

# Sinking or Swimming?

Bilingual Education and the Achievement of Immigrant Youth

Jordan D. Matsudaira  
Cornell University

Comments welcome: [jordan.matsudaira@cornell.edu](mailto:jordan.matsudaira@cornell.edu)

Texas Schools Project Seminar  
October 16, 2009

## Motivation: Growing importance of education strategies for immigrant children

- ▶ Immigration is increasing, composition is more diverse:

| Year | 1000s of legal immigrants | % Americas | % Asia |
|------|---------------------------|------------|--------|
| 1960 | 290                       | 32         | 5      |
| 1980 | 600                       | 42         | 50     |
| 2001 | 1,060                     | 51         | 33     |

## Motivation: Growing importance of education strategies for immigrant children

- ▶ Immigration is increasing, composition is more diverse:

| Year | 1000s of legal immigrants | % Americas | % Asia |
|------|---------------------------|------------|--------|
| 1960 | 290                       | 32         | 5      |
| 1980 | 600                       | 42         | 50     |
| 2001 | 1,060                     | 51         | 33     |

- ▶ Language proficiency is a strong predictor of economic outcomes
  - ▶ Poverty rate of limited English proficient immigrants in NYC and LA is more than 2 times rate of proficient speakers (at 34 percent).
  - ▶ Earnings and wages of limited English proficient immigrants also lower. Trejo (1997): accounts for 25 percent of wage-gap between Mexican Americans and non-Hispanic whites.

## Motivation: Differing Views of Role of Bilingual Education

- ▶ Civil right necessary to ensure equality of education:
  - ▶ “There is no equality of treatment merely by providing students with the same facilities, textbooks, teachers, and curriculum; for students who do not understand English are effectively foreclosed from any meaningful education”  
— Justice Douglas, *Lau v. Nichols*, 1974

vs.

## Motivation: Differing Views of Role of Bilingual Education

- ▶ Civil right necessary to ensure equality of education:
  - ▶ “There is no equality of treatment merely by providing students with the same facilities, textbooks, teachers, and curriculum; for students who do not understand English are effectively foreclosed from any meaningful education”  
— Justice Douglas, *Lau v. Nichols*, 1974

vs.

- ▶ ‘Traps’ immigrant children in substandard learning environment:
  - ▶ “It is absolutely wrong and against American concepts to have a bilingual education program that is now openly, admittedly dedicated to preserving their native language and never getting them adequate in English so they can go out into the job market and participate.”  
— former President Reagan, 1981

# Research Questions

- ▶ Are Bilingual Education programs a 'trap'?

# Research Questions

- ▶ Are Bilingual Education programs a 'trap'?
- ▶ Is English Immersion superior in promoting the achievement of LEP students?

# Why Might ELS Matter for Achievement? Mechanisms

- ▶ Instruction in the Native Language



# Why Might ELS Matter for Achievement? Mechanisms

- ▶ Instruction in the Native Language
- ▶ Peer Effects (through peers achievement, race, language, etc.)

# Why Might ELS Matter for Achievement? Mechanisms

- ▶ Instruction in the Native Language
- ▶ Peer Effects (through peers achievement, race, language, etc.)
- ▶ **Teacher Quality**

# Why Might ELS Matter for Achievement? Mechanisms

- ▶ Instruction in the Native Language
- ▶ Peer Effects (through peers achievement, race, language, etc.)
- ▶ Teacher Quality
- ▶ **Class Size**

## Outline and Definitions

- ▶ Policy and Research Background
- ▶ Research Design and Data
- ▶ Empirical Results
- ▶ Limitations and Implications

## Outline and Definitions

- ▶ Policy and Research Background
- ▶ Research Design and Data
- ▶ Empirical Results
- ▶ Limitations and Implications

### Glossary:

- ▶ Limited English Proficient (LEP)
- ▶ Bilingual education programs
- ▶ English as a Second Language (ESL) programs
- ▶ English Language Services (ELS)

# Policy Trends

- ▶ 1960s to 1970s: Increased support for Bilingual Education
  - ▶ Bilingual Education Act, 1968
  - ▶ Lau v. Nichols, 1974
  - ▶ OCR Lau Remedies, 1975

## Policy Trends

- ▶ 1960s to 1970s: Increased support for Bilingual Education
  - ▶ Bilingual Education Act, 1968
  - ▶ Lau v. Nichols, 1974
  - ▶ OCR Lau Remedies, 1975
  
