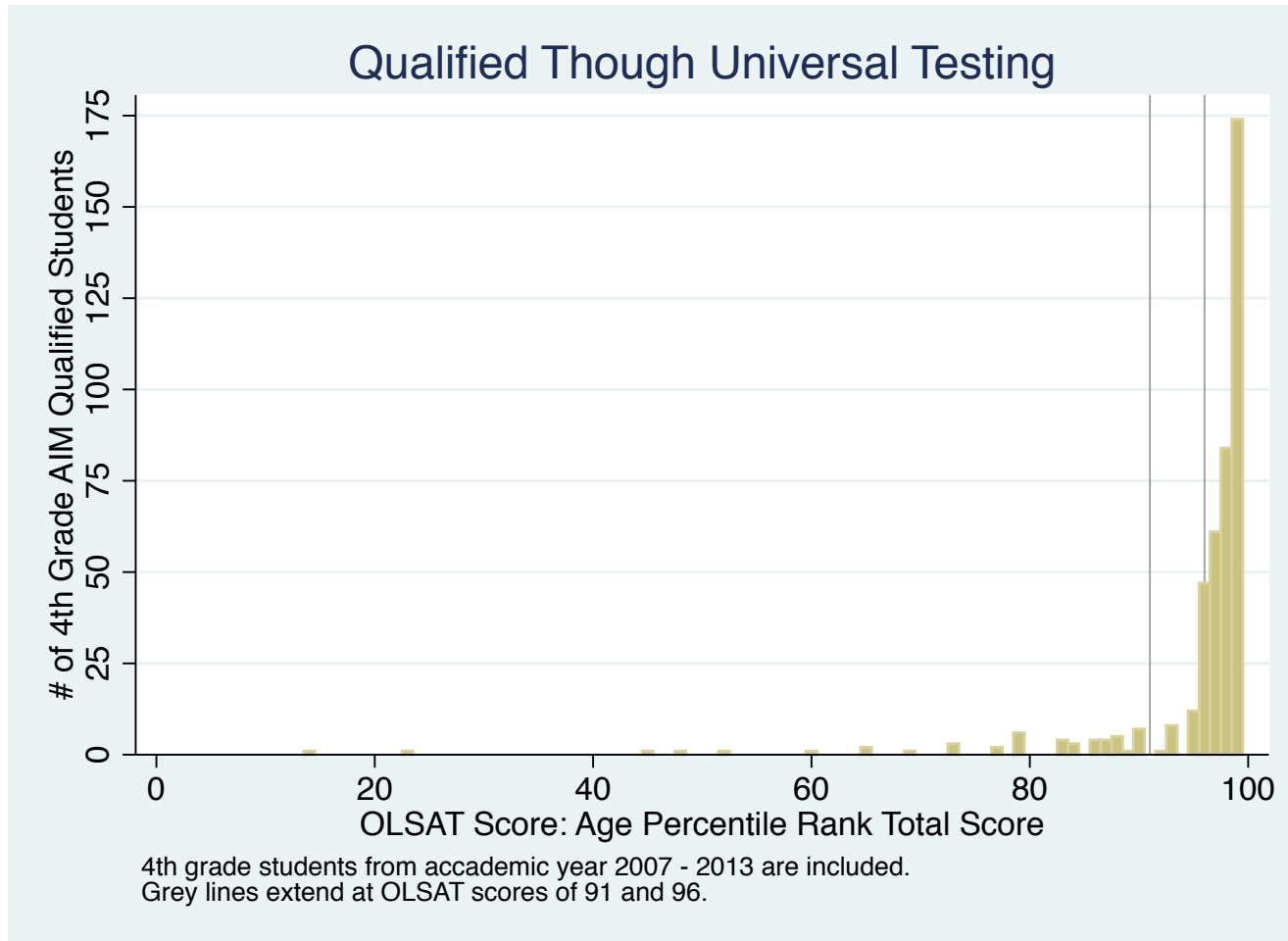


	Master Plan 2005	Master Plan 2008
■ Transfers / Other	6.8%	1.2%
■ Retesting	29.6%	39.6%
■ Private Testing	23.0%	29.8%
■ Universal Testing	40.6%	29.3%

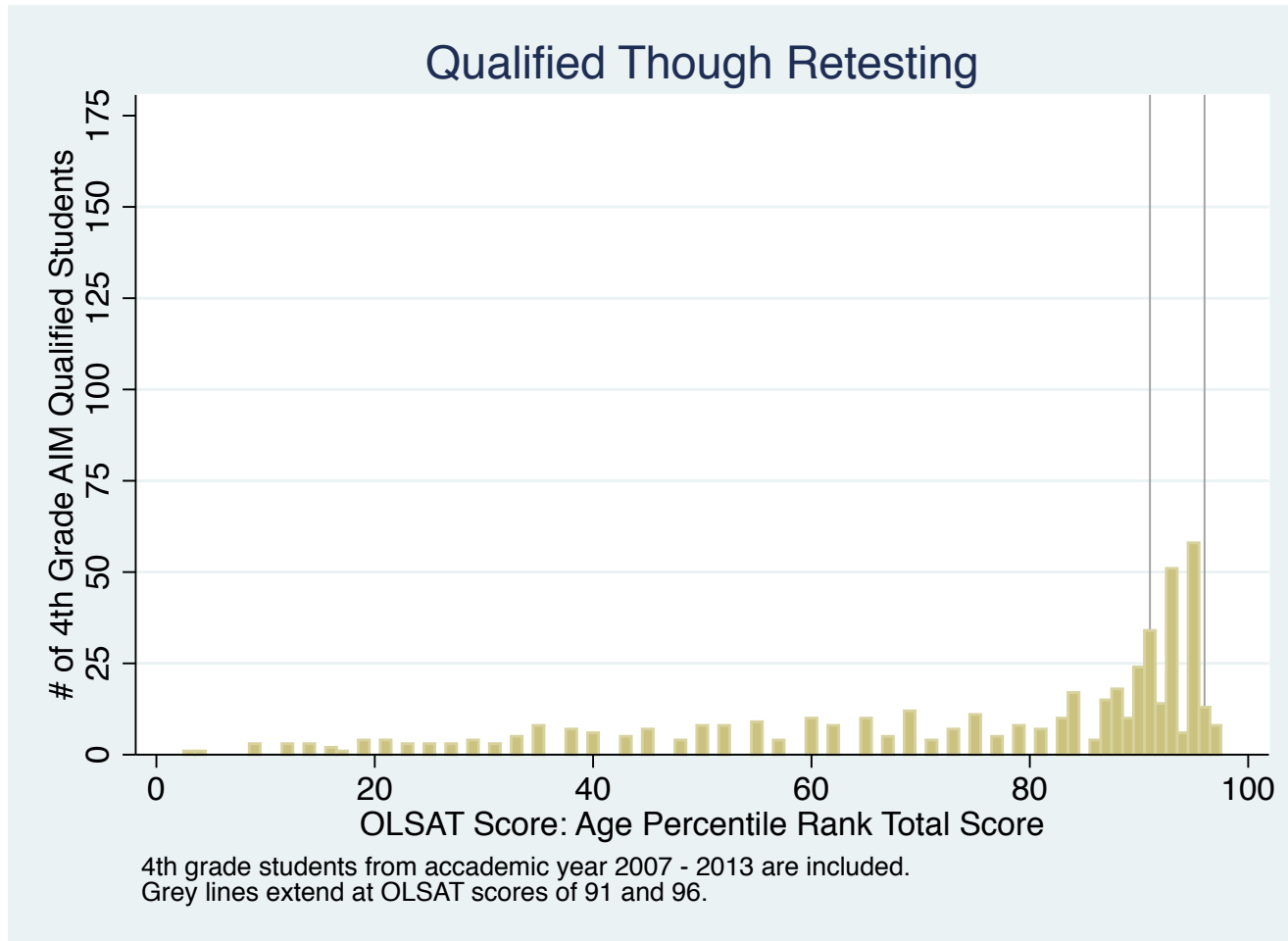
OLSAT scores of all students who qualify for AIM:



