Decoding Success

Evidence from a Federal Phonics Initiative

Alex Hollingsworth
Ohio State & NBER

Alberto Ortega
Indiana University & NBER

Nikolaos Zirogiannis
Indiana University

November 2026

Paper · Non-technical summary

Results preview I

Question and design

1. **Question:** does how reading is taught change achievement and later-life outcomes?
   - Specific focus on **phonics**-based instruction incentivized by the federal U.S. Reading First program (started in 2003)

Results preview I

Question and design

1. **Question:** does how reading is taught change achievement and later-life outcomes?
   - Specific focus on **phonics**-based instruction incentivized by the federal U.S. Reading First program (started in 2003)
2. **Design:**
   - **DiD** using Reading First's 2002 eligibility rule
   - **IV** using formula-predicted funding (fixed 2002 inputs)

Results preview I

Question and design

1. **Question:** does how reading is taught change achievement and later-life outcomes?
   - Specific focus on **phonics**-based instruction incentivized by the federal U.S. Reading First program (started in 2003)
2. **Design:**
   - **DiD** using Reading First's 2002 eligibility rule
   - **IV** using formula-predicted funding (fixed 2002 inputs)
3. **Data/setting:** Texas, 1995–2017
   - linked administrative records
   - 3rd-grade test scores $\rightarrow$ age-25 earnings
   - digitized exam questions
   - school finance/spending data; local tax elections as a benchmark

Results preview II

Findings and mechanism

1. **Findings:** reading **+0.11 s.d.** (math +0.11) for **$75–$132 per K–3 student-year**; gains persist through grade 8; treated students are more likely to finish high school on time, enroll in college, and complete a four-year degree, and earn **$509 more at age 25**

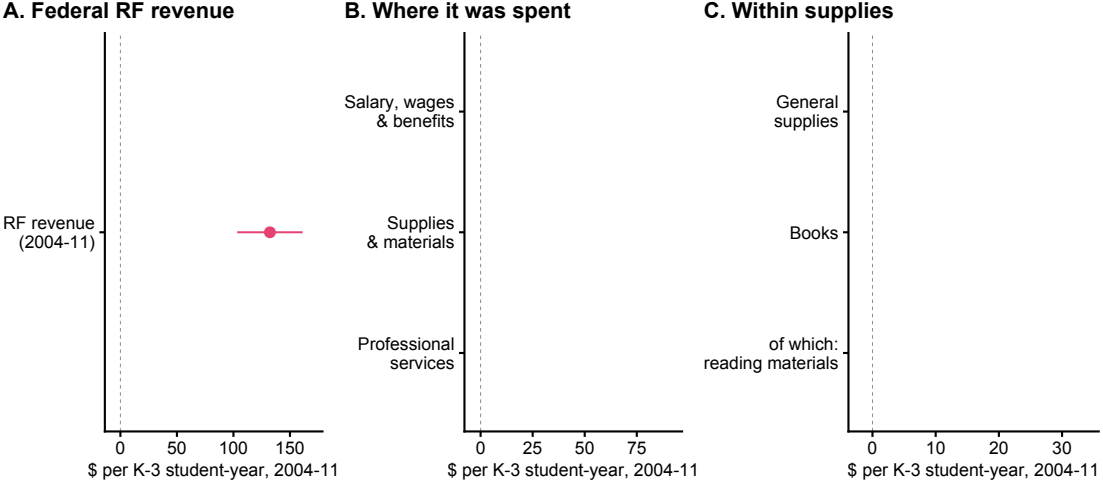
Results preview II

Findings and mechanism

1. **Findings:** reading **+0.11 s.d.** (math +0.11) for **$75–$132 per K–3 student-year**; gains persist through grade 8; treated students are more likely to finish high school on time, enroll in college, and complete a four-year degree, and earn **$509 more at age 25**

2. **Mechanism:** **digitized exam questions**, school finance data, a tax-election benchmark, and the federal evaluation all point to a change in **how reading was taught**