- ▶ 1980s & after: support for English immersion over Bilingual
  - ▶ Changes to BEA funding rules
  - ▶ No Child Left Behind Act
  - ▶ State initiatives to ban bilingual instruction

# Previous Research

- ▶ Three large national studies:
  - ▶ NAS on Burkheimer (1989) and Ramirez (1991):

“The studies do not license the conclusion that any one type of program is superior to any other nor that the programs are equally effective. Even if one of the programs was definitively superior, the studies as planned and executed could well have failed to find the effect.” (Meyer and Fienberg, 1991).



## Previous Research

- ▶ Three large national studies:
  - ▶ NAS on Burkheimer (1989) and Ramirez (1991):

“The studies do not license the conclusion that any one type of program is superior to any other nor that the programs are equally effective. Even if one of the programs was definitively superior, the studies as planned and executed could well have failed to find the effect.” (Meyer and Fienberg, 1991).
- ▶ Hundreds of Smaller Studies
  - ▶ 4-5 Competing meta-analyses of about 150 studies
  - ▶ Only 5 use random assignment
  - ▶ Of these, nothing new (4/5 older than 30 years); only 1 peer-reviewed; none with  $N > 200$ .

## Previous Research

- ▶ Three large national studies:
  - ▶ NAS on Burkheimer (1989) and Ramirez (1991):

“The studies do not license the conclusion that any one type of program is superior to any other nor that the programs are equally effective. Even if one of the programs was definitively superior, the studies as planned and executed could well have failed to find the effect.” (Meyer and Fienberg, 1991).
- ▶ Hundreds of Smaller Studies
  - ▶ 4-5 Competing meta-analyses of about 150 studies
  - ▶ Only 5 use random assignment
  - ▶ Of these, nothing new (4/5 older than 30 years); only 1 peer-reviewed; none with  $N > 200$ .
- ▶ In economics, Hoxby and Gordon (2002), Lopez & Mora (1998), Angrist, Chin, and Godoy (2006) and little else.

# Research Design

- ▶ In a Large Urban School District in the Northeast (LUSDiNE), LEP status and eligibility for ELS is based on students score on Language Assessment Test (LAT). Students scoring at or below the 40th percentile are classified as LEP and eligible for ELS; students above the 40th percentile are non-LEP & ineligible.

# Research Design

- ▶ In a Large Urban School District in the Northeast (LUSDiNE), LEP status and eligibility for ELS is based on students score on Language Assessment Test (LAT). Students scoring at or below the 40th percentile are classified as LEP and eligible for ELS; students above the 40th percentile are non-LEP & ineligible.
- ▶ If there is some randomness in LAT score, then students whose scores are 'near' the 40th percentile should be nearly identical. Generates quasi-random assignment to 'treatment' and 'control' groups:

# Research Design

- ▶ In a Large Urban School District in the Northeast (LUSDiNE), LEP status and eligibility for ELS is based on students score on Language Assessment Test (LAT). Students scoring at or below the 40th percentile are classified as LEP and eligible for ELS; students above the 40th percentile are non-LEP & ineligible.
- ▶ If there is some randomness in LAT score, then students whose scores are 'near' the 40th percentile should be nearly identical. Generates quasi-random assignment to 'treatment' and 'control' groups:
- ▶ If other determinants of achievement are smooth at the 40th percentile, differences in outcomes can be causally attributed to differences in program participation (this identification assumption is testable).

## LUSDiNE Data

- ▶ Administrative, student level data for all students while they are in grades 3 through 8 for years ending in Spring 1999-2002. Data exists for slightly less than 1 million students; there are approximately 2 million student-years of data.

# LUSDiNE Data

- ▶ Administrative, student level data for all students while they are in grades 3 through 8 for years ending in Spring 1999-2002. Data exists for slightly less than 1 million students; there are approximately 2 million student-years of data.
- ▶ Information on
  - ▶ LAT scores from 1998 through 2002 and program participation
  - ▶ Achievement tests in math and reading
  - ▶ Demographics (age, sex, race, free lunch, country of birth, language, census zip info, etc.)