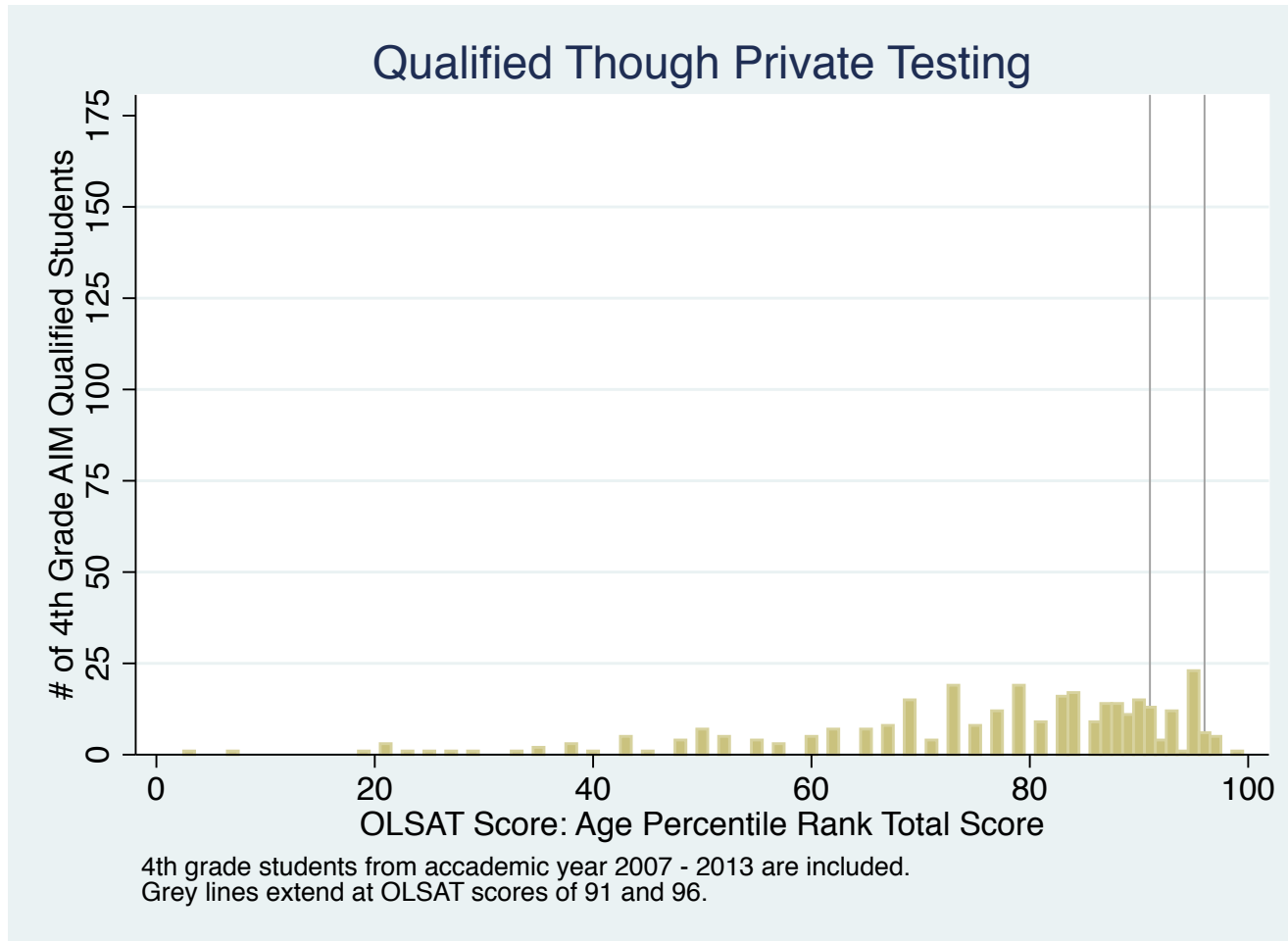
OLSAT scores of students who qualify for AIM through universal testing:



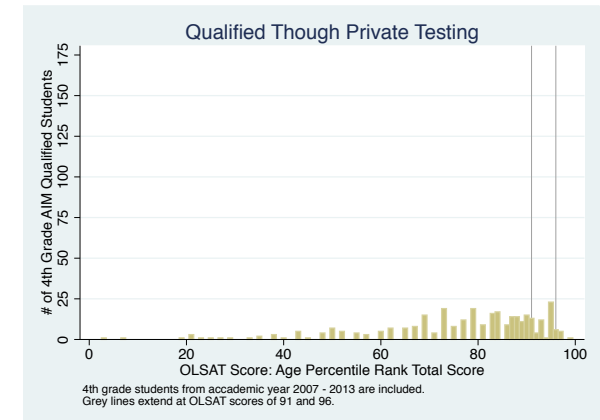
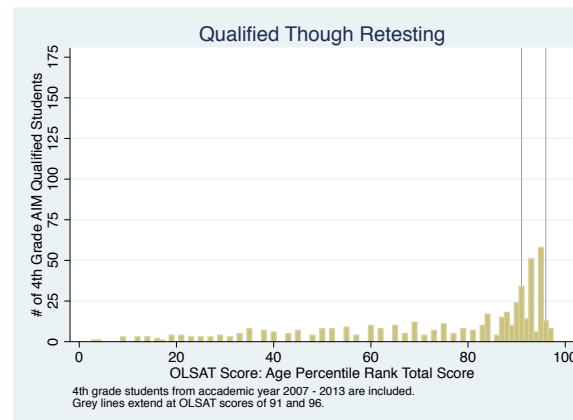
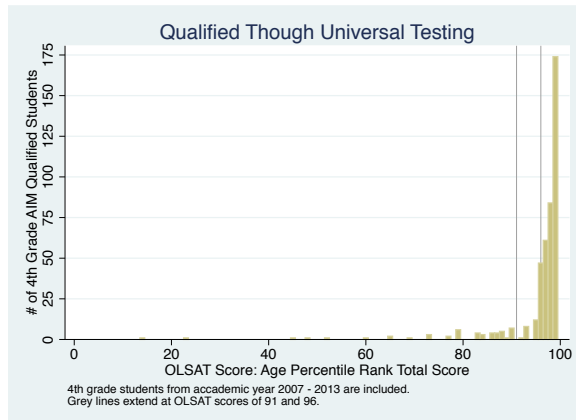
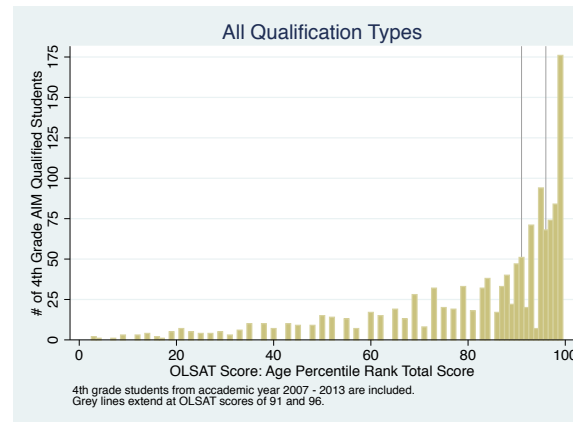
OLSAT scores of students who qualify for AIM though retesting:



OLSAT scores of students who qualify for AIM through private testing:



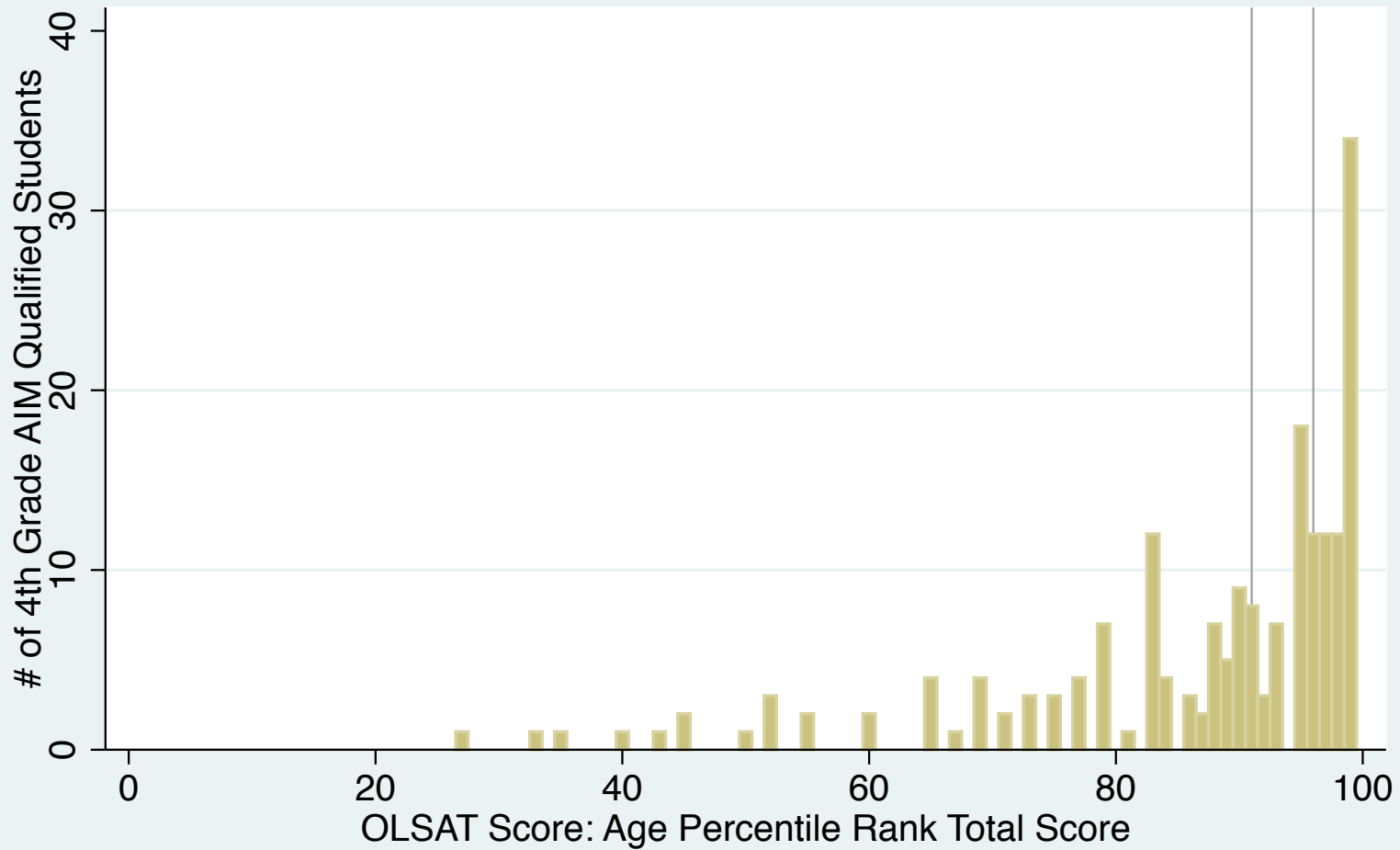
OLSAT scores by qualification type:



Mean OLSAT Score by Qualification Type

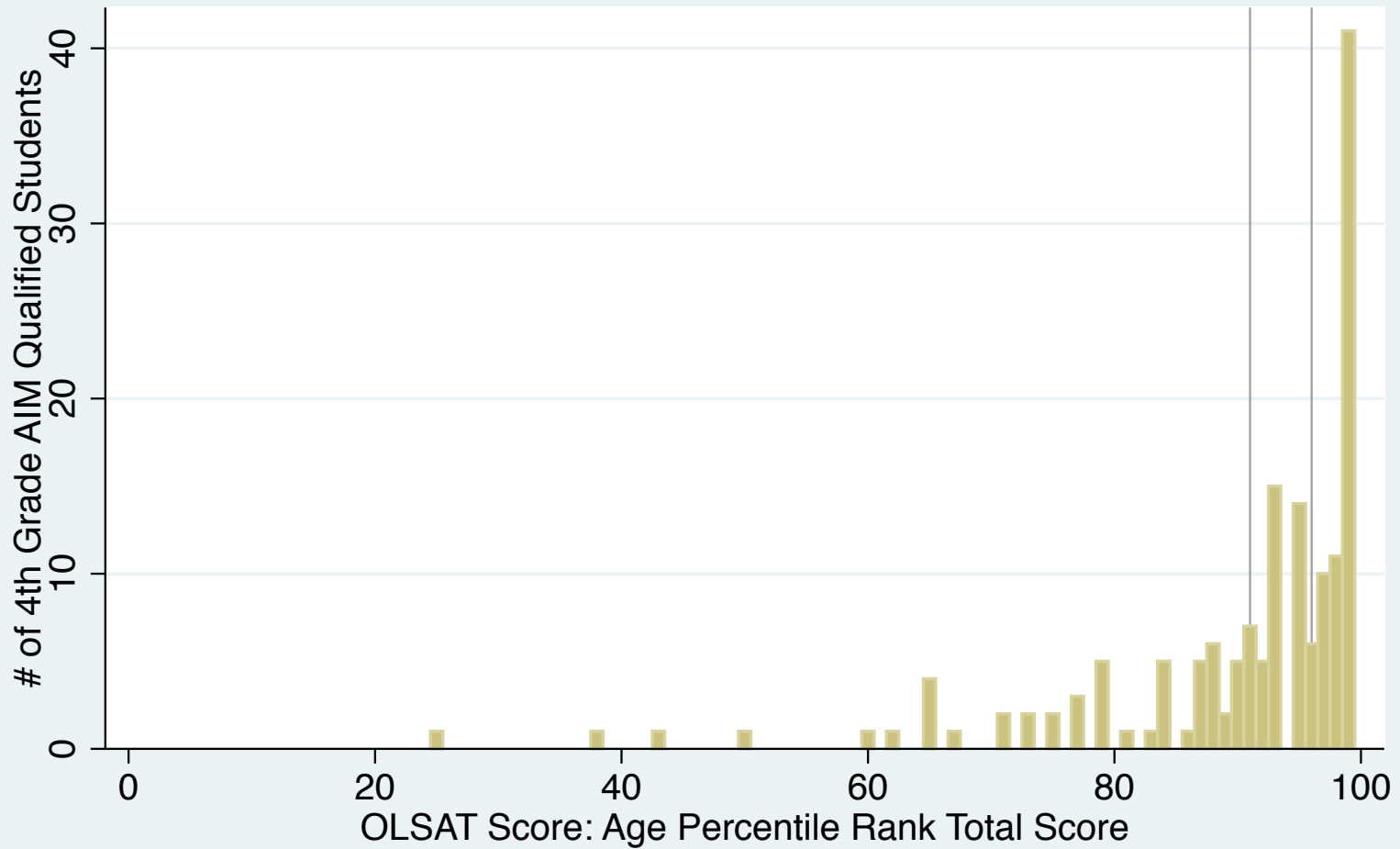
All Types	Universal Testing	Retesting	Private Testing
83	96	75	77

OLSAT Scores: 4th Grade AIM Identified Students, 2007



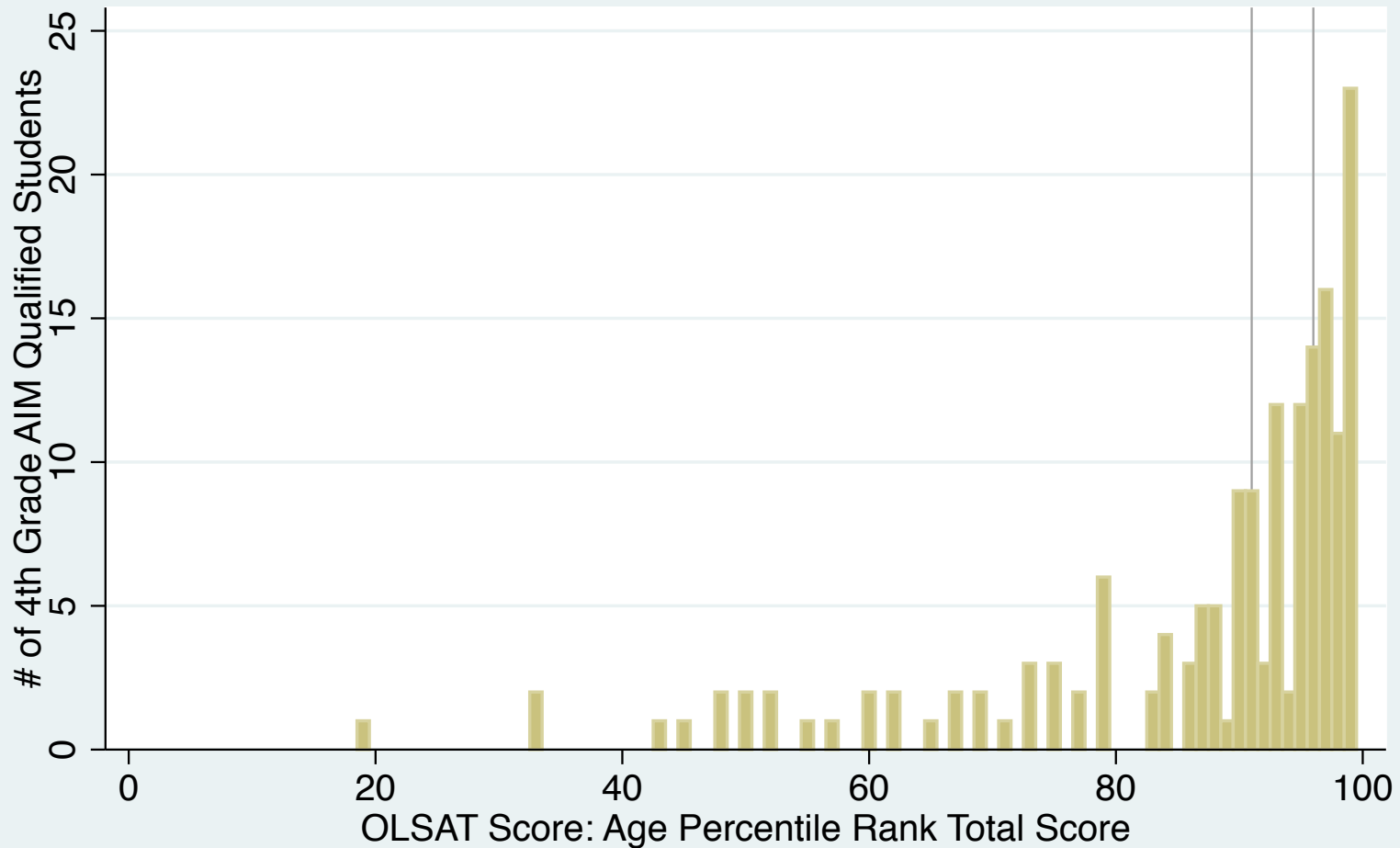
There were 211 4th grade AIM (GATE) Identified students in 2007.
Grey lines extend at OLSAT scores of 91 and 96.