Results preview: district finances



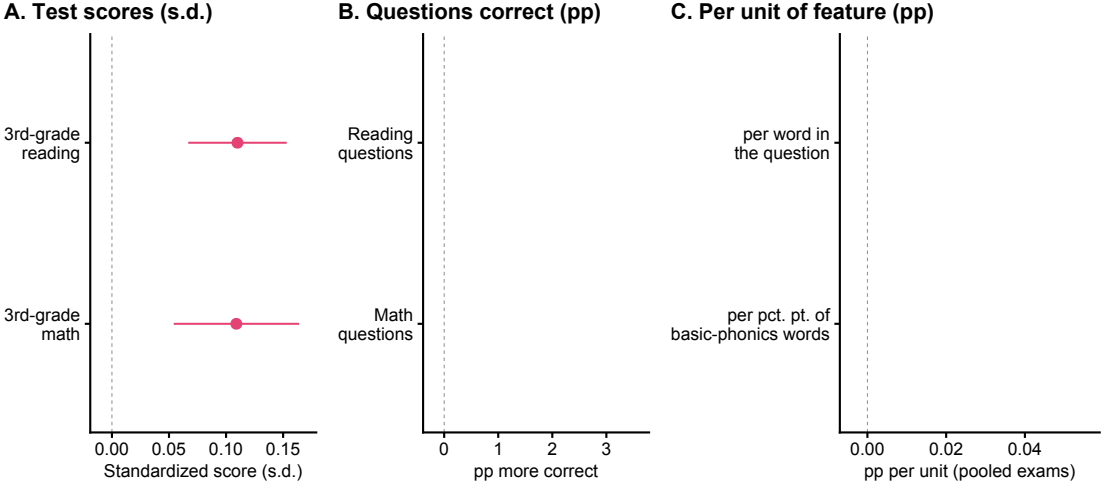
Results preview: district finances



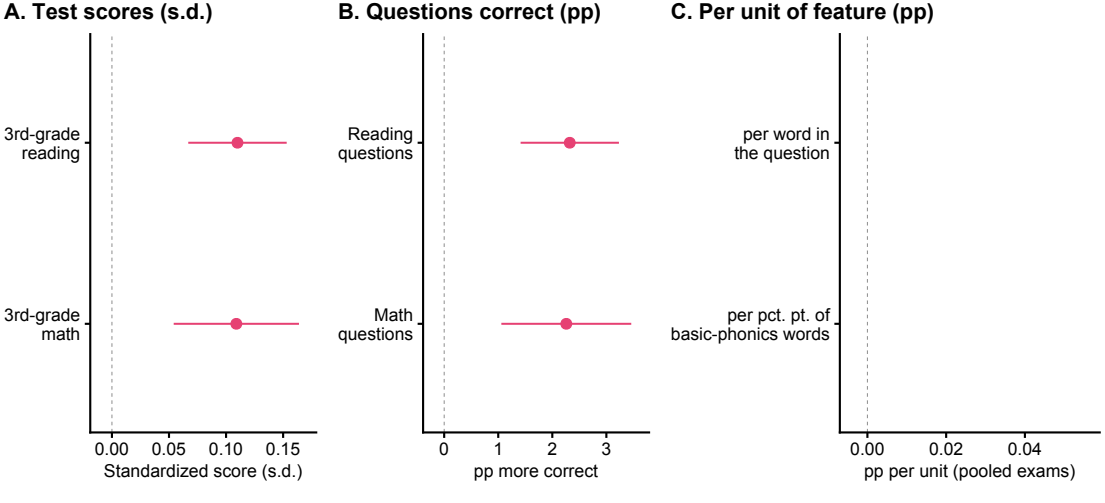
Results preview: district finances



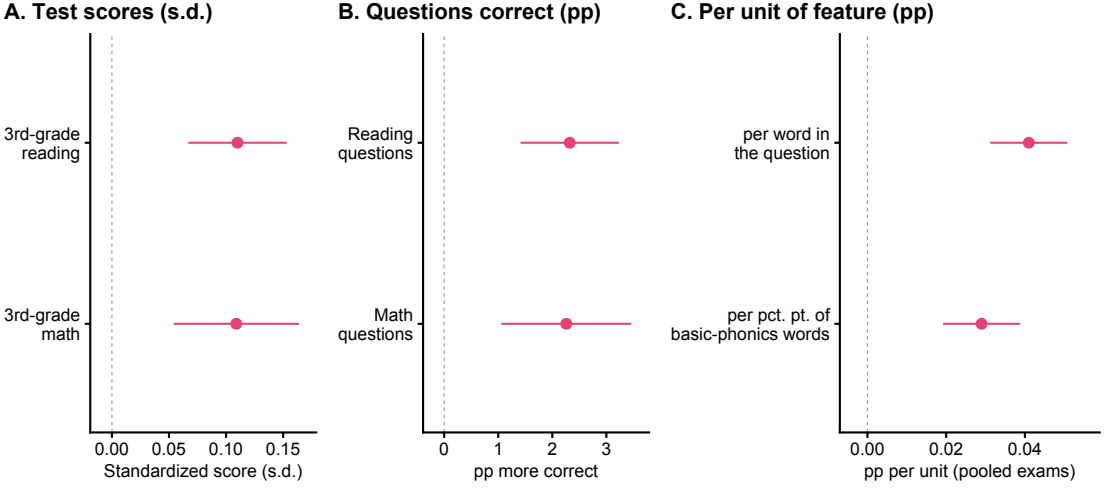
Results preview: the classroom



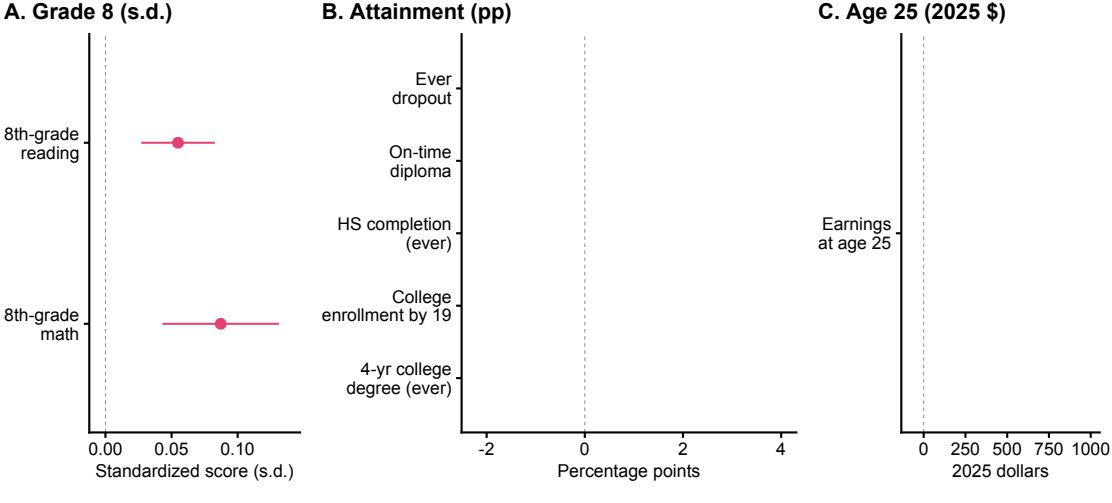
Results preview: the classroom



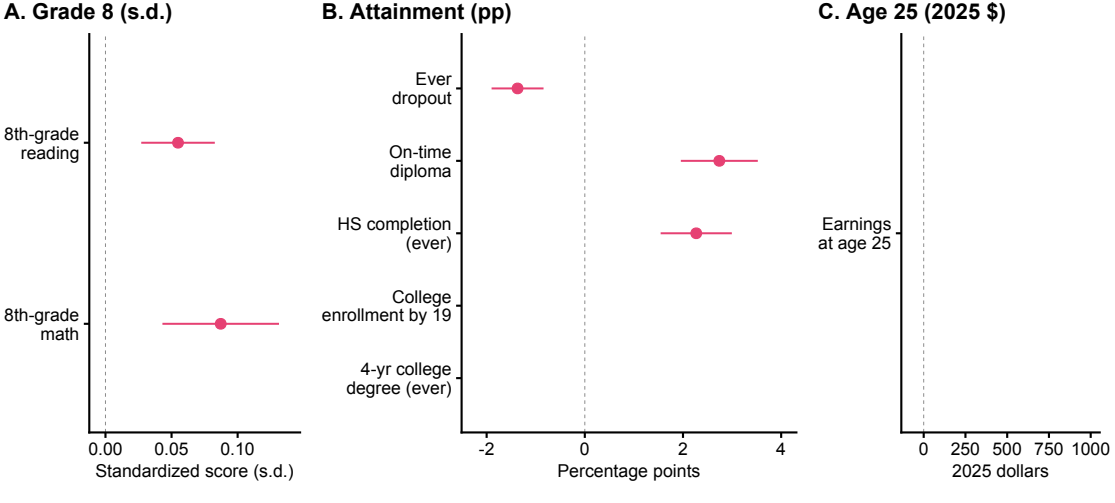
Results preview: the classroom



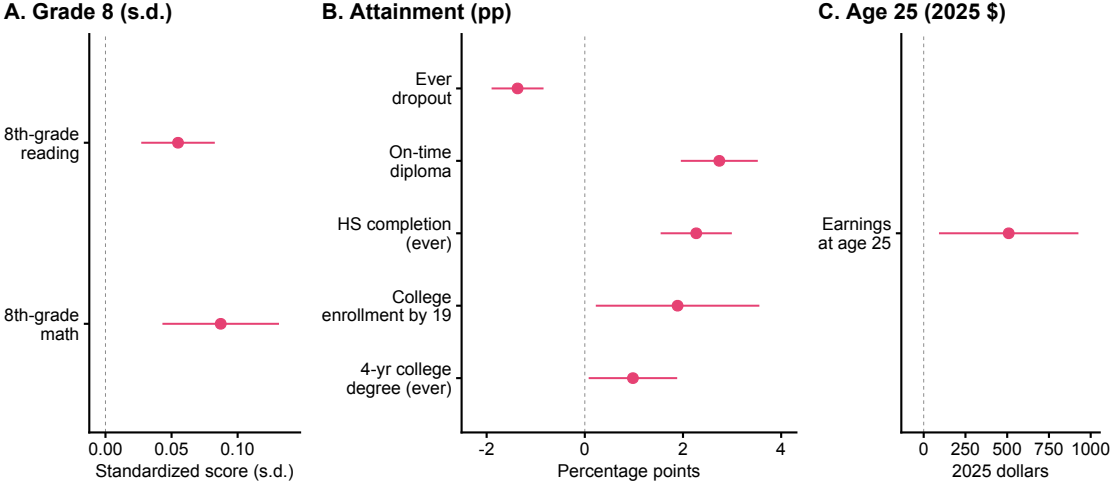
Results preview: the long run



Results preview: the long run



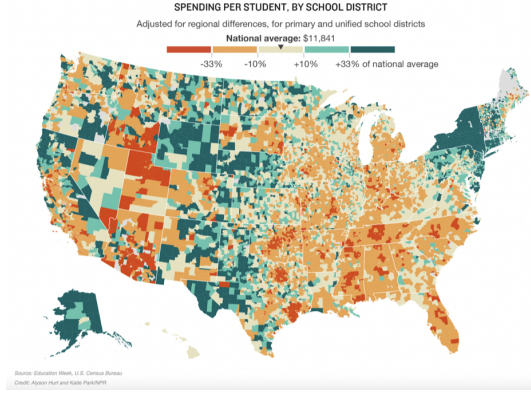
Results preview: the long run



Motivation

The US spends a lot on public education

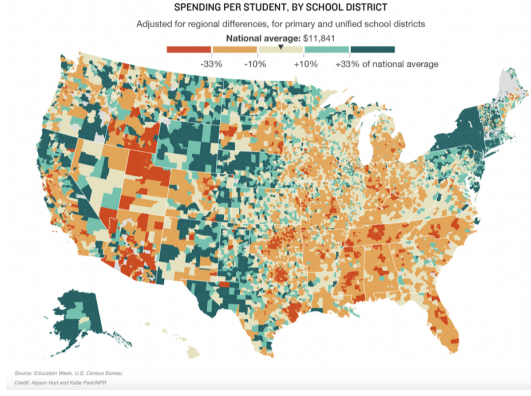
- Exceeded **$12,000 per student** in 2017



Map: NPR School Money, "Why America's Schools Have a Money Problem" (Apr. 2016), Alyson Hurt and Katie Park/NPR; data: Education Week Research Center, U.S. Census Bureau (FY2013)

The US spends a lot on public education

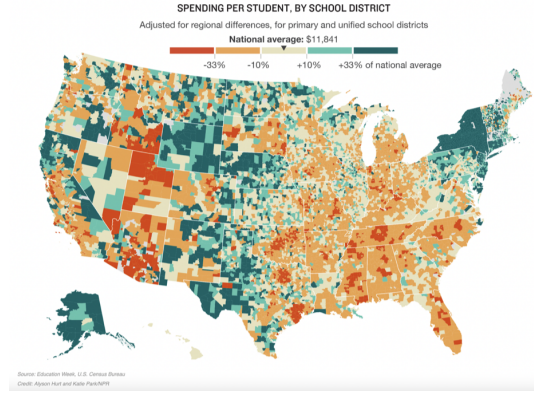
- Exceeded **$12,000 per student** in 2017
- 30% of state and 8% of national expenditures



Map: NPR School Money, "Why America's Schools Have a Money Problem" (Apr. 2016), Alyson Hurt and Katie Park/NPR; data: Education Week Research Center, U.S. Census Bureau (FY2013)

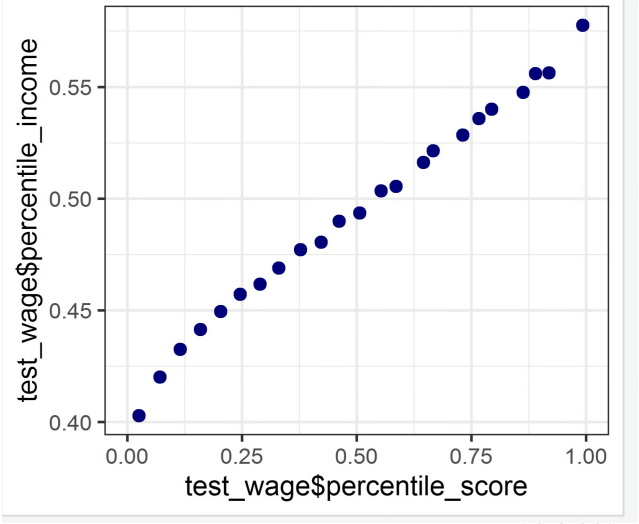
The US spends a lot on public education

- Exceeded **$12,000 per student** in 2017
- 30% of state and 8% of national expenditures
- Money **works** on average (Jackson, Johnson & Persico 2016; Jackson & Mackevicius 2024) — far less is known about **how** the marginal dollar should be spent



Map: NPR School Money, "Why America's Schools Have a Money Problem" (Apr. 2016), Alyson Hurt and Katie Park/NPR; data: Education Week Research Center, U.S. Census Bureau (FY2013)

Test scores are related to future income

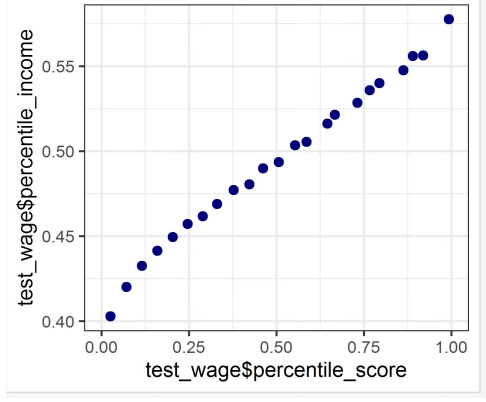


Authors' calculations from the Texas administrative data: mean earnings percentile at age 25 by third-grade reading score percentile

Third-grade scores and future income

Not a causal relationship

- Many factors affect academic performance *and* income:
- class size, pedagogy, peers, technology, sleep, financial stress, home environment, nutrition, pollution, …

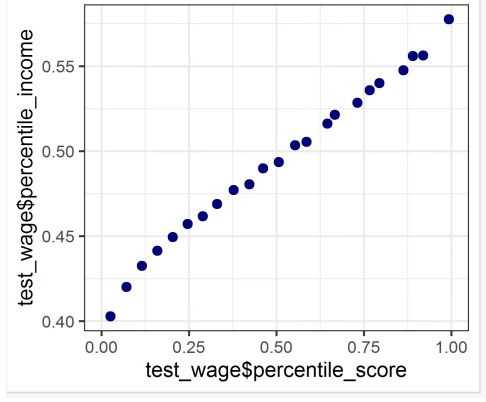


Authors' calculations from the Texas administrative data: mean earnings percentile at age 25 by third-grade reading score percentile

Third-grade scores and future income

Not a causal relationship

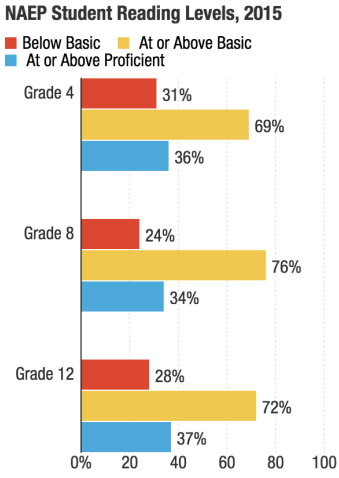
- Many factors affect academic performance *and* income:
- class size, pedagogy, peers, technology, sleep, financial stress, home environment, nutrition, pollution, …
- We exploit **exogenous movement along the line**: a (possibly) **permanent** shifter — **pedagogy**

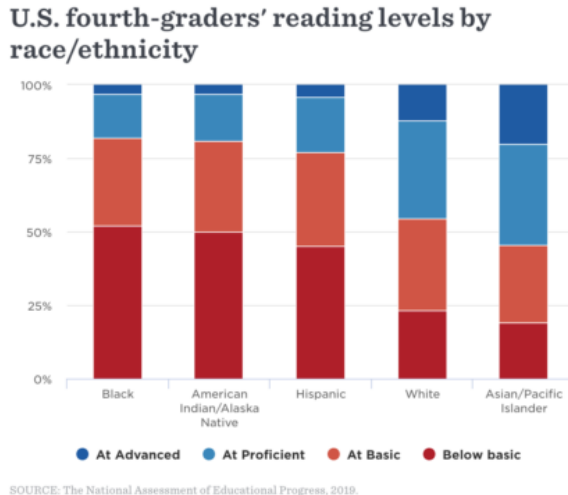


Authors' calculations from the Texas administrative data: mean earnings percentile at age 25 by third-grade reading score percentile

Despite all that spending…

We may not be doing so well





SOURCE: The National Assessment of Educational Progress, 2019.

Data: NCES, National Assessment of Educational Progress, reading 2015 (left) and grade-4 reading 2019 (right); charts reproduced from press coverage of the Nation's Report Card

One state is the exception

Bottom-decile 4th-grade NAEP reading, 2013–2024

- **Mississippi** mandated phonics-based “science of reading” instruction
- The **only state** whose struggling readers gained

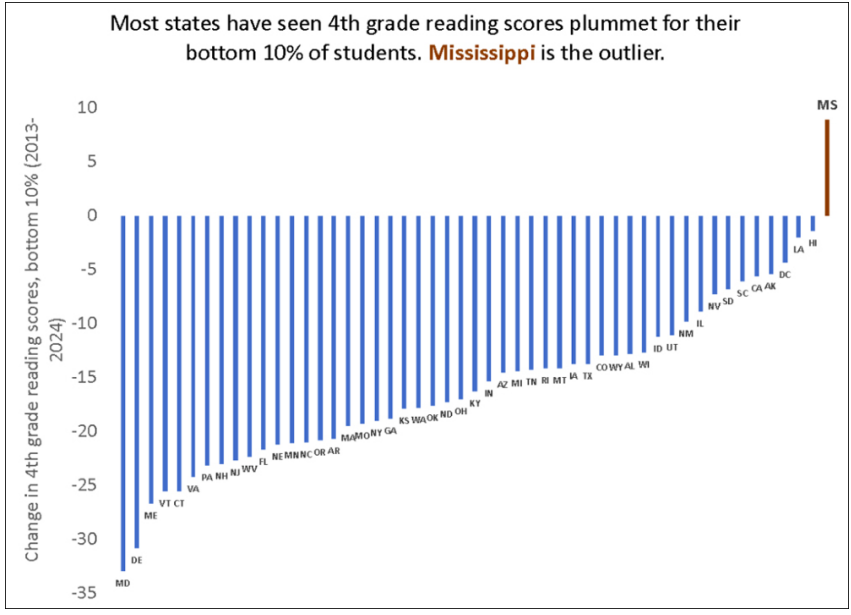


Chart: Chad Aldeman, “There Really Was a ‘Mississippi Miracle’ in Reading,” The 74, Feb. 25, 2025; data: NCES/NAEP

Reading (and especially phonics) is having a moment

Forty-plus states have passed science-of-reading laws since 2013













Clockwise from top left: National Review (Rich Lowry, Sept. 2025); Ohio Capital Journal (Megan Henry, Aug. 22, 2024); The 74 (Patrick O'Donnell, May 10, 2023); mythreereaders.com; Emily Hanford and the *Sold a Story* art, APM Reports (2022)

What is Reading First?