# LUSDiNE Data

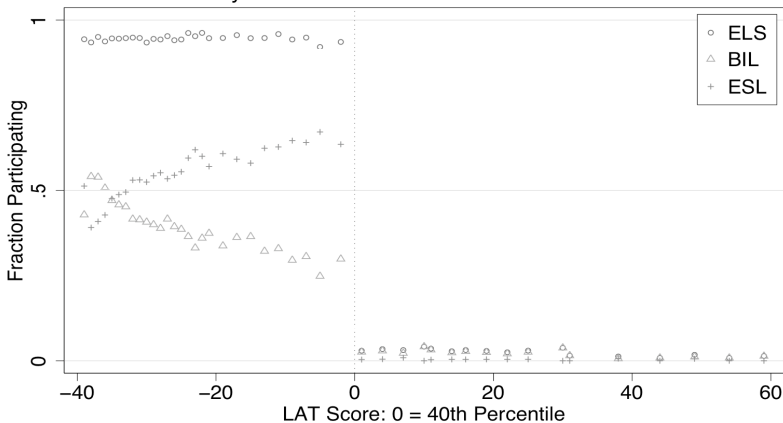
- ▶ Administrative, student level data for all students while they are in grades 3 through 8 for years ending in Spring 1999-2002. Data exists for slightly less than 1 million students; there are approximately 2 million student-years of data.
- ▶ Information on
  - ▶ LAT scores from 1998 through 2002 and program participation
  - ▶ Achievement tests in math and reading
  - ▶ Demographics (age, sex, race, free lunch, country of birth, language, census zip info, etc.)
- ▶ For this study, I use all students ( $n=183,254$  student years) who
  - ▶ Are 'at risk' for needing ELS: take the LAT test between 1998 and 2001 prior to entering grade 3-8
  - ▶ Remain in the school system at least one year later
  - ▶ Are not new entrants to the school system



## Sample Descriptives

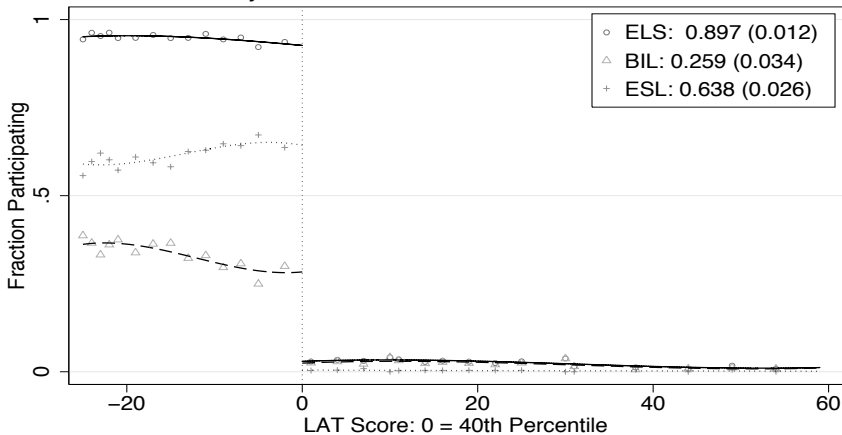
| Variable                        | Full Sample     | ELS             | No ELS          | Diff.           |
|---------------------------------|-----------------|-----------------|-----------------|-----------------|
| <i>Following Year Outcomes</i>  |                 |                 |                 |                 |
| Reading z-score                 | -.627<br>(.003) | -.872<br>(.003) | -.237<br>(.004) | -.635<br>(.005) |
| Math z-score                    | -.432<br>(.002) | -.611<br>(.003) | -.07<br>(.004)  | -.541<br>(.005) |
| Promoted                        | .914<br>(.001)  | .901<br>(.001)  | .946<br>(.001)  | -.045<br>(.002) |
| <i>Baseline Characteristics</i> |                 |                 |                 |                 |
| Asian                           | .154<br>(.001)  | .144<br>(.001)  | .181<br>(.002)  | -.037<br>(.002) |
| Hispanic                        | .706<br>(.001)  | .731<br>(.001)  | .639<br>(.002)  | .093<br>(.002)  |
| Black                           | .056<br>(.001)  | .053<br>(.001)  | .064<br>(.001)  | -.011<br>(.001) |
| White                           | .083<br>(.001)  | .071<br>(.001)  | .115<br>(.001)  | -.045<br>(.002) |
| Free/Red Lunch                  | .966<br>(.001)  | .972<br>(.001)  | .950<br>(.001)  | .023<br>(.002)  |
| Foreign Born                    | .475<br>(.001)  | .493<br>(.001)  | .430<br>(.002)  | .063<br>(.003)  |
| Home Lang = Spanish             | .698<br>(.001)  | .731<br>(.001)  | .612<br>(.002)  | .119<br>(.002)  |