OLSAT Scores: 4th Grade AIM Identified Students, 2008



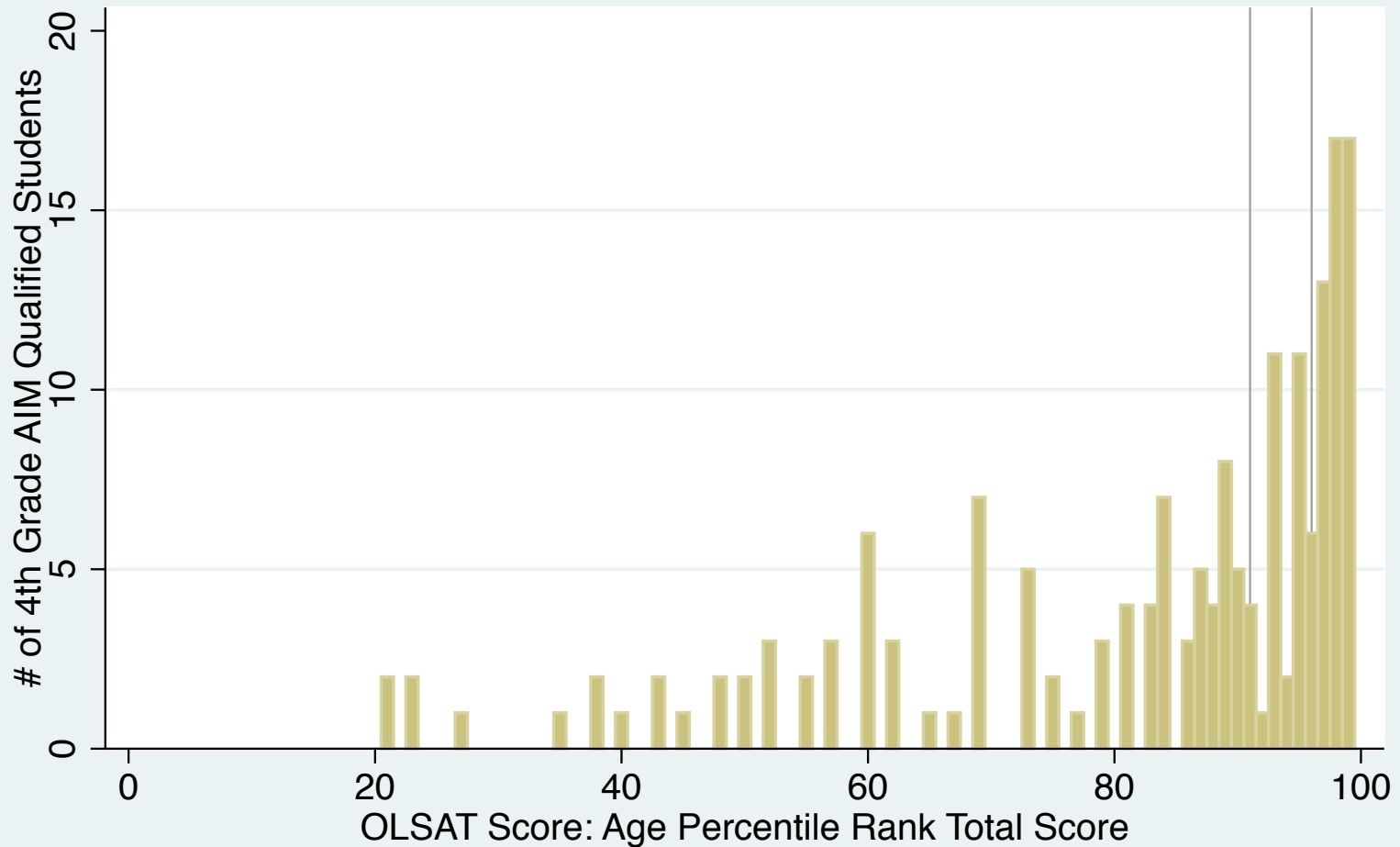
There were 183 4th grade AIM (GATE) Identified students in 2008.
Grey lines extend at OLSAT scores of 91 and 96.

OLSAT Scores: 4th Grade AIM Identified Students, 2009



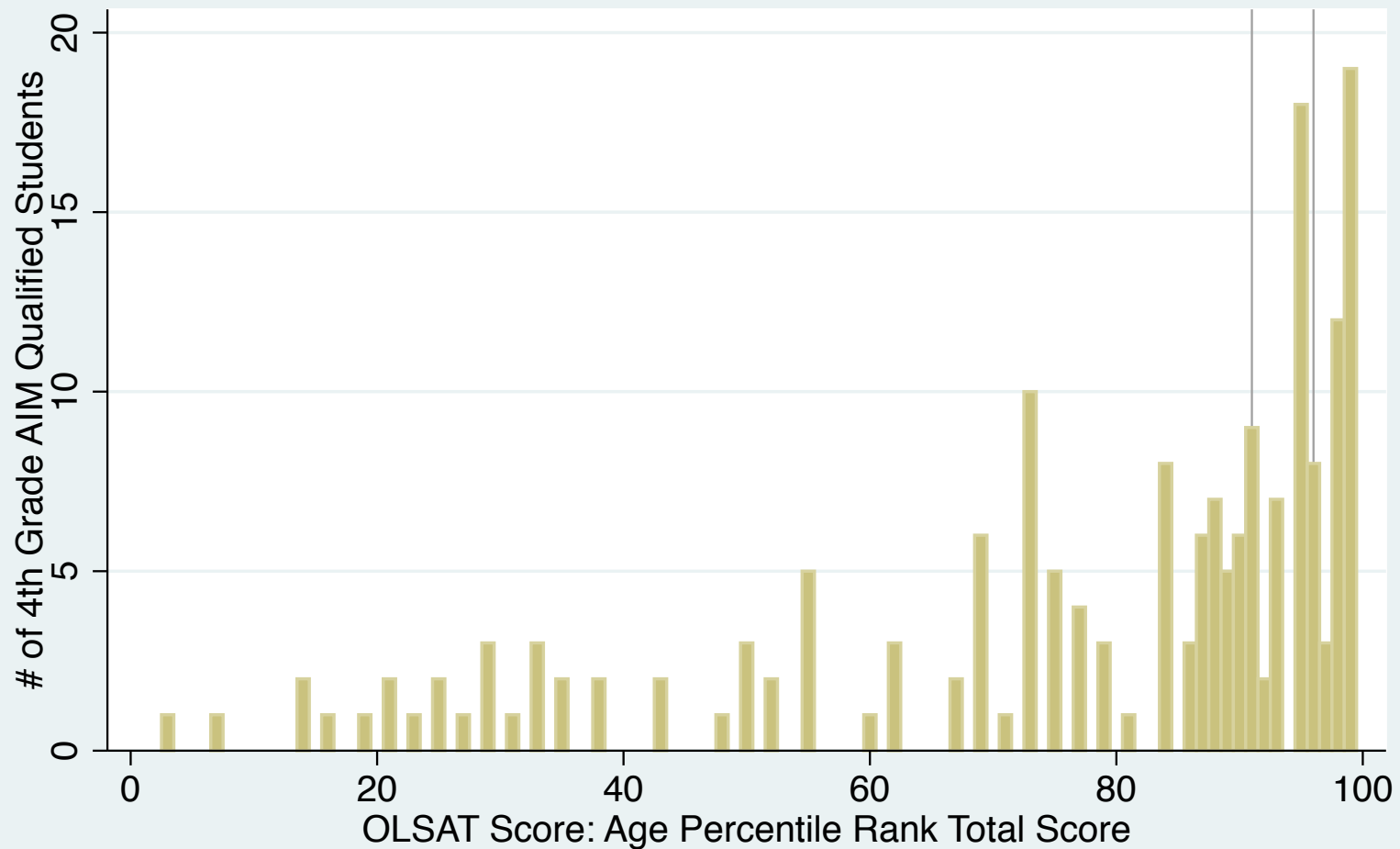
There were 197 4th grade AIM (GATE) Identified students in 2009.
Grey lines extend at OLSAT scores of 91 and 96.