What do we mean by mode of instruction?

Hooked on Phonics

VS

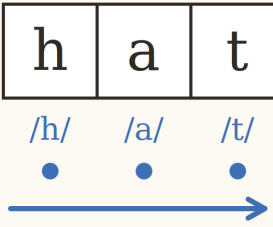


- **Phonics**: systematic letter–sound correspondence
- **Whole language**: meaning first; word identity from context (cueing)

Hooked on Phonics logo, trademark of its owner; photo: Marie Clay (Reading Recovery), 1974, via Wikipedia

Two theories of how children read

Two routes to the same word: decode it, or guess it from the picture



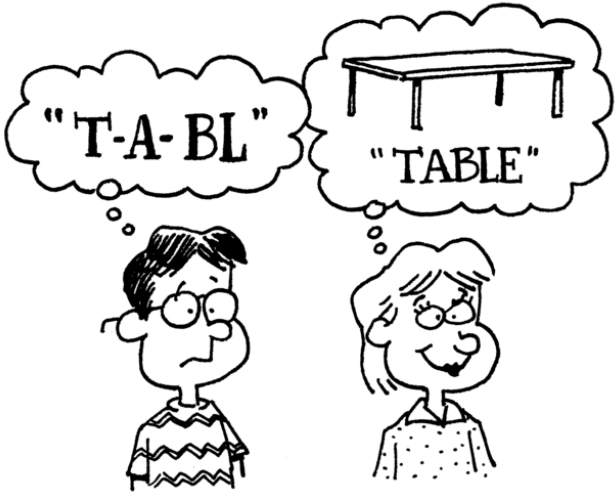


- **Phonics**: teach systematic letter–sound correspondence, so a child can sound out a word they have never seen
- **Whole language**: meaning first; infer an unfamiliar word from context, grammar, and pictures (“cueing”)

Illustrations: the authors'

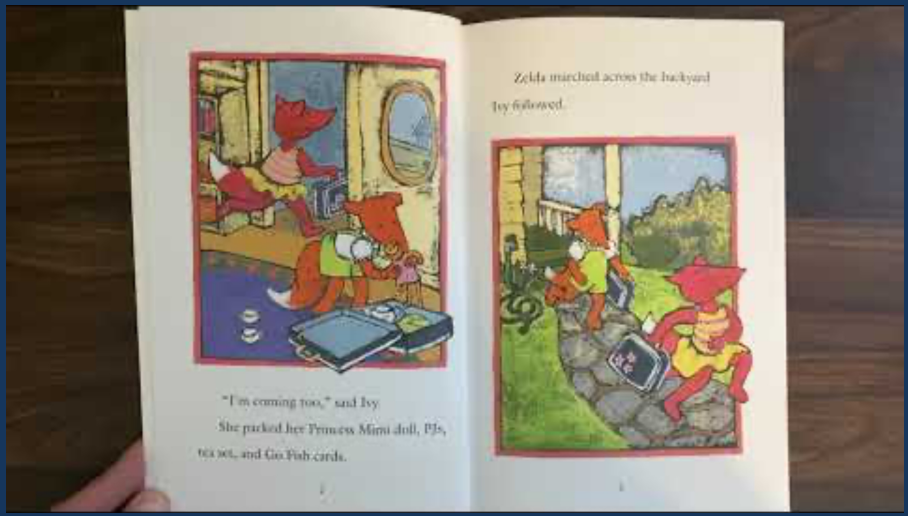
Decoding vs. recognizing

Same word, two theories of reading



Cartoon reproduced from the authors' earlier talks; original source not recorded

Whole-language approach uses cueing



Still from a classroom read-aloud video posted to YouTube (youtu.be/5wZ_bTlNXKs, 1:21); book shown: *Zelda and Ivy*, Laura McGee Kvasnosky (Candlewick)

Debates about how to teach reading are not new

- Early 1800s: ABC Method
- Mid-1800s: Horace Mann & John Dewey, "Whole Word Method" — phonics "dull and tedious drilling …destroying children's innate love of learning" (Kogan 2025)
- Cold War: *Why Johnny Can't Read* (1955) ⇒ spread of phonics
- 1980s: **"Whole Language"**

Spread of whole language in the late 1980s

> *"Pick your metaphor — an epidemic, wildfire, manna from heaven — whole language has spread so rapidly throughout North America that it is a fact of life in literacy curriculum and research."*

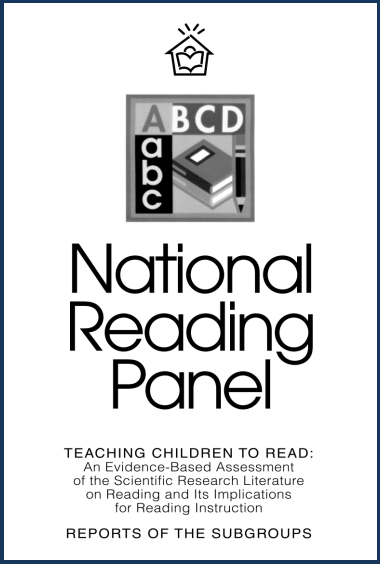
P. David Pearson, Dean, UC Berkeley School of Education, 1989 (via Kogan 2025)

Enter politics

- By 2000, phonics was clearly a Republican policy; whole language more embraced by Democrats
- Whole language seen as a “means for advancing a political agenda …the classroom as a model for an egalitarian society” (McKenna, Stahl & Reinking 1995)

Return of phonics

- Congress requests a study in 1997; National Reading Panel reports in 2000
- “Strong endorsement for systematic and explicit phonics instruction” (Kogan 2025)
- “Scientific” evidence favors phonics (lab studies, eye tracking, …)
- But **limited social-science research about policy**



Cover: National Reading Panel, *Teaching Children to Read* (NICHD, 2000)

Identifying variation via Reading First

- Federal grants to encourage phonics instruction: **$1 billion per year**
- Vary across states, schools, grades, and years
- State applications in April 2003

Reading First is a K–3 intervention

Materials

- Textbooks and reading materials that encourage phonics

Instruction

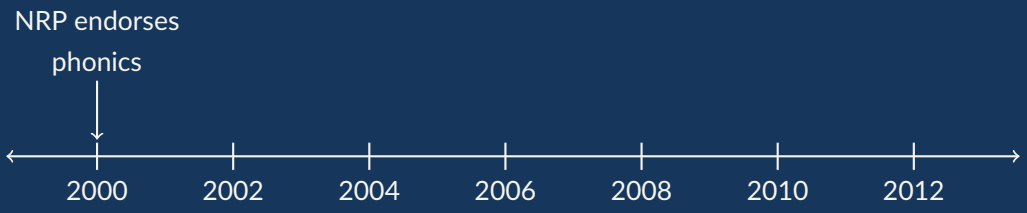
- Change classroom practice; diagnostic assessments

People

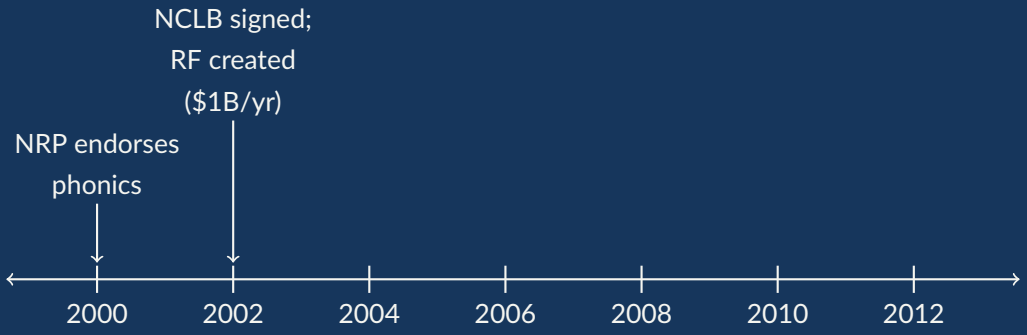
- Hiring more teachers
- Hiring reading coaches
- Training: seminars, conferences, …

The rare case where federal money arrived with a pedagogy attached. That condition is what we end up identifying.