Assignment to ELS, Bilingual and ESL for Grade 4  
One year after LAT (All Years)  
by LAT Score Relative to 40th Percentile



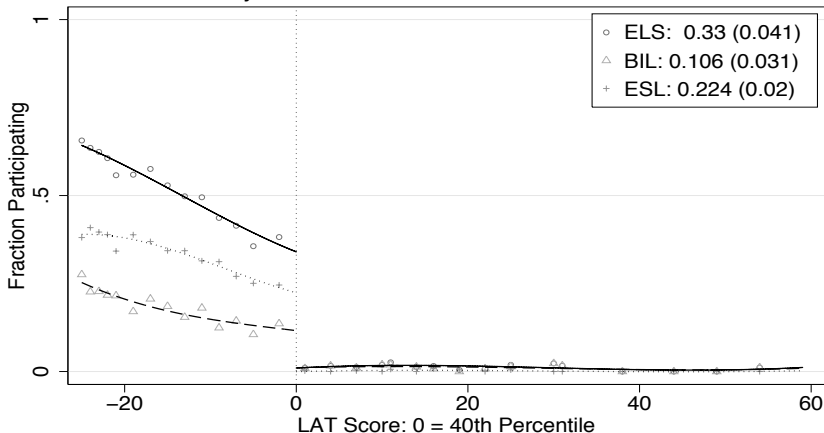
Note: Discontinuities estimated using polynomial of degree 3 on 1 year sample.

Assignment to ELS, Bilingual and ESL for Grade 4  
One year after LAT (All Years)  
by LAT Score Relative to 40th Percentile



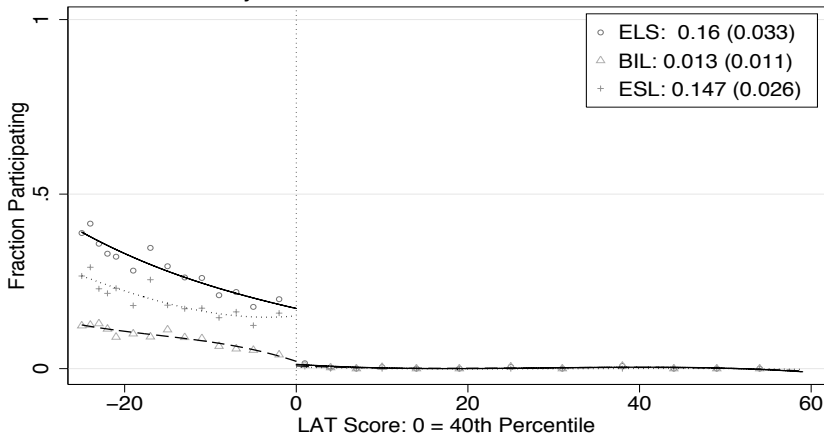
Note: Discontinuities estimated using polynomial of degree 3 on 1 year sample.

Assignment to ELS, Bilingual and ESL for Grade 4  
Two years after LAT (All Years)  
by LAT Score Relative to 40th Percentile



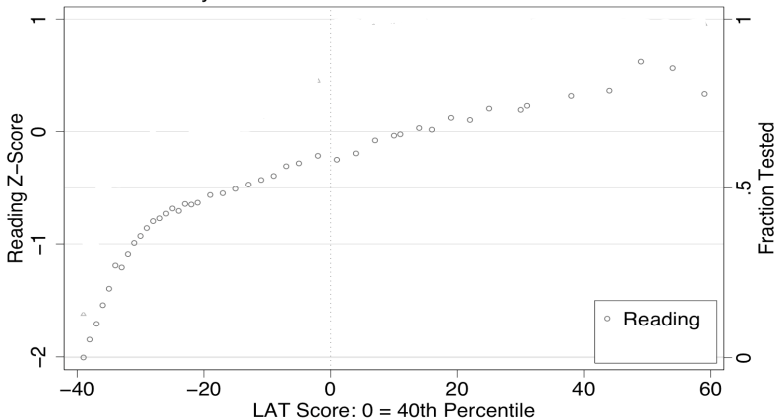
Note: Discontinuities estimated using polynomial of degree 3 on 2 year sample.