OLSAT Scores: 4th Grade AIM Identified Students, 2010



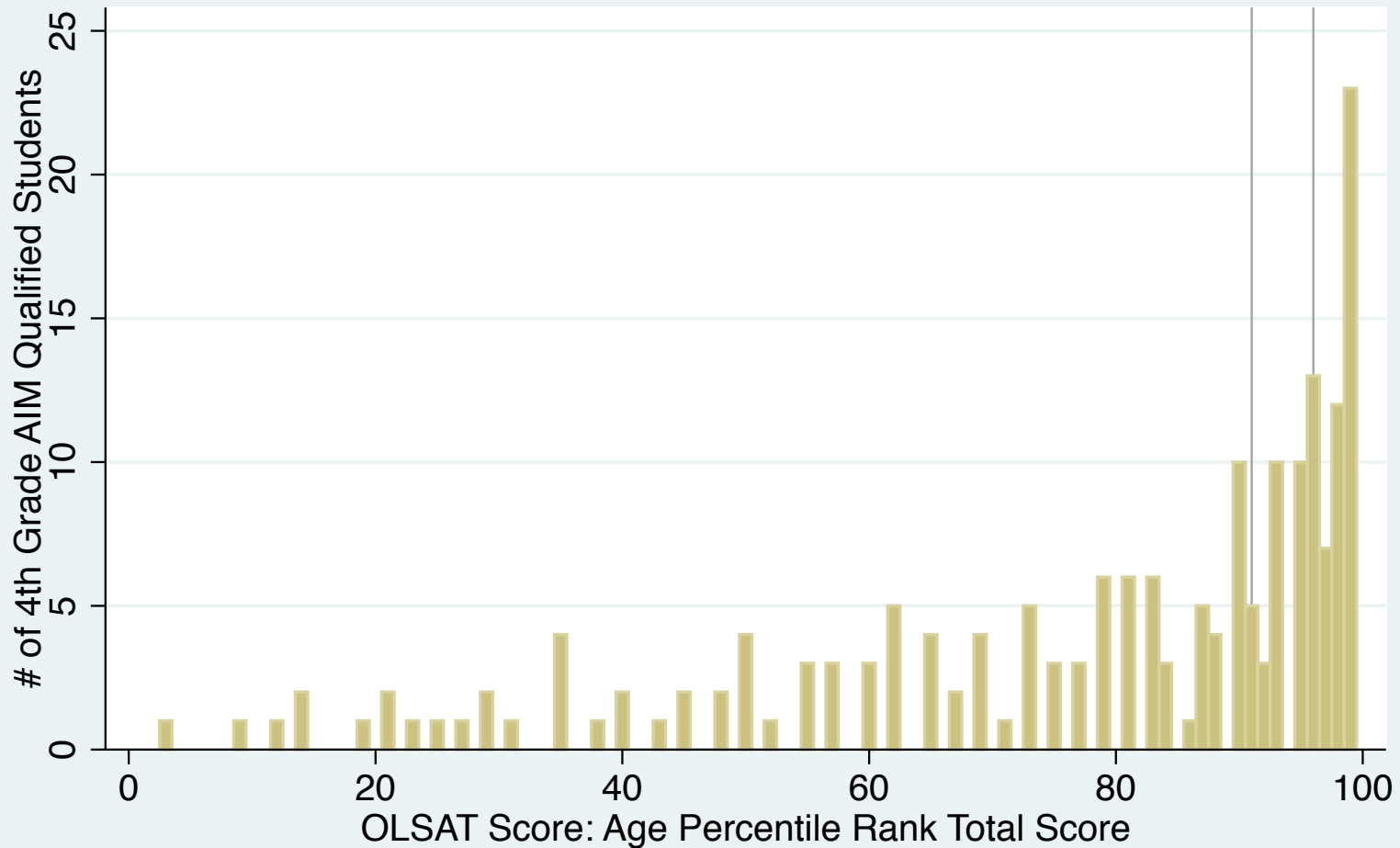
There were 191 4th grade AIM (GATE) Identified students in 2010.
Grey lines extend at OLSAT scores of 91 and 96.

OLSAT Scores: 4th Grade AIM Identified Students, 2011



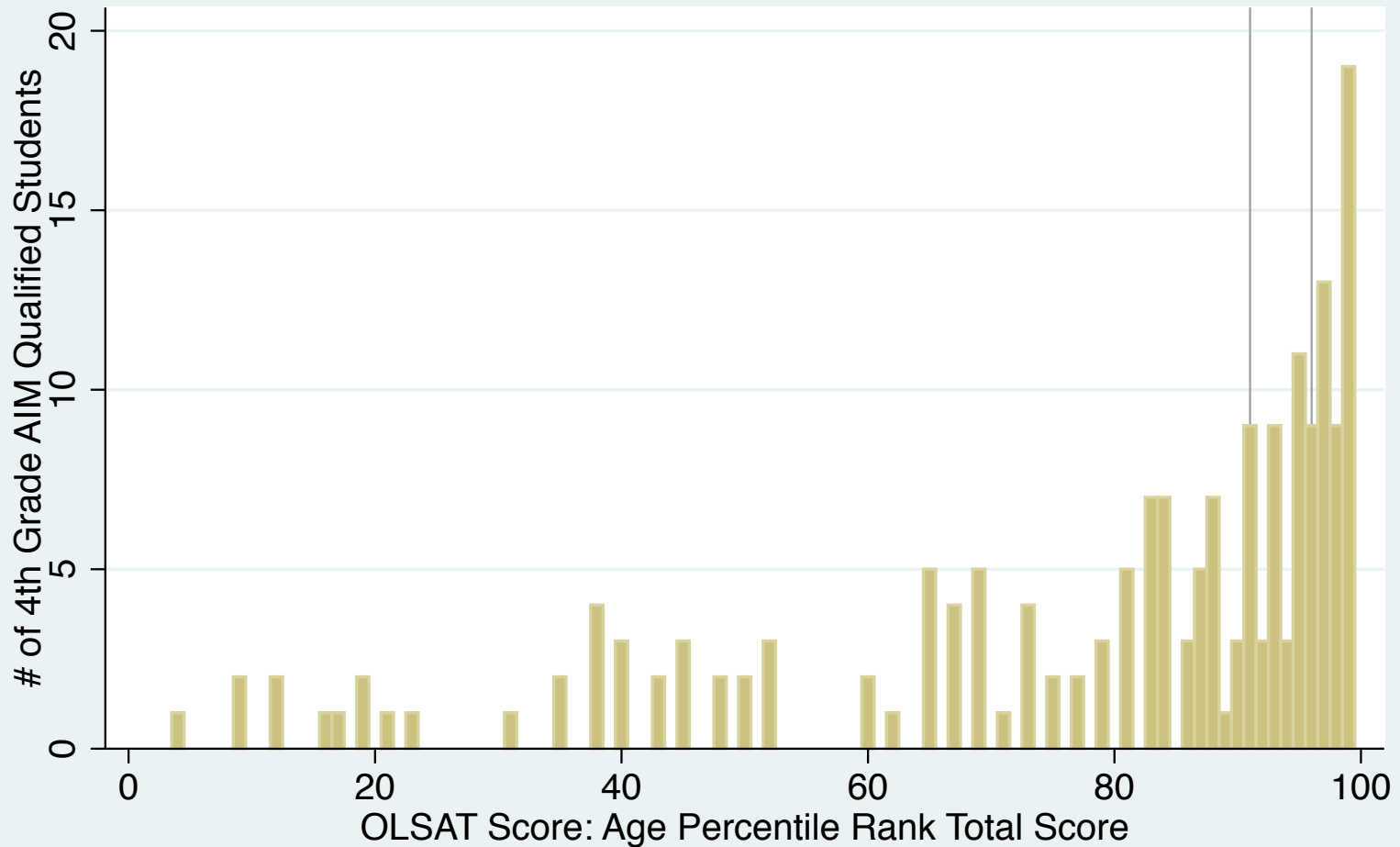
There were 203 4th grade AIM (GATE) Identified students in 2011.
Grey lines extend at OLSAT scores of 91 and 96.