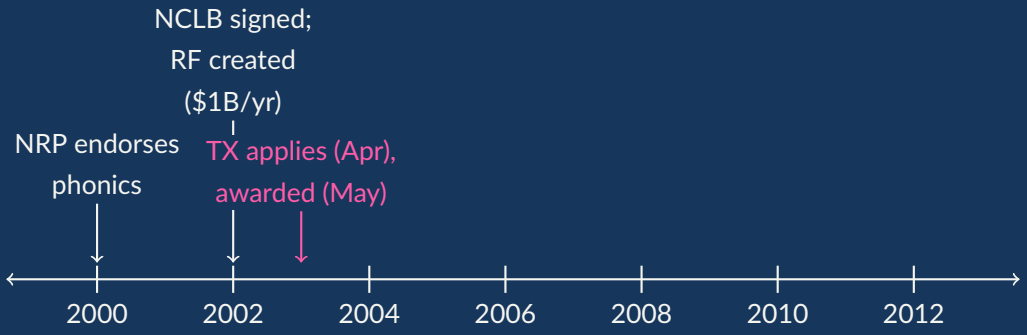
The Reading First timeline



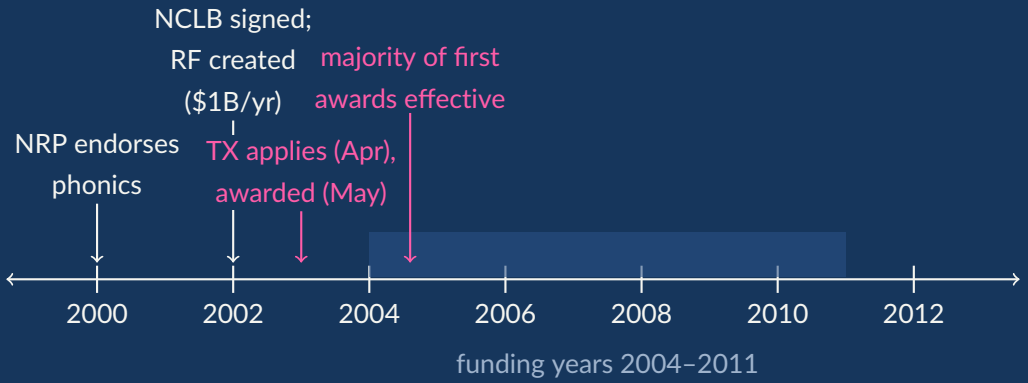
The Reading First timeline



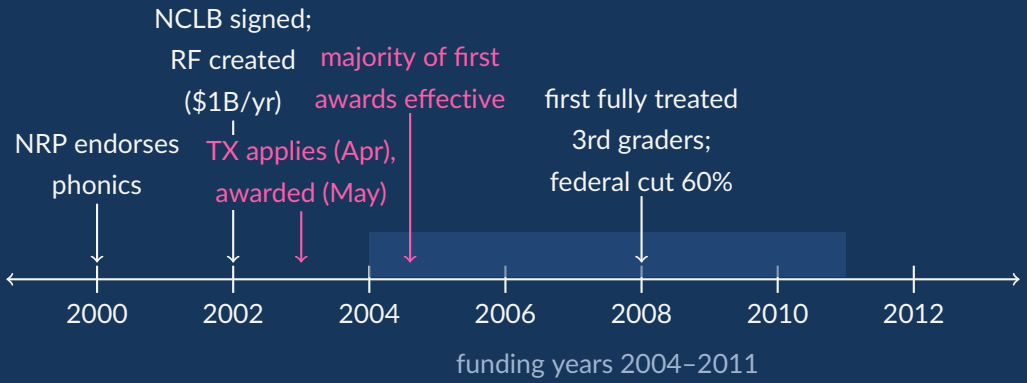
The Reading First timeline



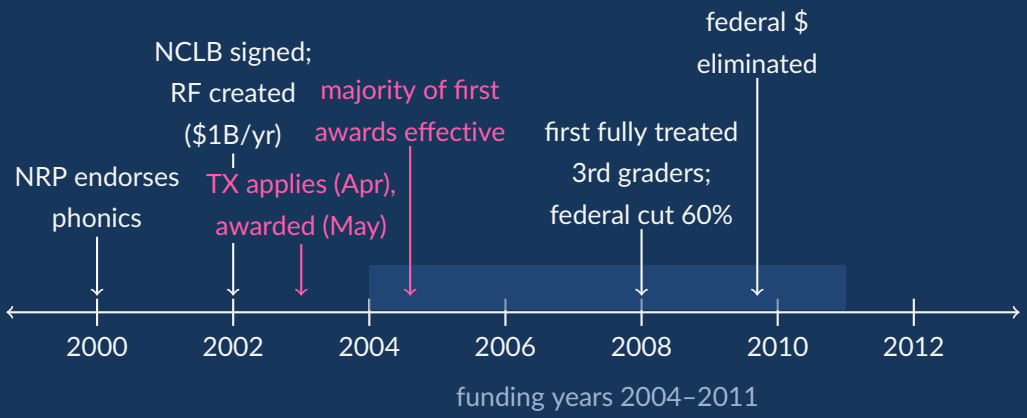
The Reading First timeline



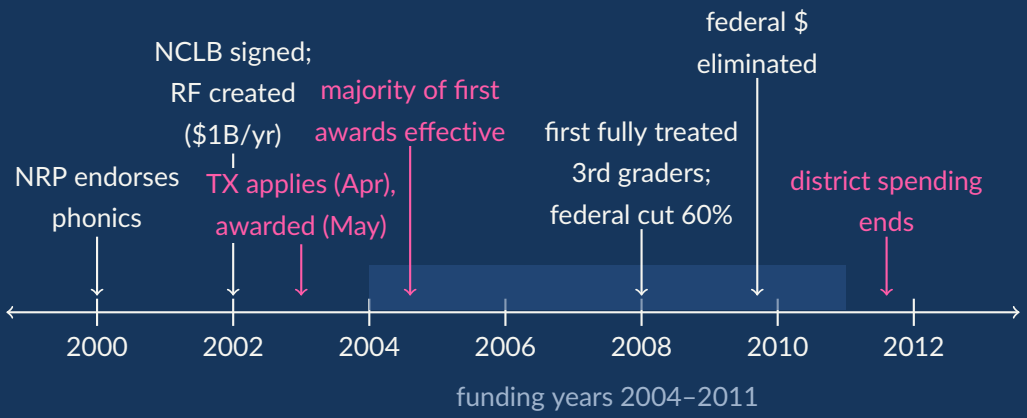
The Reading First timeline



The Reading First timeline



The Reading First timeline



Reading First in Texas

- Award notification May 2003; districts apply to the state
- **Strict eligibility criteria** (next slide)
- First awards summer 2003; majority of first awards in 2004–05
- Program “ends” in 2011: **over $630 million** in all, ∼$100 million per year at the peak, roughly 700 campuses

Eligibility & data

Reading First eligibility

Built from 2002, before the program

- Eligible if the district's **2002 grade-3 reading failure rate $\geq$ 13.6%** (the statewide average) *and* one of: empowerment zone; a Title I improvement campus; $\geq$15% or $>$6,500 Title I-eligible students

Reading First eligibility

Built from 2002, before the program

- Eligible if the district's **2002 grade-3 reading failure rate $\geq$ 13.6%** (the statewide average) *and* one of: empowerment zone; a Title I improvement campus; $\geq$15% or $>$6,500 Title I-eligible students
- Everything fixed at **2002 (pre-program) levels**:

$$\text{Eligible}_d = \mathbf{1}\left[\overline{\text{Fail}}_d^{2002} \geq 13.6\%\right] \times \mathbf{1}[\text{need criterion}_d]$$

Reading First eligibility

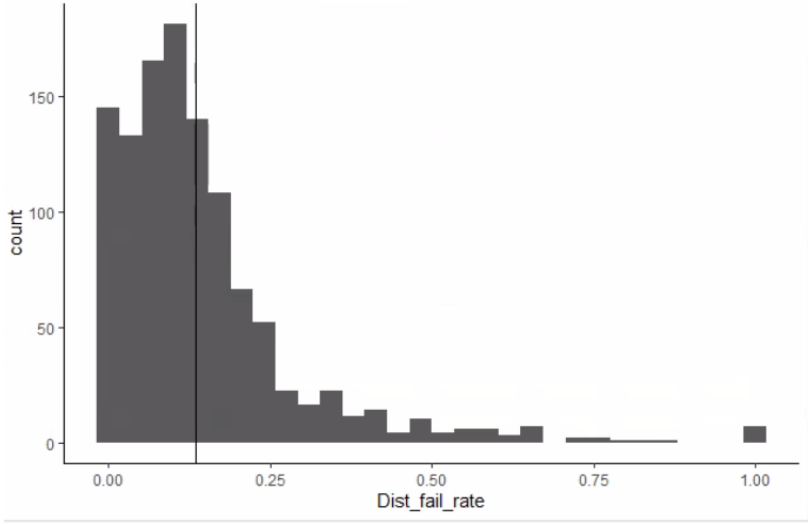
Built from 2002, before the program

- Eligible if the district's **2002 grade-3 reading failure rate $\geq$ 13.6%** (the statewide average) *and* one of: empowerment zone; a Title I improvement campus; $\geq$15% or $>$6,500 Title I-eligible students
- Everything fixed at **2002 (pre-program) levels**:

$$\text{Eligible}_d = \mathbf{1}\left[\overline{\text{Fail}}_d^{2002} \geq 13.6\%\right] \times \mathbf{1}[\text{need criterion}_d]$$

- Predicted grant: $\text{Grant}_{dt} = \max\{\$170\text{k},\ \text{RF funds}_t \times \text{Title I share}_{d,2002}\} \times \text{Eligible}_d$ — **our instrument**

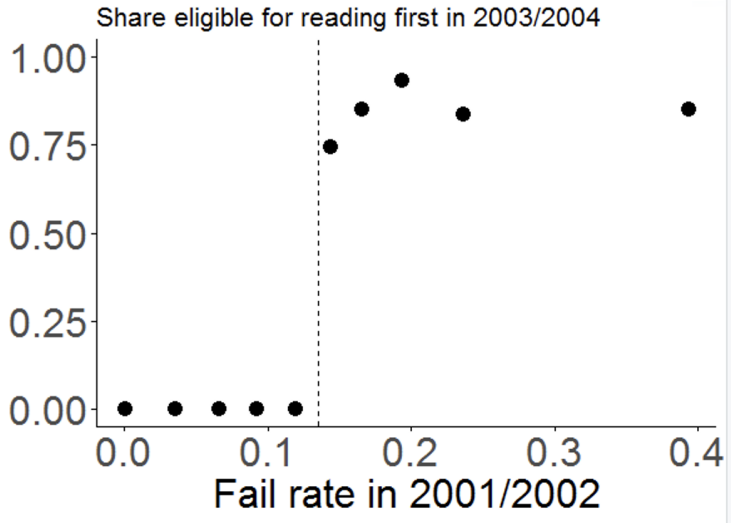
Failure rate around the 13.6% threshold



Authors' calculations from Texas Education Agency data

Eligibility jumps at the threshold

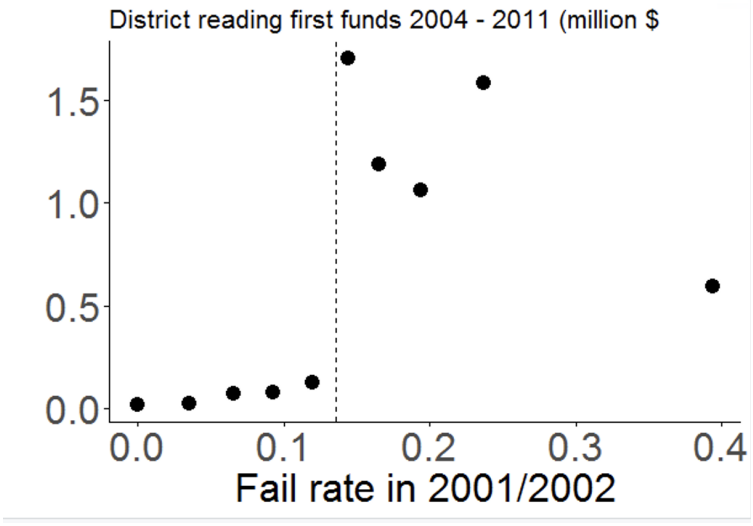
Share of districts receiving RF, by 2002 fail rate



Authors' calculations from Texas Education Agency data

And so do Reading First dollars

District RF funds 2004–2011, by 2002 fail rate



Authors' calculations from Texas Education Agency data

Wayback machine data + PIR + FOIA

Cycle-1 awardee lists, award amounts, eligibility rules

CYCLE 1 YEAR 1 TEXAS READING FIRST AWARDEES-2003-2004

No.	Applicant Agency	Co. Dist.	Region	Amt Awarded
1	Eagle Pass ISD	159-901	20	$815,000
2	Port Arthur ISD	123-907	05	$582,183
3	Kermit ISD	248-901	18	$218,194
4	Mount Calm ISD	109-910	12	$170,000
5	Bexar County Academy	015-809	20	$170,000
6	Taneha ISD	219-904	07	$217,020

Texas Education Agency, "Cycle 1 Year 1 Texas Reading First Awardees 2003–2004," retrieved through the Internet Archive Wayback Machine

Administrative data

Texas, late 1980s to today; best data start 1994

Test scores

- Stable student ID across time; exact school
- Reading/math scores, grades 3–8
- **Every question** on 13 reading and 13 math grade-3 exams, digitized

Later life and money

- High school completion; college enrollment and degrees
- Quarterly job-level wages + NAICS industry
- District finances (PEIMS), incl. a dedicated RF fund; tax elections 2000–2019

Linked: third-grade cohorts followed all the way to age-25 earnings.

Design

Simple analyses

Average standardized third-grade scores, campus $\times$ year

$$y_{ct} = \beta \underbrace{\mathbf{1}[\text{Eligible}_c] \times \mathbf{1}[t > 2003]}_{\text{DiD term}} + \gamma_c + \delta_t + \phi\left(\text{ED share}_{c,2002} \times \delta_t\right) + \varepsilon_{ct}$$

Simple analyses

Average standardized third-grade scores, campus $\times$ year

$$y_{ct} = \beta \underbrace{\mathbf{1}[\text{Eligible}_c] \times \mathbf{1}[t > 2003]}_{\text{DiD term}} + \gamma_c + \delta_t + \phi\left(\text{ED share}_{c,2002} \times \delta_t\right) + \varepsilon_{ct}$$

- **First-time test takers** (baseline; Texas tied grade-3 retention to the reading test from 2003), standardized within test-grade-year

Simple analyses

Average standardized third-grade scores, campus × year

$$y_{ct} = \beta \underbrace{\mathbf{1}[\text{Eligible}_c] \times \mathbf{1}[t > 2003]}_{\text{DiD term}} + \gamma_c + \delta_t + \phi\left(\text{ED share}_{c,2002} \times \delta_t\right) + \varepsilon_{ct}$$

- **First-time test takers** (baseline; Texas tied grade-3 retention to the reading test from 2003), standardized within test-grade-year
- **Balanced campus panel 1995–2017**; campus and year FE **+ 2002 econ-disadvantaged share × year** controls; SEs clustered by district; weighted by test takers

Simple analyses

Average standardized third-grade scores, campus × year

$$y_{ct} = \beta \underbrace{\mathbf{1}[\text{Eligible}_c] \times \mathbf{1}[t > 2003]}_{\text{DiD term}} + \gamma_c + \delta_t + \phi\left(\text{ED share}_{c,2002} \times \delta_t\right) + \varepsilon_{ct}$$

- **First-time test takers** (baseline; Texas tied grade-3 retention to the reading test from 2003), standardized within test-grade-year
- **Balanced campus panel 1995–2017**; campus and year FE **+ 2002 econ-disadvantaged share × year** controls; SEs clustered by district; weighted by test takers
- Event studies use **2003 as the reference year** (spring-2003 exams predate any disbursement)

Simple analyses

Average standardized third-grade scores, campus × year

$$y_{ct} = \beta \underbrace{\mathbf{1}[\text{Eligible}_c] \times \mathbf{1}[t > 2003]}_{\text{DiD term}} + \gamma_c + \delta_t + \phi\left(\text{ED share}_{c,2002} \times \delta_t\right) + \varepsilon_{ct}$$

- **First-time test takers** (baseline; Texas tied grade-3 retention to the reading test from 2003), standardized within test-grade-year
- **Balanced campus panel 1995–2017**; campus and year FE **+ 2002 econ-disadvantaged share × year** controls; SEs clustered by district; weighted by test takers
- Event studies use **2003 as the reference year** (spring-2003 exams predate any disbursement)
- **IV**: cumulative RF dollars instrumented by cumulative formula-predicted dollars; reported **per $20 million**, about what the average eligible student's district received ($18M)