Assignment to ELS, Bilingual and ESL for Grade 4  
Three years after LAT (All Years)  
by LAT Score Relative to 40th Percentile



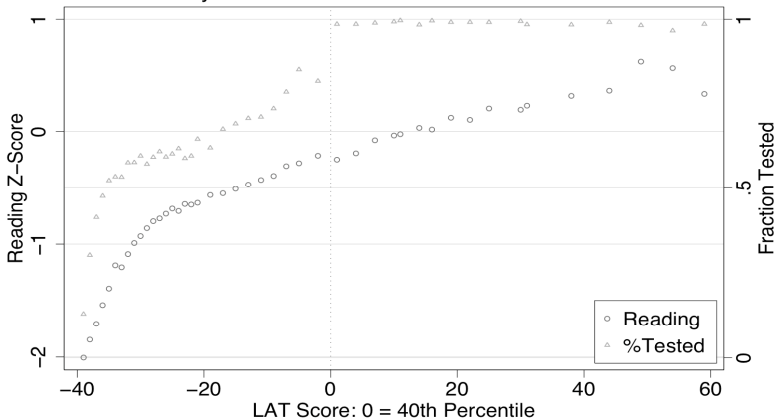
Note: Discontinuities estimated using polynomial of degree 3 on 3 year sample.

### Reading Achievement Outcomes for Grade 4 One year after LAT (All Years) by LAT Score Relative to 40th Percentile



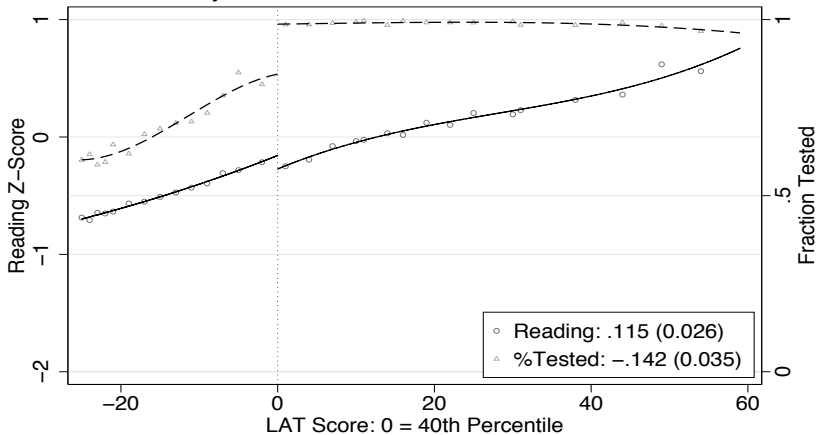
Discontinuities estimated using polynomial of degree 3 on 1-year sample.

### Reading Achievement Outcomes for Grade 4 One year after LAT (All Years) by LAT Score Relative to 40th Percentile



Discontinuities estimated using polynomial of degree 3 on 1-year sample.

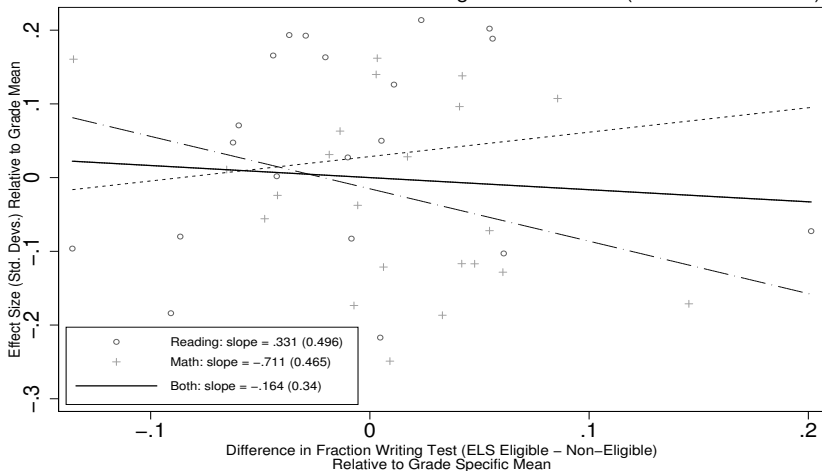
### Reading Achievement Outcomes for Grade 4 One year after LAT (All Years) by LAT Score Relative to 40th Percentile



Discontinuities estimated using polynomial of degree 3 on 1-year sample.