OLSAT Scores: 4th Grade AIM Identified Students, 2012



There were 211 4th grade AIM (GATE) Identified students in 2012.
Grey lines extend at OLSAT scores of 91 and 96.

OLSAT Scores: 4th Grade AIM Identified Students, 2013

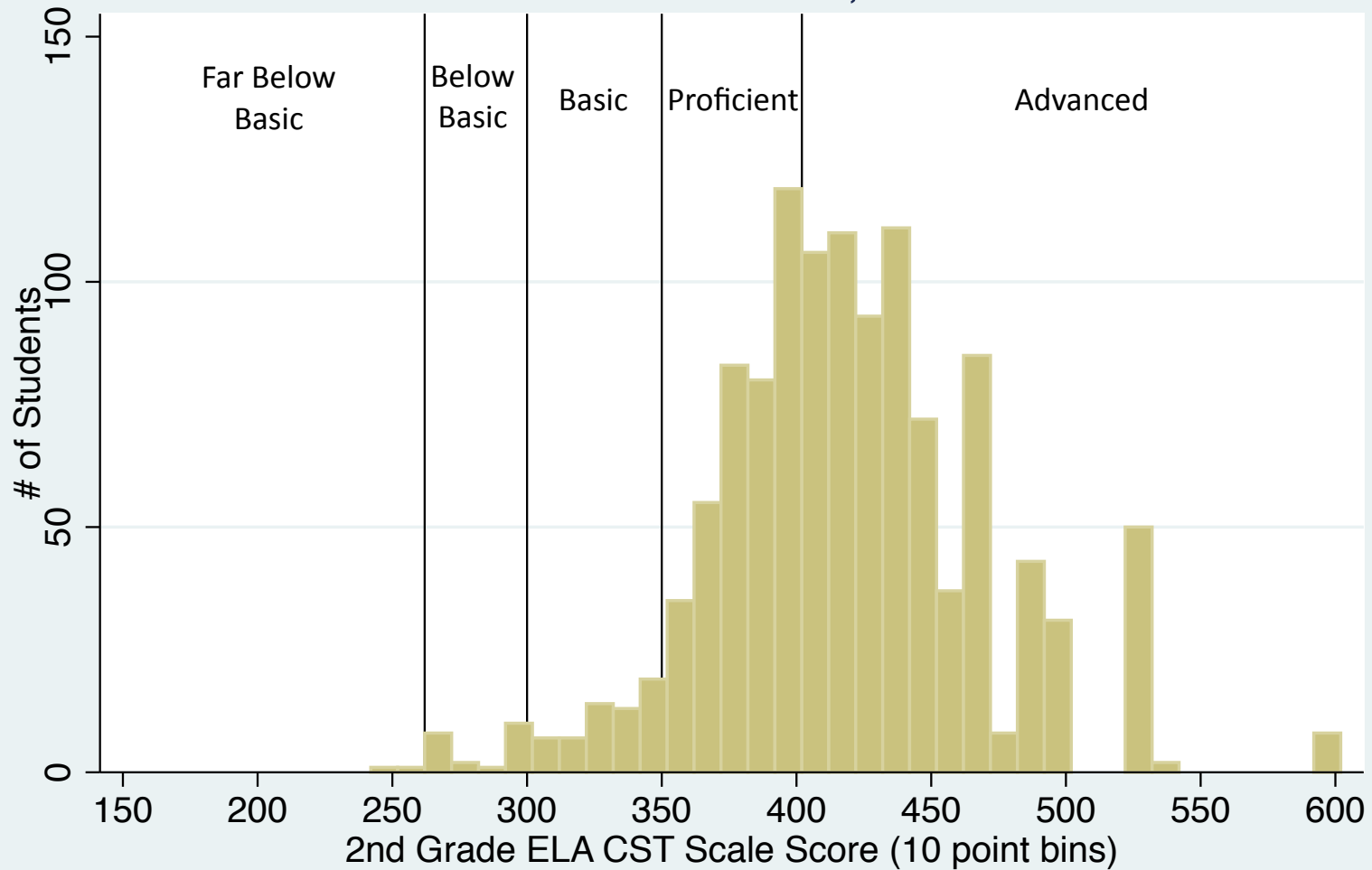


There were 196 4th grade AIM (GATE) Identified students in 2013.
Grey lines extend at OLSAT scores of 91 and 96.

Are all the students retested who have OLSAT scores outside the SEM, students with risk factors?

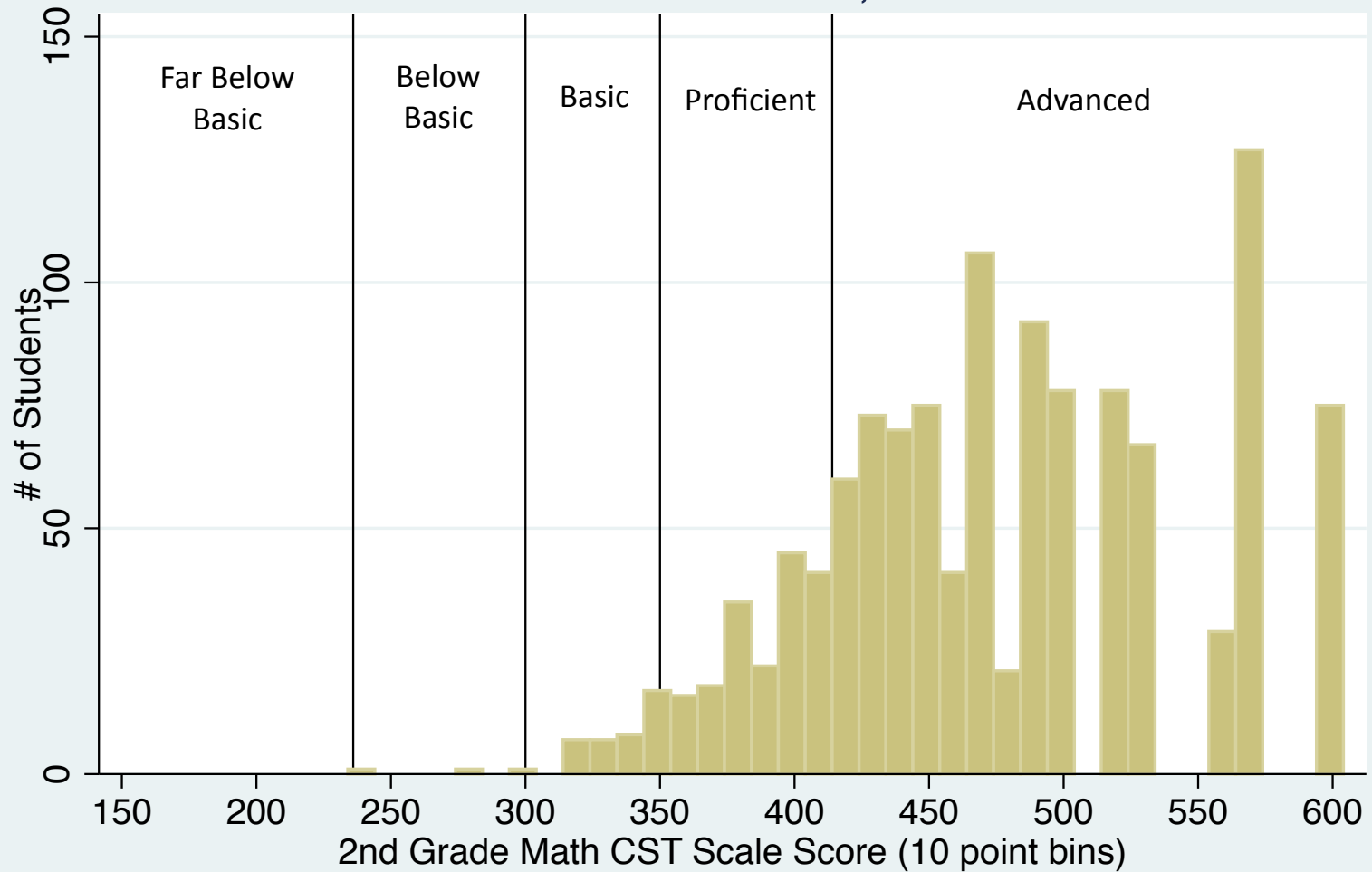
- No
 - Of the 492 students who qualify for AIM by retesting, 331 (67%) have no risk factors.
 - Of the students who qualify for AIM by retesting, with original OLSAT scores that fall outside of the standard error of measurement, 58% have no risk factors.