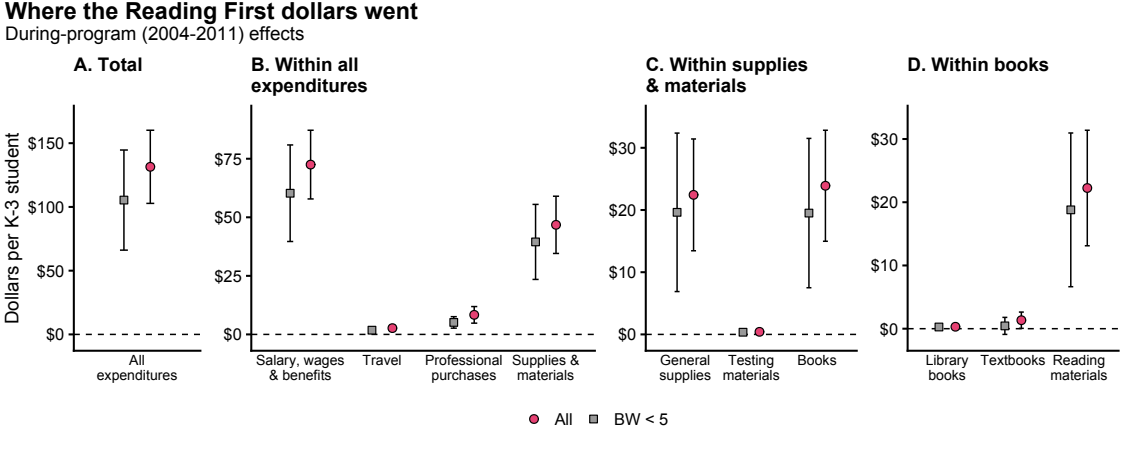
First stage, district-level finances

Revenue per K–3 student; first stage $1.41 per predicted dollar, F = 1,688

	Full sample	Bandwidth < 5
	(1)	(2)
A. Revenue per K–3 student, DID		
1(RF Eligible) × 1(year>2003)	74.75***	58.85***
	(8.51)	(11.49)
B. Revenue per K–3 student, separate periods		
1(RF Eligible) × 1(2004–2011)	132.27***	105.51***
	(14.77)	(20.20)
1(RF Eligible) × 1(2012 and after)	1.57	0.01
	(1.47)	(1.69)
C. First stage, cumulative RF revenue		
Cumulative predicted RF $ ($ millions)	1.41***	1.40***
	(0.03)	(0.04)
First-stage F stat	1,687.92	1,158.48
Observations	23,831	8,628
District FE	Yes	Yes
Year FE	Yes	Yes
Weights	K–3 students	K–3 students

Money was spent on instruction and reading materials

During-program effects per K–3 student: $131 in all, $73 payroll, $47 supplies, of which $22 reading materials



Should we expect to find an effect on reading?

Money didn't have an immediate effect



Reading First report big topic at DISD board meeting

By Elizabeth Craft
Anvil Herald Staff

Tricia Perez, a TEA technical assistant, gave a report on the Reading First program in D'Hanis. Perez reported how well the student are doing.

Reading First is an initiative that will ensure a student can read at his or her grade level or above by the third grade.

The premise behind this program is that all but a small handful of children can be taught to read successfully, prevention of reading problems is far more cost effective and efficient that remediation of an existing problem, and reading failure can be prevented by relying on the research base in readi.g.

The Texas Reading First Initiative includes these key elements:

• Reading instruction and reading programs that are grounded in scientifically-based reading research.

• Assessment driven instruction to be able to identify struggling readers, plan instruction, and provide early intervention.

• Ongoing professional development with follow-up support through campus coaching and regional technical assistance.

There are five essential components of reading: phonemic awareness, phonics, fluency, vocabulary, and finally comprehension.

In D'Hanis the percentage of students meeting proficiency requirements went up from the 2004- 2005 school year to the 2005-2006.

In 1st grade the percentage went up from 67% to 91%. 2nd grade went up from 81% to 91%. 3rd grade, however, went down from 91% to 83%. The reason for the decrease is because the 3rd graders didn't have a consistent program from kindergarten to 3rd. In 2003-2004 D'Hanis received the grant for the Reading First, but the children could not be tested yet. In 2004-2005 the school got the required materials and started to use them. The teachers had to learn the new model and implement it.. On top of all that, the 3rd grade suffered the temporary loss of their teacher due to illness and they had a substitute for about half the school year. The expectation was to reach the 40th percentile. They far exceeded that number. The average for other state schools was 55% for kinder and 50% for 1st grade. The state was unable to get an average for 3rd grade. The 3rd graders also had to begin TAKS testing and the theory was that the 3rd grade numbers across the state would go down.

Pam Siepp, D'Hanis superintendent, said the school will follow up in all grades with consistency of the learning materials.

After the Reading First report Lana Rollans gave an update on the planning process for the technology plan.

Administrators Dorothy Saunders, elementary principal, and Todd Craft, athletic director, gave details of fall activities and calendar items.

Other items on the agenda were: monthly bills and expenditures, tax collection report, and quarterly investment report.

Pam Seipp, D'Hanis ISD superintendent, reported on assessing enrollment, considering annual TEA monitoring, evaluating formative administrators appraisals, and reviewing "facility needs" research.

The business items for the meeting were:

• Approval of the facility upgrade project and begin advertisement for bids to air condition the Roy Kile Gymnasium.

• Approval of the board planning policies (legal and local BQ series) as currently written and found in the Board Policy Manual.

• Approval of the district Textbook Committee to be served by teachers from the District Improvement Team: Barbara Pacini, Lora Garrison, Jim McCleery, Troy Langfeld, Lance Montgomery, and Jaime Whitwell.

Sunday Buffet $8.99
Seniors $7.99, Youth (9-12) $5.99, Kids $3.99, 2 & under Free!
Fried Catfish and Roast Beef
We Cater Parties & Provide Party Trays
Call Us for Prices & Available Dates!
THE FLIGHTLINE CAFE
426-4020 • Hwy. 90 in Hondo • Landmark Shopping Center

Grand Opening Weekend!
The Rustic Courtyard
Ribbon Cutting: Oct. 20th • 5 p.m. w/ refreshments
Open House: October 21st • 3 p.m. - 7 p.m.
Floral Design ~ Antiques ~ Garden Center ~ Gifts
Home Decor ~ Event Planning ~ Handmade Items

Hondo Anvil Herald (Hondo, Tex.), Oct. 19, 2006, p. 3A, "Reading First report big topic at DISD board meeting," by Elizabeth Craft; via The Portal to Texas History, UNT Libraries

Money didn't have an immediate effect



Hondo Anvil Herald • Thursday, October 19, 2006 • 3A

Reading First report big topic at DISD board meeting

By Elizabeth Craft
Anvil Herald Staff

Tricia Perez, a TEA technical assistant, gave a report on the Reading First program in D'Hanis. Perez reported how well the student are doing.

Reading First is an initiative that will ensure a student can read at his or her grade level or above by the third grade.

The premise behind this program is that all but a small handful of children can be taught to read successfully, prevention of reading problems is far more cost effective and efficient that remediation of an existing problem, and reading failure can be prevented by relying on the research base in readi..g.

The Texas Reading First Initiative includes these key elements:

• Reading instruction and reading programs that are grounded in scientifically-based reading research.

• Assessment driven instruction to be able to identify struggling readers, plan

instruction, and
intervention.

• Ongoing pro
velopment with
support throu
coaching and reg
cal assistance.

There are fi
components of r
nemic awarene
fluency, vocabu
nally comprehen

In D'Hanis the
of students me
ciency requirem
from the 2004-
year to the 2005-

In 1st grade th
went up from 67%
grade went up f
91%. 3rd grade, h
down from 91%
reason for the de
cause the 3rd gr
have a consiste
from kindergarte
2003-2004 D'Ha
the grant for the R
but the children
tested yet. In 2004-2005 the
school got the required mate-
rials and started to use them.
The teachers had to learn the

pal, and Todd Craft, athletic
director, gave details of fall
activities and calendar items.

Other items on the agenda

son, Jim McCleery, Troy
Langfeld, Lance Montgom-
ery, and Jaime Whitwell.

from kindergarten to 3rd. In 2003-2004 D'Hanis received the grant for the Reading First, but the children could not be tested yet. In 2004-2005 the school got the required materials and started to use them. The teachers had to learn the

new model and implement it.

Sunday Buffet $8.99
Fried Catfish
and
Roast Beef
Cater Parties & Provide Party Trays
Call Us for Prices & Available Dates!
THE FLIGHTLINE CAFE
Hwy. 90 in Hondo • Landmark Shopping Center

Grand Opening Weekend!
The Rustic Courtyard
Ribbon Cutting: Oct. 20th • 5 p.m. w/ refreshments
Open House: October 21st • 3 p.m. - 7 p.m.
Floral Design ~ Antiques ~ Garden Center ~ Gifts

Hondo Anvil Herald (Hondo, Tex.), Oct. 19, 2006, p. 3A, "Reading First report big topic at DISD board meeting," by Elizabeth Craft; via The Portal to Texas History, UNT Libraries

Hiring continued even after the “first” dollars

Reading coaches, 2006



Too Late To Classify

1960 FORD TRACTOR - Excellent condition overall, completely overhauled, 5 ft. shredder & blade. 210-684-4678. 05-11-1tv

D'HANIS ISD - Is seeking a Full-Time Elementary Reading First Coach. Qualifications may be seen on TASANET.ORG. Applicants may call 830-363-7215 for an application, or email request to bvoigt@dhanisisd.net. 05-11-2tc

PART-TIME SITTER NEEDED - Hondo, after school and Saturdays. References req'd. Call 210-324-3242. 05-11-4tp

PRODUCTION POSITIONS - Crane Plumbing is now accept-

TEXAS ST

TexSCAN Week of May 7, 2006

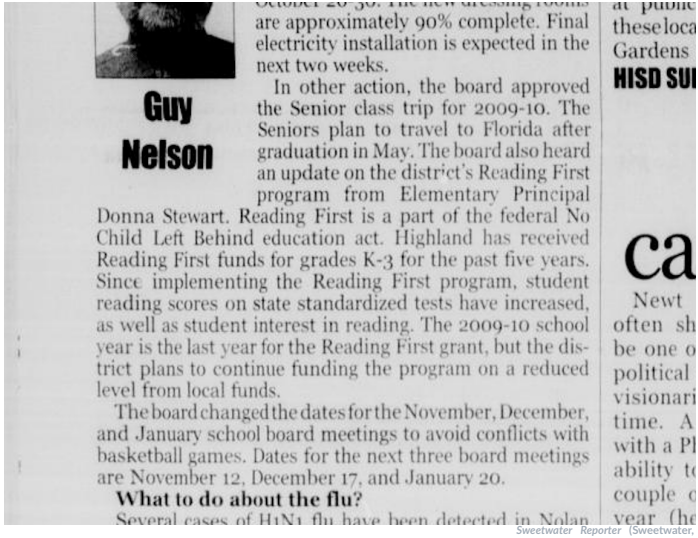
ADOPTION

Note: It is illegal to be paid for anything beyond medical and legal expenses in Texas adoption.

PREGNANT? CONSIDERING

Classified advertisement, *Hondo Anvil Herald* (Hondo, Tex.), May 11, 2006; via The Portal to Texas History, UNT Libraries

Curricula changes outlived the federal program



Guy Nelson

are approximately 90% complete. Final electricity installation is expected in the next two weeks.

In other action, the board approved the Senior class trip for 2009-10. The Seniors plan to travel to Florida after graduation in May. The board also heard an update on the district's Reading First program from Elementary Principal Donna Stewart. Reading First is a part of the federal No Child Left Behind education act. Highland has received Reading First funds for grades K-3 for the past five years. Since implementing the Reading First program, student reading scores on state standardized tests have increased, as well as student interest in reading. The 2009-10 school year is the last year for the Reading First grant, but the district plans to continue funding the program on a reduced level from local funds.

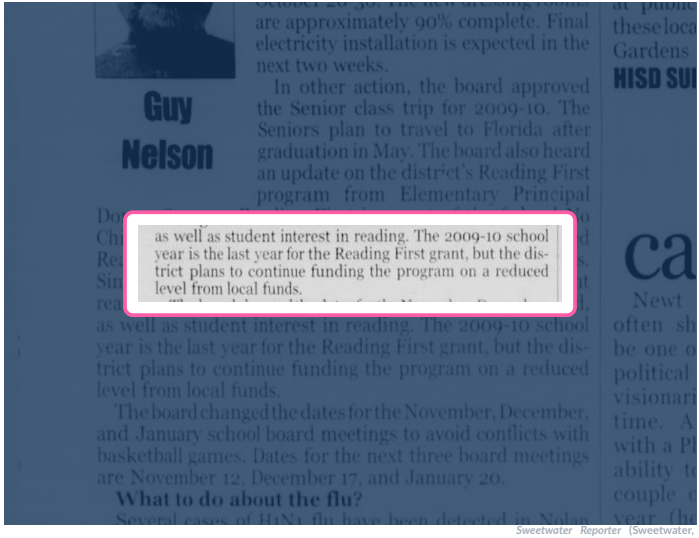
The board changed the dates for the November, December, and January school board meetings to avoid conflicts with basketball games. Dates for the next three board meetings are November 12, December 17, and January 20.