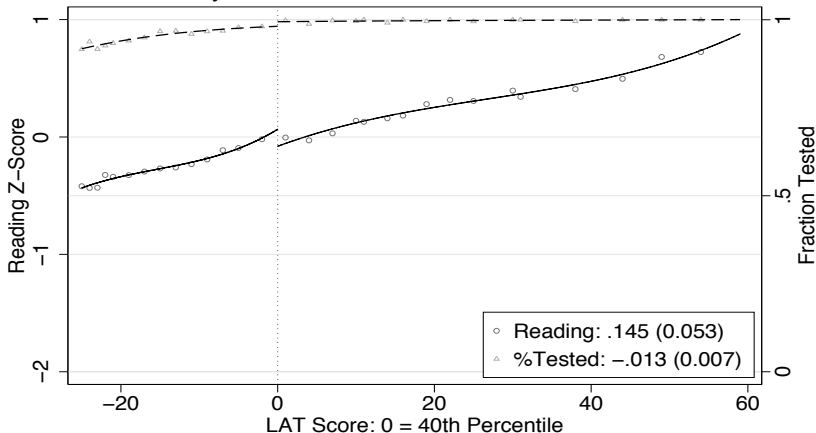
### The Relationship between Effect Size and The Differential Fraction of Students Taking Outcome Tests (Within Grade Level)



Note: Effect and attrition sizes are estimated from models using polynomial of degree 3.

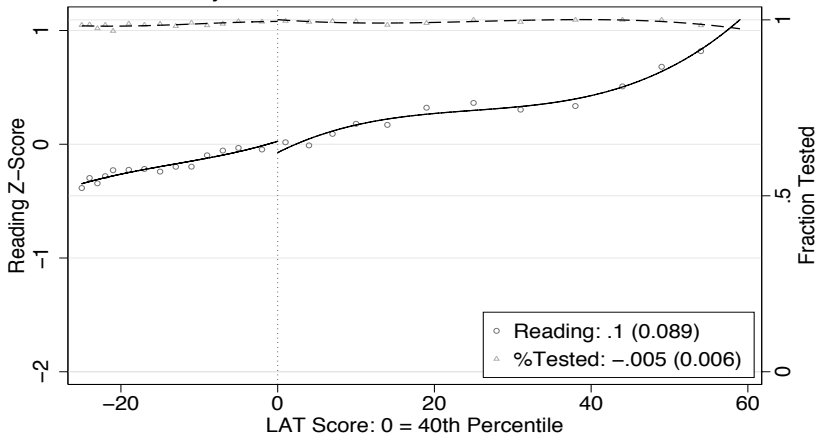


### Reading Achievement Outcomes for Grade 4 Two years after LAT (All Years) by LAT Score Relative to 40th Percentile



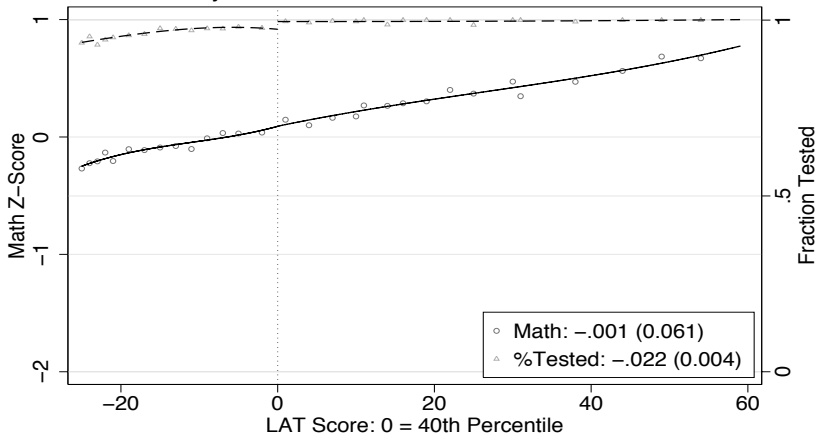
Discontinuities estimated using polynomial of degree 3 on 2-year sample.