2nd Grade ELA CST Scale Scores, AIM Qualified Students



This sample is based on AIM (GATE) qualified 4th grade students from 2007 - 2013

2nd Grade Math CST Scale Scores, AIM Qualified Students



This sample is based on AIM (GATE) qualified 4th grade students from 2007 - 2013

What are the effects of AIM on students in AIM?

- What are the effects of participation in AIM on student outcomes for those students in AIM self contained classrooms?
- How do these effects differ by gender, race/ethnicity, free/reduced lunch eligibility and, previous academic performance?

How do we measure the affects of AIM on students in AIM?

We **cannot** compare the outcomes of students in AIM to the students not in AIM because:

- We would expect that the students in AIM self contained classrooms to have different characteristics from students not in AIM self contained classrooms.
- Students in AIM classrooms are, on average, higher achievers than students not in AIM classrooms. Students in AIM classrooms are also likely to differ in ways that are unobservable to researchers. For example, they may have parents that have more experience navigating educational resources.
- So even in the absence of AIM, we would expect the students in AIM classrooms to have, on average, better academic outcomes than the students not in AIM classrooms.

How do we measure the affects of AIM on students in AIM?

Under ideal research conditions, we would run a randomized control trial (RCT) to evaluate AIM.

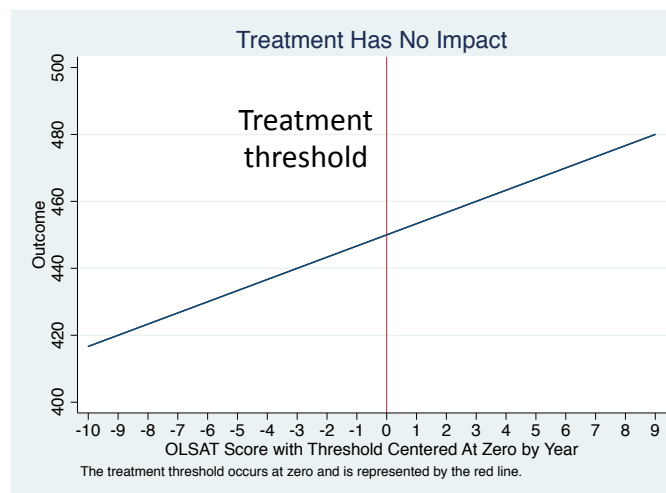
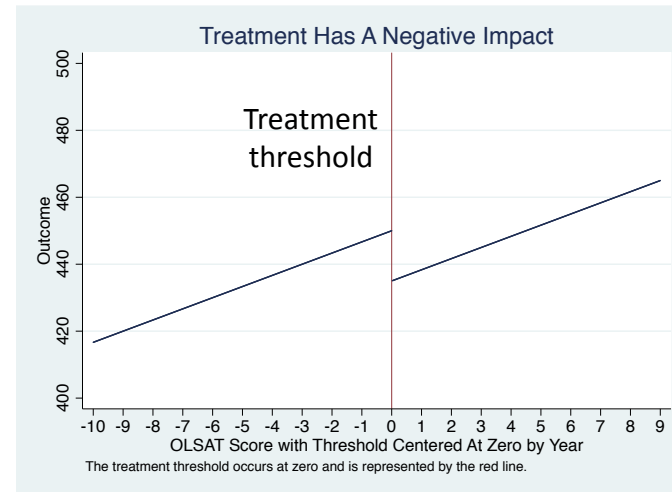
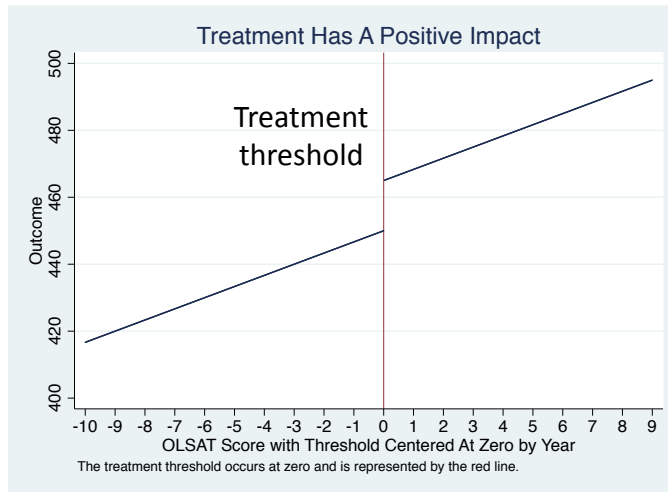
- In a randomized control trial, a researcher randomly divides the subjects into treatment and control groups.
- Since group assignment is randomly assigned, we would expect the two groups to have similar characteristics before treatment. As such, we would expect the two groups to have similar outcomes in the absence of treatment.
- If the two groups had different outcomes, we could then conclude that the treatment ***caused*** the difference in outcomes.
- There are many ethical and logistical reasons why a school district might not want to run a RCT to evaluate a program.

How do we measure the affects of AIM on students in AIM?

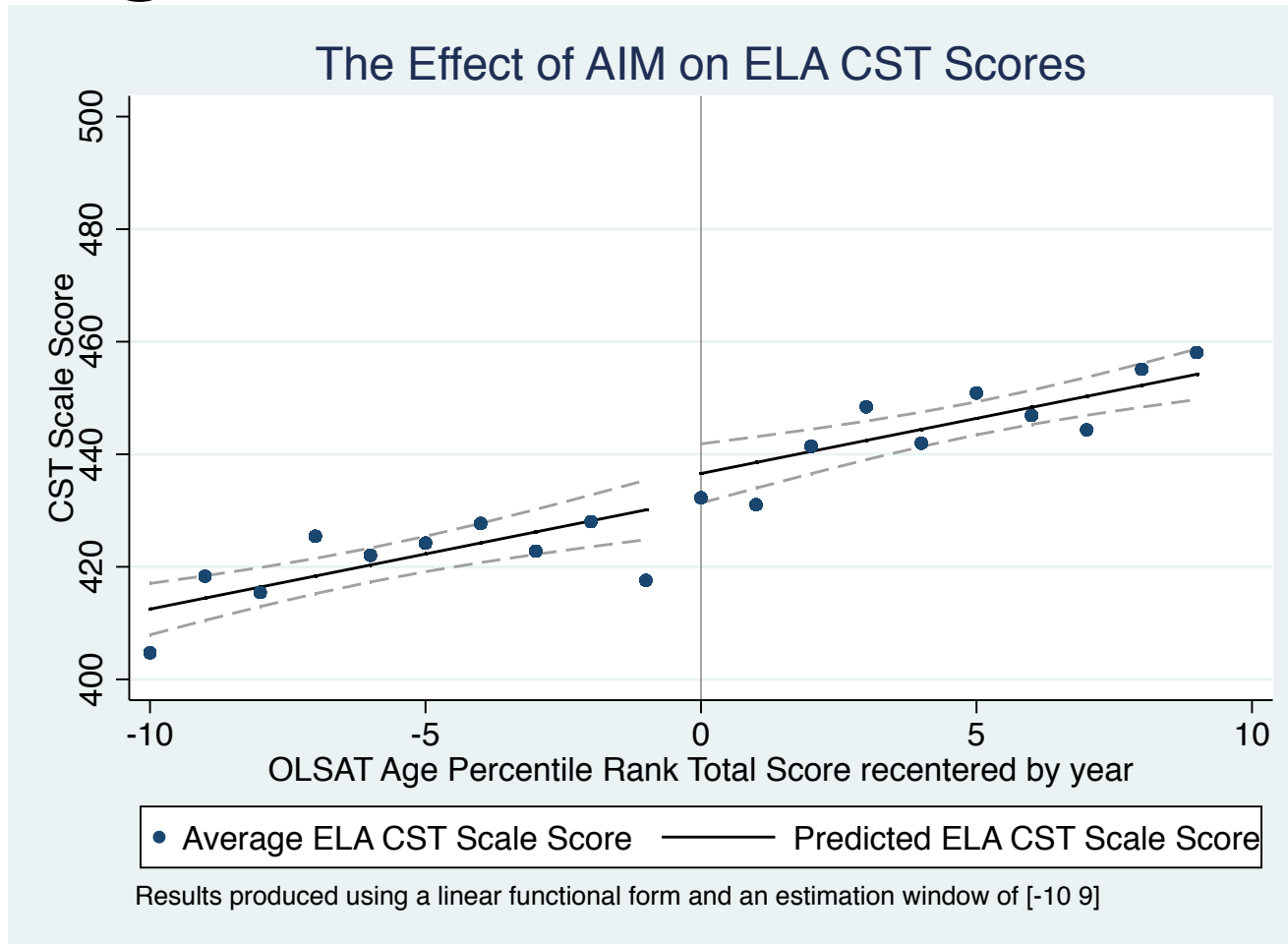
In the absence of a randomized control trial, how can we evaluate AIM?