What to do about the flu?

Several cases of H1N1 flu have been detected in Nolan

Sweetwater Reporter (Sweetwater, Tex.), Highland ISD column by superintendent Guy Nelson, fall 2009; via The Portal to Texas History, UNT Libraries

Curricula changes outlived the federal program



Guy Nelson

as well as student interest in reading. The 2009-10 school year is the last year for the Reading First grant, but the district plans to continue funding the program on a reduced level from local funds.

are approximately 90% complete. Final electricity installation is expected in the next two weeks.

In other action, the board approved the Senior class trip for 2009-10. The Seniors plan to travel to Florida after graduation in May. The board also heard an update on the district's Reading First program from Elementary Principal

as well as student interest in reading. The 2009-10 school year is the last year for the Reading First grant, but the district plans to continue funding the program on a reduced level from local funds.

The board changed the dates for the November, December, and January school board meetings to avoid conflicts with basketball games. Dates for the next three board meetings are November 12, December 17, and January 20.

What to do about the flu?

Sweetwater Reporter (Sweetwater, Tex.), Highland ISD column by superintendent Guy Nelson, fall 2009; via The Portal to Texas History, UNT Libraries

Prior evidence found nothing

2008 federal impact study

- 2004–07 data; 248 schools in 17 districts across 13 states
- By 2007, **zero** cohorts had completed a fully treated K–3 program
- No significant change in reading scores…
- …but **strong evidence on mechanisms**: decoding, phonics instruction, training and practice all moved


Reading First Impact Study
Final Report

NOVEMBER 2008

Beth C. Gamse, Project Director, Abt Associates
Robin Tepper Jacob, Abt Associates/University of Michigan
Megan Horst, Abt Associates
Beth Boulay, Abt Associates
Fatih Unlu, Abt Associates

Cover: Gamse, Jacob, Horst, Boulay and Unlu, *Reading First Impact Study: Final Report* (NCEE 2009-4038), IES, U.S. Department of Education, Nov. 2008

2008 report finds no evidence of a treatment effect

Changes in reading scores?

No.

Exhibit 2.5: Estimated Impacts on Reading Comprehension: Spring 2005, 2006, and 2007 (Pooled)

Construct	Actual Mean with Reading First	Estimated Mean without Reading First	Impact	Effect Size of Impact	Statistical Significance of Impact (p-value)
Panel 1					
All Sites					
Reading Comprehension Scaled Score					
Grade 1					
Scaled Score	543.8	539.1	4.7	0.10	(0.083)
Corresponding Grade Equivalent[a]	1.7	1.7			
Corresponding Percentile	*44*	*41*			
Grade 2					
Scaled Score	584.4	582.8	1.7	0.04	(0.462)
Corresponding Grade Equivalent[a]	*2.5*	2.4			
Corresponding Percentile	*39*	*38*			
Grade 3					
Scaled Score	609.1	608.8	0.3	0.01	(0.887)
Corresponding Grade Equivalent[a]	*3.3*	*3.3*			
Corresponding Percentile	*39*	*39*			
Panel 2					
All Sites					
Percent Reading At or Above Grade Level[b]					
Grade 1	46.0	41.8	4.2		(0.104)
Grade 2	38.9	37.3	1.6		(0.504)
Grade 3	38.7	38.8	-0.1		(0.973)

NOTES:

The complete Reading First Impact Study sample includes 248 schools from 18 sites (17 school districts and 1 state) located in 13 states. 125 schools are Reading First schools and 123 are non-Reading First schools. For grade 2 in 2006, one non-RF school could not be included in the analysis because test score data were not available. For grade 3 in 2007, one RF school could not be included in the analysis because test score data were not available.

Exhibit 2.5 of Gamse, Jacob, Horst, Boulay and Unlu, *Reading First Impact Study: Final Report* (NCEE 2009-4038), IES, U.S. Department of Education, Nov. 2008

2008 report finds lots of evidence supporting mechanisms

Changes in students' ability to **decode**

Exhibit 2.6: Estimated Impacts of Reading First on Decoding Skill: Grade One, Spring 2007

	Actual Mean with Reading First	Estimated Mean without Reading First	Impact	Effect Size of Impact	Statistical Significance of Impact (p-value)
Decoding Skill					
Standard Score	96.9	94.4	2.5 *	0.17 *	(0.025)
Corresponding Grade Equivalent[a]	1.7	1.4			
Corresponding Percentile	42	35			

NOTES:

The Test of Silent Word Reading Fluency (TOSWRF) sample includes first-graders in 248 schools from 18 sites (17 school districts and 1 state) located in 13 states. 125 schools are Reading First schools and 123 are non-Reading First schools.

Exhibit 2.6 of Gamse, Jacob, Horst, Boulay and Unlu, *Reading First Impact Study: Final Report* (NCEE 2009-4038), IES, U.S. Department of Education, Nov. 2008

2008 report finds lots of evidence supporting mechanisms

Changes in "phonics" instruction

Exhibit 2.2: Estimated Impacts On the Number of Minutes in Instruction in Each of the Five Dimensions of Reading: 2005, 2006, and 2007 (pooled)

Number of minutes of instruction in:	Actual Mean With Reading First	Estimated Mean Without Reading First	Impact	Effect Size of Impact	Statistical Significance of Impact (p-value)
Grade 1					
Phonemic Awareness	2.32	1.71	0.61*	0.23*	(0.030)
Phonics	21.32	18.45	2.86*	0.21*	(0.048)
Vocabulary	7.92	7.35	0.57	0.09	(0.386)
Fluency	4.67	3.43	1.24*	0.20*	(0.043)
Comprehension	23.01	21.23	1.78	0.12	(0.247)
Grade 2					
Phonemic Awareness	0.49	0.37	0.12	0.10	(0.319)
Phonics	13.92	10.65	3.27*	0.31*	(0.006)
Vocabulary	11.79	10.06	1.73*	0.20*	(0.036)
Fluency	4.14	3.56	0.58	0.11	(0.297)
Comprehension	28.74	24.73	4.01*	0.24*	(0.019)

NOTES:

The complete Reading First Impact Study sample includes 248 schools from 18 sites (17 districts and 1 state) located in 13 states. 125 schools are Reading First schools and 123 are non-Reading First schools.

Exhibit 2.2 of Gamse, Jacob, Horst, Boulay and Unlu, *Reading First Impact Study: Final Report* (NCEE 2009-4038), IES, U.S. Department of Education, Nov. 2008

2008 report finds lots of evidence supporting mechanisms

Changes in **training and teaching practices**

Exhibit ES.4: Estimated Impacts on Key Components of Scientifically Based Reading Instruction (SBRI): Spring 2007

Domain	Actual Mean With Reading First	Estimated Mean Without Reading First	Impact	Effect Size of Impact	Statistical Significance of Impact (p-value)
Professional Development (PD) in SBRI					
Amount of PD in reading received by teachers (hours) [a]	25.84	13.71	12.13*	0.51*	(<0.001)
Teacher receipt of PD in the five essential components of reading instruction (0-5) [a]	4.30	3.75	0.55*	0.31*	(0.010)
Teacher receipt of coaching (proportion) [a]	0.83	0.63	0.20*	0.41*	(<0.001)
Amount of time dedicated to serving as K-3 reading coach (percent) [b,c]	91.06	57.57	33.49*	1.03*	(<0.001)
Amount of Reading Instruction					
Minutes of reading instruction per day [a]	105.71	87.24	18.47*	0.63*	(<0.001)
Supports for Struggling Readers					
Availability of differentiated instructional materials for struggling readers (proportion) [b]	0.98	0.97	0.01	0.15	(0.661)
Provision of extra classroom practice for struggling readers (0-4) [a]	3.79	3.59	0.19*	0.20*	(0.018)
Use of Assessments					
Use of assessments to inform classroom practice (0-3) [a]	2.63	2.45	0.18	0.19	(0.090)

Exhibit ES.4 of Gamse, Jacob, Horst, Boulay and Unlu, *Reading First Impact Study: Final Report* (NCEE 2009-4038), IES, U.S. Department of Education, Nov. 2008

Why we find effects where the 2008 report could not

- RF is a **K–3 intervention** with first funds mid-2003: almost no 2007 third grader was fully treated; fully treated classrooms only arrive in **2008**
- Money and implementation were slow; curricula changes continued after the federal program stopped
- Our own event study agrees: in the report's 2004–07 window the eligible–ineligible gap is only **0.03–0.05 s.d.** and still rising; it averages **0.07 s.d.** while money flowed (2004–11) and **0.16 s.d.** after it stopped (2012–17)

Other evidence

Mostly qualitative — and very influential

- No large-scale study finds meaningful Reading First effects
- **Huge** media attention (APM's *Sold a Story*)
- Recent work on a California block grant (75 schools) finds score gains — money and curriculum arrive together right during COVID

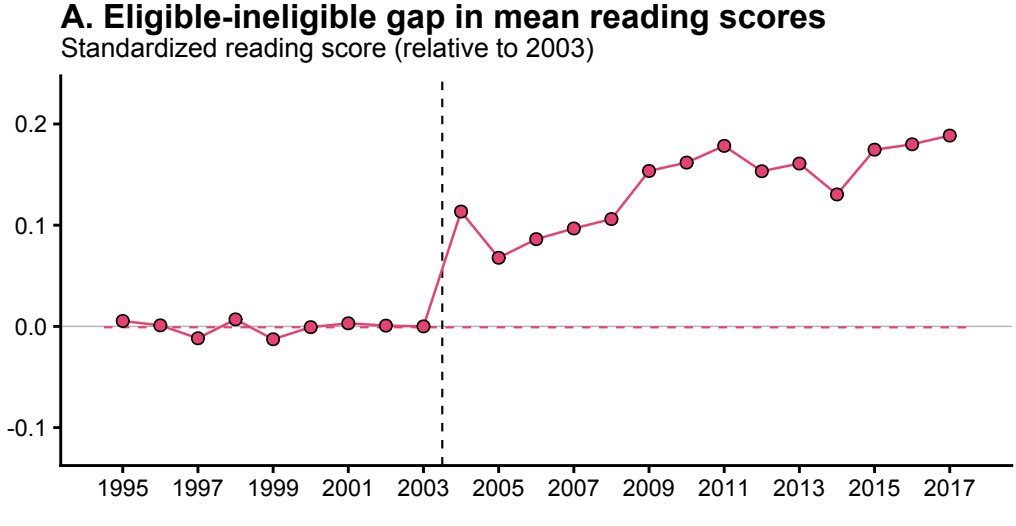


Screenshot: APM Reports, *Sold a Story*, Emily Hanford (2022), features.apmreports.org/sold-a-story

Effects on reading

Treated third-grade classrooms improve

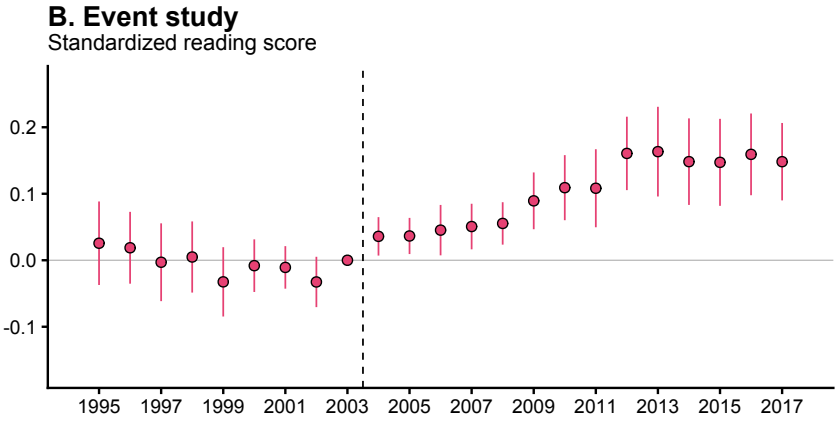
Eligible–ineligible gap in mean reading scores, relative to 2003