### Reading Achievement Outcomes for Grade 4 Three years after LAT (All Years) by LAT Score Relative to 40th Percentile



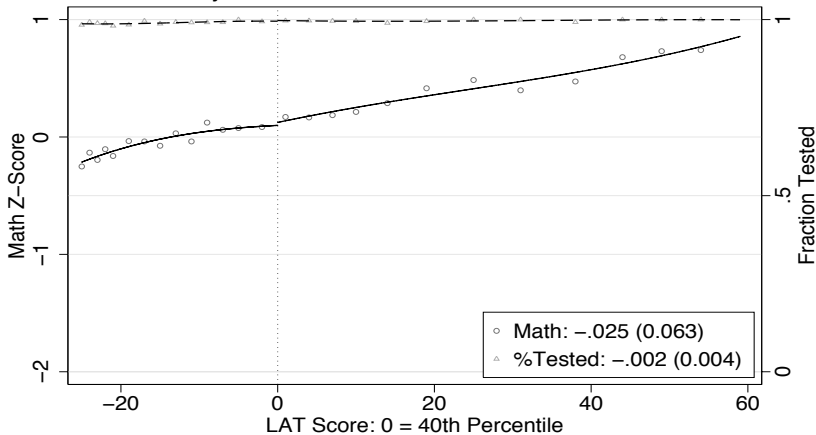
Discontinuities estimated using polynomial of degree 3 on 3-year sample.

### Math Achievement Outcomes for Grade 4 Two years after LAT (All Years) by LAT Score Relative to 40th Percentile



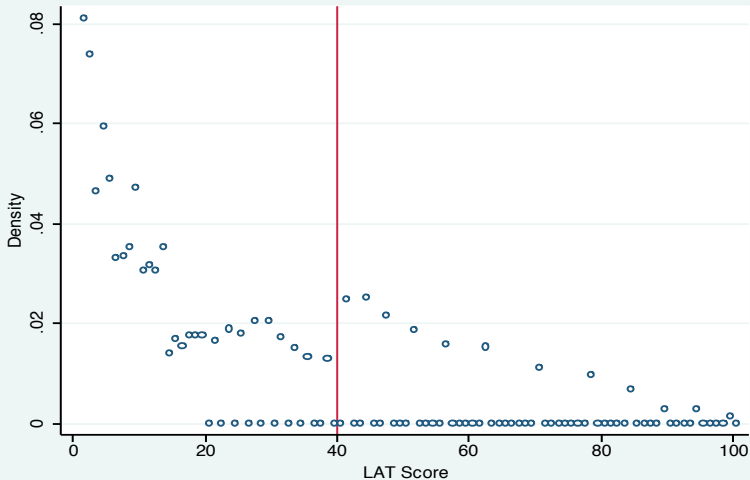
Discontinuities estimated using polynomial of degree 3 on 2-year sample.

### Math Achievement Outcomes for Grade 4 Three years after LAT (All Years) by LAT Score Relative to 40th Percentile

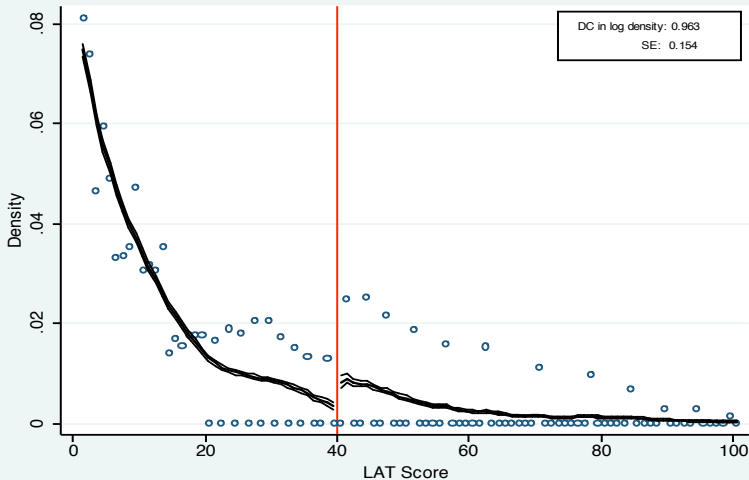


Discontinuities estimated using polynomial of degree 3 on 3-year sample.