- Methodology: Regression Discontinuity Design (RDD)
- Intuition: Scoring a 95% vs. a 96% on a test is largely random. So we can compare students who just qualify for AIM (or retesting) with students who just miss qualifying.
- RDDs are a design commonly used in education research to establish causality in the absence of random assignment. This methodology meets the U.S. Department of Education, Institute of Education Science's *What Works Clearinghouse* evidence standards.

Methodology: regression discontinuity (hypothetical scenarios)



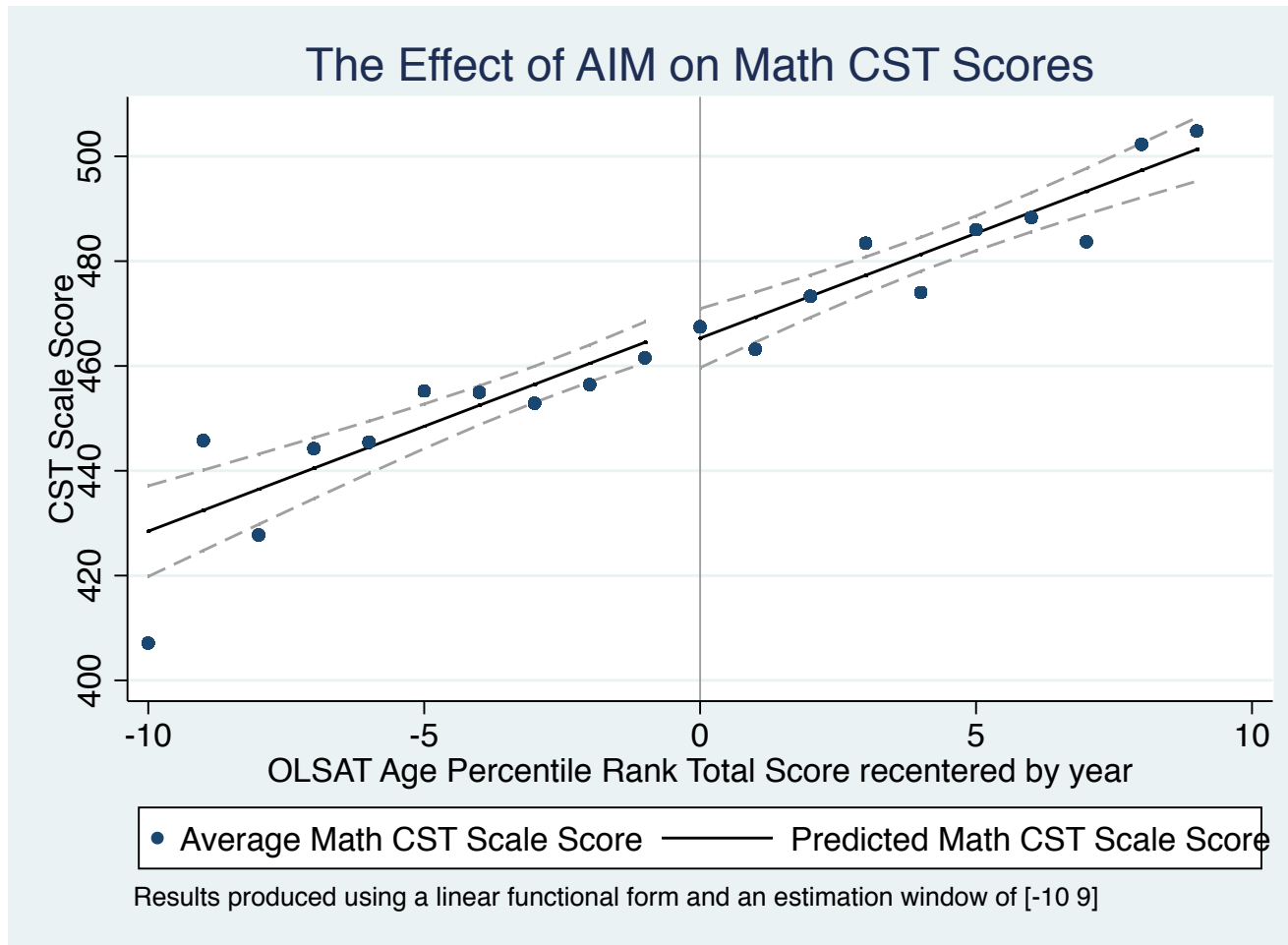
Findings: ELA



ELA CST scores from 4th, 5th, and 6th graders who took the OLSAT as 3rd graders from AY 2006 – AY 2012 are included.

AIM has no effect on the ELA CST scores of AIM students.
(effects do not differ by race/gender/free lunch/ELL)

Findings: Math



ELA CST scores from 4th, 5th, and 6th graders who took the OLSAT as 3rd graders from AY 2006 – AY 2012 are included.

AIM has no effect on the Math CST scores of AIM students.
(effects do not differ by race/gender/free lunch/ELL)

Why do we care about standardized test scores?

- Growing literature connecting achievement test scores with outcomes later in life.
 - Interventions like smaller class sizes and better teachers that raise test scores also increase the likelihood of college attendance, increase future earnings, and decrease teenage pregnancy.
(Chetty et al. 2011, Chetty et al. 2014)
- Test scores are the basis of school accountability systems.

Why do we find no effect?

Possible explanations:

- AIM has no effect.
 - Our results are consistent with recent literature on this topic, see: Bui et al. (2014) and Card & Giuliano (2014).
- AIM may affect other potential outcomes.

Previous studies meeting *What Works Clearinghouse* evidence standards:

- Bui et al. (2014), “Is Gifted Education a Bright Idea?”
 - The authors find that GATE has no effect on seventh grade test scores of GATE students in math, reading, language. There is some evidence of positive effects on science test scores.
- Card and Giuliano (2014), “Does Gifted Education Work?”
 - The authors find that GATE has no effect on fourth and fifth grade math and reading test scores. Some evidence of positive effects of GATE classes on the test scores of students who do not “qualify” as gifted but who are high achieving.

What about those not in AIM?

AIM could affect students not in AIM by causing:

- Differences in Resources
 - e.g., parental involvement
- Differences in instruction
 - e.g., teacher training/experience, differentiation, and the curriculum
- Peer group changes
 - e.g., achievement, behavior, and motivation

What are the effects of AIM on students ***not*** in AIM?

- What are the effects of participation in AIM on student outcomes for those students not eligible for AIM?
- How do these effects differ by gender, race/ethnicity, free/reduced lunch eligibility, and previous academic performance?

How do we measure the effects of AIM on students ***not*** in AIM?

In the absence of an RCT, how can we evaluate AIM?

- Methodology: Instrumental Variable (IV) Student Growth Model
- Intuition: Compare the outcomes of students with similar demographics and test scores in the same schools in years when more students score above an AIM threshold with years when fewer students score above an AIM threshold.*
- IV growth models are a quasi-experimental design commonly used in education research to establish causality in the absence of random assignment. This methodology meets the U.S. Department of Education, Institute of Education Science's *What Works Clearinghouse* evidence standards.

* We focus on schools that have self contained AIM classrooms

Findings:

We find no evidence that the AIM program has any effect on students who do not qualify for AIM.

- Our estimated effects overall are small and not significant.
 - However, we find significant negative effects for Hispanic students
- Departure of students for the AIM program lowers the test scores of students left behind:
 - 4th grade ELA CST scores drop by approximately 1.7 points, on average, per student.
 - 4th grade math CST scores drop by approximately 3.9 points, on average, per student.

Conclusions:

Identification into AIM is inconsistent.

- Students who qualify for AIM through retesting and private testing have a wide range of OLSAT scores.
- It is likely that many students who are not given the opportunity to retest or do not have the resources to be privately tested could also qualify for AIM.

Conclusions:

Who qualifies for AIM has changed over time.

- Fewer students now qualify through universal testing. More students now qualify through retesting and private testing.
- The average and minimum OLSAT scores of students eligible for the AIM program has decreased over time.

Conclusions:

There is no evidence that the AIM program affects academic performance.

- The AIM program does not improve the academic performance of students in AIM self contained classrooms.
- The AIM program does not help or hurt the academic performance of students who do not qualify to be in AIM.