Event study: reading

Standardized score, all districts

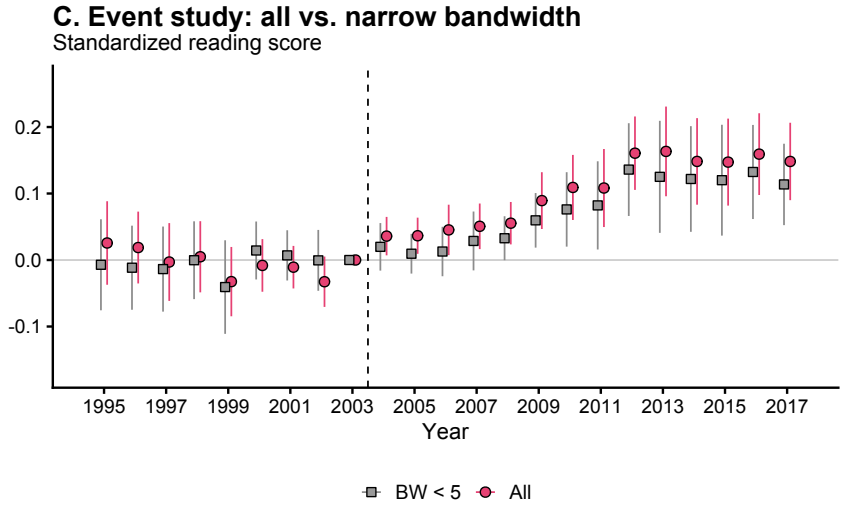
- No pre-trends
- Break at 2004, grows, then **stays** after the money stops
- Pooled DiD: **0.110 s.d.** (SE 0.022); 0.07 during 2004–11, 0.16 after



Event study: reading

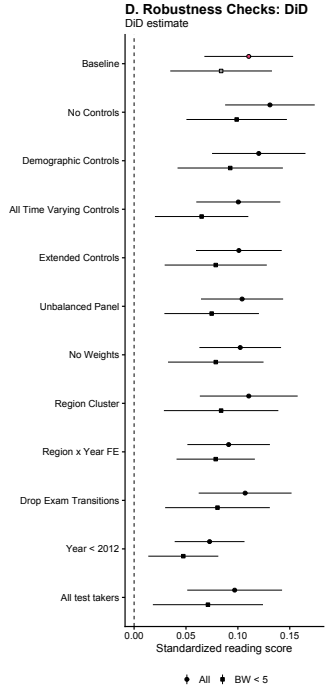
All vs. bandwidth < 5

- Same pattern inside the eligibility bandwidth
- In-bandwidth pooled DiD: **0.084 s.d.** (SE 0.025)



Robustness: DiD across specifications

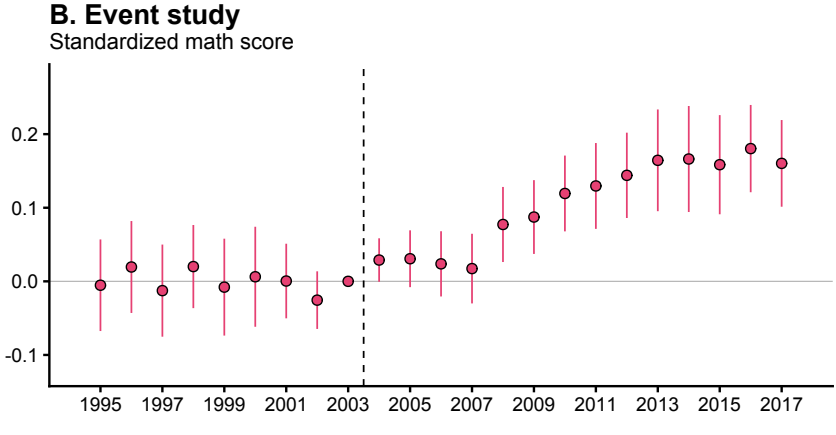
- Controls, weights, clustering, panel and sample choices
- **Positive and similar** in every one



Math moves too

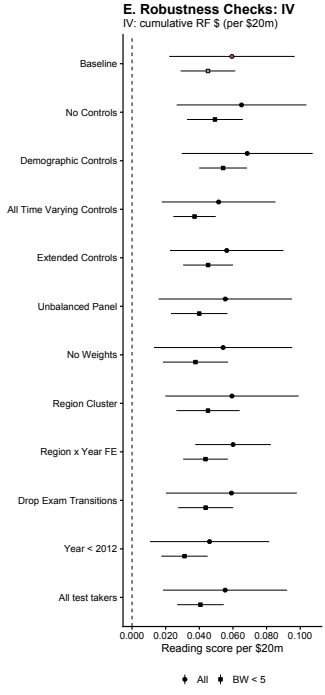
Event study: standardized math score

- Pooled DiD: **0.109 s.d.** (SE 0.028)
- Almost identical to reading
- Persists across grades at the student level (later)



IV: reading score per RF dollar

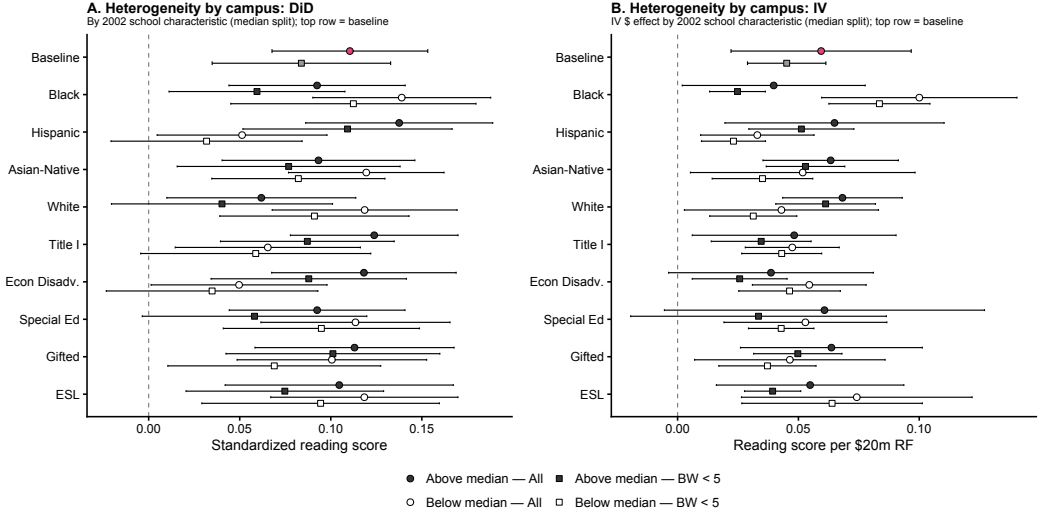
- Cumulative RF $, instrumented by formula-predicted $
- **0.059 s.d. per $20M** (SE 0.019)
- First-stage *F* = 1,711
- Stable across specifications



Heterogeneity & persistence

Who gains? Campuses

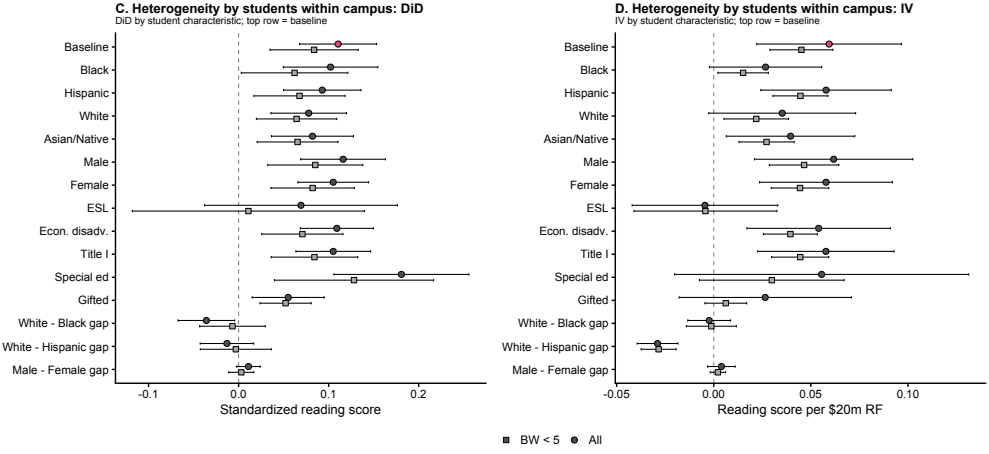
Above-median Hispanic, Title I, and econ.-disadvantaged shares gain most



Formula ties dollars to Title I share: dose vs. dose-response not separable.

Who gains? Students

Positive for every group; largest for special-education, econ.-disadvantaged, and Title I students

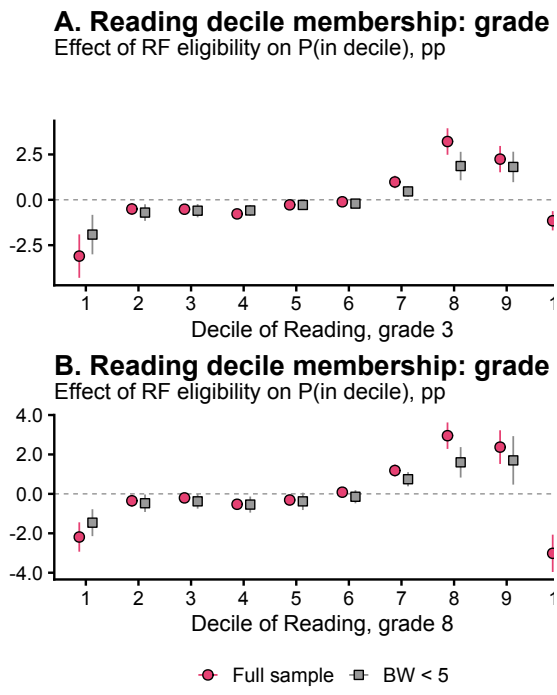


Suggestive narrowing of the White–Black gap (DiD) and the White–Hispanic gap (IV).

Where in the distribution?

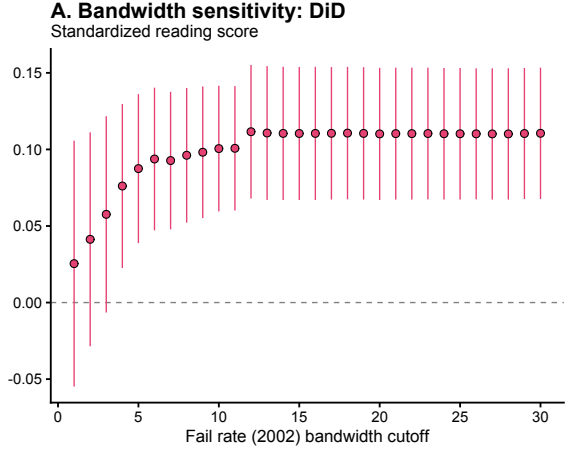
Moved struggling readers out of the bottom

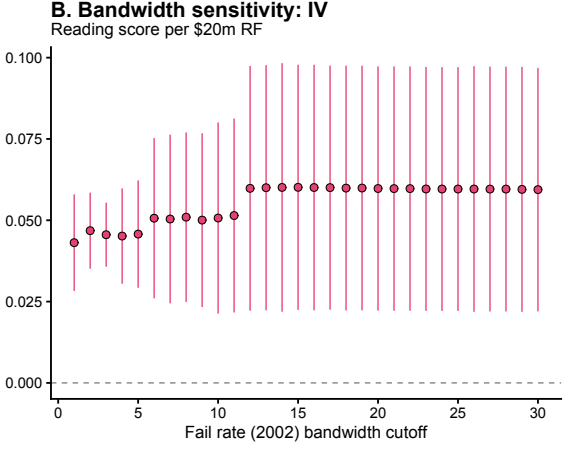
- Lowest decile: **−3 pp**
- Deciles 7–9: **+1–3 pp**
- Same inside the bandwidth
- Still there in grade 8



Bandwidth sensitivity

Positive at every bandwidth; rises to a bandwidth of ∼12 pp, then flat





Mechanisms

We digitized every exam question

A new item-level dataset

- Every released TAAS/TAKS/STAAR grade-3 reading and math exam we could recover (13 of each; cross-validated against answer keys)
- Coded per question: **word count**; the share of its words that are **decodable with basic phonics** (CIERA decodability levels 1–4); reading objective (vocabulary vs. comprehension); whether a math item's answer options are all numbers
- Lets us ask **where** the gains come from, not just whether scores moved

What an item looks like

Both “math” — but only one is also a reading test

12 Which number is divisible by 3?

F 53

G 57

H 59

J 61

40 Laura brushed her teeth 3 times each day for 7 days. Each time, she brushed them for 2 minutes. What is one way to find the total number of minutes Laura brushed her teeth during these 7 days?

F Multiply 3 by 7 and then multiply the product by 2

G Multiply 3 by 2

H Add 3 and 2 and 7

J Add 2 to the product of 3 and 7

reading is required!