## Validity of Design: LAT Score Density

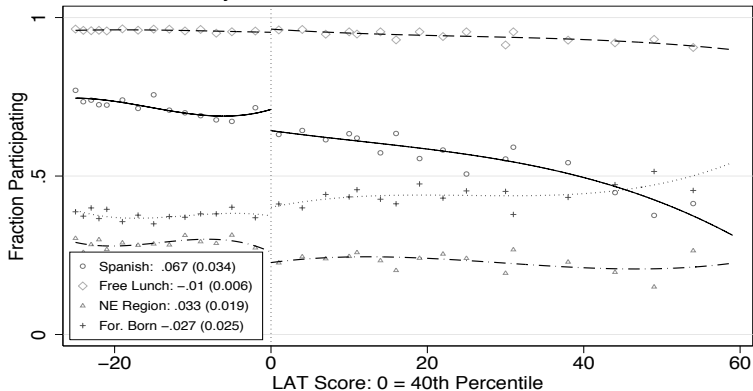


## Validity of Design: Discontinuous Density?



# Is exchangeability assumption compromised?

Selected Covariates for Grade 4  
One year after LAT (All Years)  
by LAT Score Relative to 40th Percentile



Note: Discontinuities estimated using polynomial of degree 3 on 1 year sample.



## Estimates for All Grades—Covariate Adjusted

|            | Reading (w/covariates) | Math (w/covariates) |
|------------|------------------------|---------------------|
| Grade 3    | .091<br>(.103)         | .27<br>(.098)       |
| Grade 4    | .123<br>(.035)         | .022<br>(.043)      |
| Grade 5    | .035<br>(.07)          | -.098<br>(.055)     |
| Grade 6    | -.07<br>(.098)         | .01<br>(.024)       |
| Grade 7    | .06<br>(.043)          | .178<br>(.05)       |
| All Grades | .08<br>(.024)          | .03<br>(.018)       |

# Implications

- ▶ Marginal reductions in the number of children participating in ELS are unlikely to affect math achievement, but may have an adverse effect on reading achievement:
  - ▶ Time limits for ELS eligibility have been considered (but not yet passed)
  - ▶ NCLB incentives schools to rapidly push ELS participants into mainstream classes

# Implications

- ▶ Marginal reductions in the number of children participating in ELS are unlikely to affect math achievement, but may have an adverse effect on reading achievement:
  - ▶ Time limits for ELS eligibility have been considered (but not yet passed)
  - ▶ NCLB incentives schools to rapidly push ELS participants into mainstream classes
- ▶ If the benefit of ELS is inversely related to initial proficiency, then the results suggest that the effect of ELS for the average participant are positive (at least non-negative).
  - ▶ A ban of ELS may not raise achievement.
  - ▶ Further work using data from the 1980s may help to provide information on the treatment effect on the treated.

## Limitations and Further Research

- ▶ Treatment effect for students with LAT score at 40th percentile vs. average treatment effect of treated. Applying the design in contexts where eligibility threshold is different would yield important information:
  - ▶ In the 1980s, the 20th percentile was used as the eligibility cutoff in LUSDiNE
  - ▶ Other districts use different cutoffs: e.g., in Florida a 32th percentile rule is followed

## Limitations and Further Research

- ▶ Treatment effect for students with LAT score at 40th percentile vs. average treatment effect of treated. Applying the design in contexts where eligibility threshold is different would yield important information:
  - ▶ In the 1980s, the 20th percentile was used as the eligibility cutoff in LUSDiNE
  - ▶ Other districts use different cutoffs: e.g., in Florida a 32th percentile rule is followed
- ▶ Sample limited to students already in ELS as of 3rd Grade. Effects in earlier grades may be more pronounced, and long-term outcomes are arguably more important measures of success:
  - ▶ Follow students from school entry to graduation/dropout, college enrollment
  - ▶ Link to data on labor market outcomes like wages, sector of employment (ethnic enclave?), etc.
  - ▶ Measures of civic integration (voting), etc.

## Limitations and Further Research ... cont.

- ▶ Limited information about ELS classroom environment and therefore which mechanisms are at work. Class room experience could differ in other ways besides role of native language in the curriculum:
  - ▶ Teacher Quality, Class Size, Peers
  - ▶ Need to get creative, but exploring heterogeneity of effects across different contexts should yield insights

# Inside the Black Box... mechanisms (peer quality)

Average Reading Achievement of Peers for Grade 4  
by LAT Score Relative to 40th Percentile