Scoring the phonics content of each item

CIERA decodability levels

Figure: CIERA Text Analysis Values for Levels of Decodability

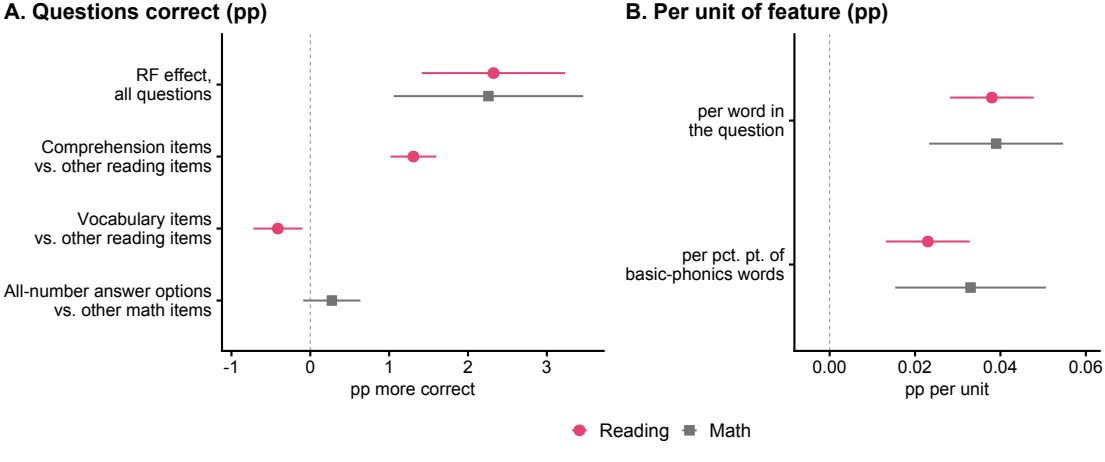
Level	Pattern	Excludes	Example
1	A, I C-V		A, I Me, we, be, he, my, by, so, go, no
2	C-V-C V-C	No words ending in R or L	Man, cat, hot Am, an, as, at, ax, if, in, is, it, of, on, ox, up, us
3	C-C-V V-C-C-[C] C-C-[C]-V-C C-V-C-C-[C] C-C-[C]-V-C-C-[C]	No words ending in R or L r-C or 1-C (e.g. fort, mild) or V-gh (e.g. sigh)	She, the, who, why, cry, dry Ash, itch That, chat, brat, scrap Back, mash, catch Crash, track, scratch
4	[C]-[C]-[C]-V-C-e		Bake, ride, mile, plate, strike, ate
5	C-[C]-V-V-[C]-[C] V-V-C-[C]	No words ending in –gh (e.g., laugh, through, though)	Beat, tree, say, paid Eat, each
6	C-[C]-V-r [C]-[C]-V-r-C [C]-[C]-V-11 C-[C]-V-1-C C-[C]-V-V-1-C		Car, scar, fir Farm, start, art, arm All, ball, shall, tell, will, Told, child Could, should, field, build
7	Diphthongs		Boy, oil, draw, cloud
8	Everything else		

Each word scored 1–8; a question's "basic-phonics share" = share of its words at levels 1–4, decodable with basic letter–sound rules.

Decodability levels from Hoffman, Sailors and Patterson (2002), CIERA Report 1-016 / *Journal of Literacy Research* 34(3); table as reproduced in the paper

Question-level DiD: reading and math

Share of a campus's third graders answering a question correctly (Table 3)



What the questions tell us

Digitized exam questions (Table 3)

- RF students answer **+2.3 pp** more questions correctly — reading and math alike

What the questions tell us

Digitized exam questions (Table 3)

- RF students answer **+2.3 pp** more questions correctly — reading and math alike
- The advantage **grows with question length**: +0.04 pp per word, so a question one s.d. longer (∼25 words) carries ∼+1 pp more, as large as the average effect itself

What the questions tell us

Digitized exam questions (Table 3)

- RF students answer **+2.3 pp** more questions correctly — reading and math alike
- The advantage **grows with question length**: +0.04 pp per word, so a question one s.d. longer (∼25 words) carries ∼+1 pp more, as large as the average effect itself
- …and with the share of **easily decodable words**: +0.03 pp per percentage point of basic-phonics words (∼+0.3 pp per s.d.), on reading and math items alike

What the questions tell us

Digitized exam questions (Table 3)

- RF students answer **+2.3 pp** more questions correctly — reading and math alike
- The advantage **grows with question length**: +0.04 pp per word, so a question one s.d. longer (∼25 words) carries ∼+1 pp more, as large as the average effect itself
- …and with the share of **easily decodable words**: +0.03 pp per percentage point of basic-phonics words (∼+0.3 pp per s.d.), on reading and math items alike
- On the reading exam the edge is larger on **comprehension** items (+1.3 pp) and smaller on vocabulary items (−0.4 pp)

Math as the out-of-sample test

- Math gains concentrate in **word problems**: the RF advantage grows +0.04 pp with every word in the question

31 Victor went jogging 4 days last week. The longest distance he jogged in 1 day was 3.3 miles, and the shortest distance he jogged was 2.4 miles. Which is a reasonable total number of miles Victor jogged during the 4 days?

A Less than 4 mi

B Between 4 mi and 8 mi

C Between 8 mi and 13 mi

D Between 13 mi and 16 mi

E More than 16 mi

Item: released TAKS grade-3 mathematics test, Texas Education Agency

Math as the out-of-sample test

- Math gains concentrate in **word problems**: the RF advantage grows +0.04 pp with every word in the question
- The same decoding gradient as in reading (+0.03 pp per percentage point of basic-phonics words); little difference for items whose answer options are all numbers

31 Victor went jogging 4 days last week. The longest distance he jogged in 1 day was 3.3 miles, and the shortest distance he jogged was 2.4 miles. Which is a reasonable total number of miles Victor jogged during the 4 days?

A Less than 4 mi
B Between 4 mi and 8 mi
C Between 8 mi and 13 mi
D Between 13 mi and 16 mi
E More than 16 mi

Item: released TAKS grade-3 mathematics test, Texas Education Agency

Math as the out-of-sample test

- Math gains concentrate in **word problems**: the RF advantage grows +0.04 pp with every word in the question
- The same decoding gradient as in reading (+0.03 pp per percentage point of basic-phonics words); little difference for items whose answer options are all numbers
- Hard to reconcile with a pure-resources story — **students got better at reading**, and it spilled into math

31 Victor went jogging 4 days last week. The longest distance he jogged in 1 day was 3.3 miles, and the shortest distance he jogged was 2.4 miles. Which is a reasonable total number of miles Victor jogged during the 4 days?

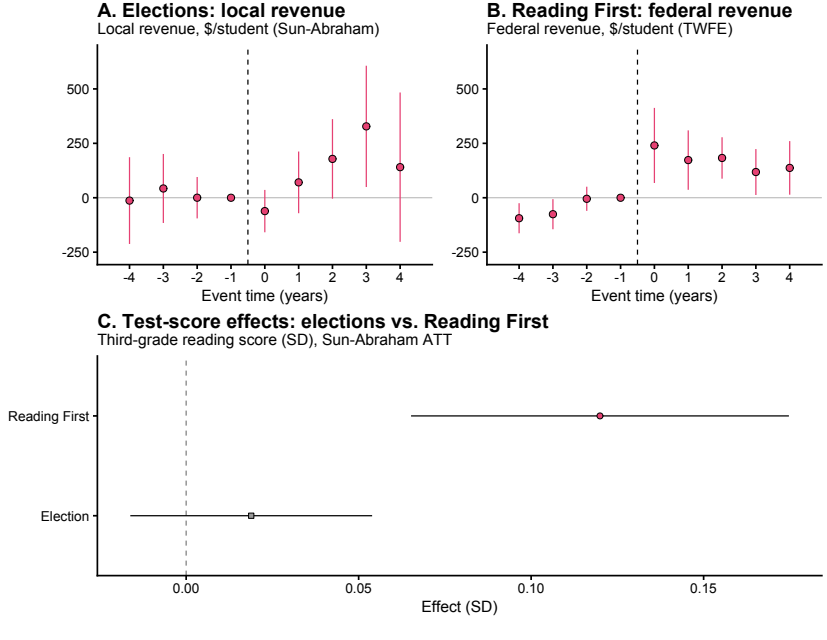
A Less than 4 mi
B Between 4 mi and 8 mi
C Between 8 mi and 13 mi
D Between 13 mi and 16 mi
E More than 16 mi

Item: released TAKS grade-3 mathematics test, Texas Education Agency

Is it just money?

Elections vs. Reading First dollars

- A passed tax election: **+$321** local revenue per student
- Reading: **0.02 vs. 0.12 s.d.**
- The **conditional** grant mattered



Not simply money: tax elections as a comparison

- Passed school tax elections raise **local revenue by ∼$321 per student** (total ∼$428; flat pre-trends); 580 districts, 2000–2019

Not simply money: tax elections as a comparison

- Passed school tax elections raise **local revenue by ∼$321 per student** (total ∼$428; flat pre-trends); 580 districts, 2000–2019
- The windfall goes to **salaries** (+$190) and buys **no detectable reading materials or supplies**

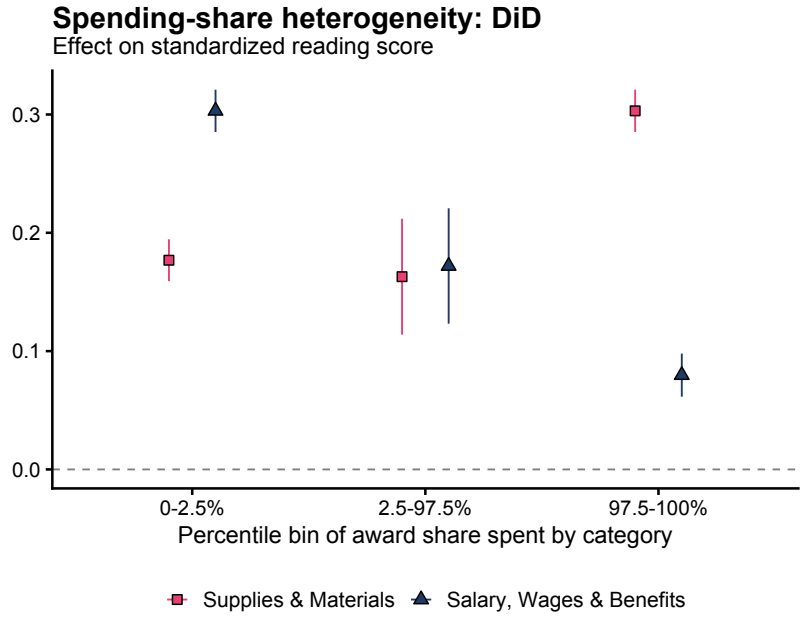
Not simply money: tax elections as a comparison

- Passed school tax elections raise **local revenue by ∼$321 per student** (total ∼$428; flat pre-trends); 580 districts, 2000–2019
- The windfall goes to **salaries** (+$190) and buys **no detectable reading materials or supplies**
- Same panel, same estimator: election money moves reading by an insignificant **0.02 s.d.**; RF moves it **0.12 s.d.**

Does the spending mix matter?

Effect by district spending composition

- For the middle 95% of districts the mix barely matters (**0.16–0.17 s.d.**)
- The **instruction switch itself** does the work



The value of Reading First per dollar

Study	Use of funds	Spending	Impact (s.d.)	Cost per 0.1 s.d.
Environmental conditions				
Yoder (2026)	Bus retrofits, grades 3–8	$109 per student	0.05	$218
Gilraine (2023)	Air filters, grades 4–5	$40 per student-year	0.2	$20
School inputs				
Reading First	**Phonics curricula + coaching, K–3**	**$75–$132 per K–3 student-year**	**0.11–0.07**	**$70–$184**
Krueger (1999)	Class size −1/3 (STAR), K–3	$863 per student-year	0.2	$432
Jackson & Mackevicius (2024)	Overall school budget, K–12	$1,000 per student-year, 4 years	0.03	$3,333

Comparison table adapted with permission from Maranna Yoder (2026).

The value of Reading First per dollar

Study	Use of funds	Spending	Impact (s.d.)	Cost per 0.1 s.d.
Environmental conditions				
Yoder (2026)	Bus retrofits, grades 3–8	$109 per student	0.05	$218
Gilraine (2023)	Air filters, grades 4–5	$40 per student-year	0.2	$20
School inputs				
Reading First	**Phonics curricula + coaching, K–3**	**$75–$132 per K–3 student-year**	**0.11–0.07**	**$70–$184**
Krueger (1999)	Class size −1/3 (STAR), K–3	$863 per student-year	0.2	$432
Jackson & Mackevicius (2024)	Overall school budget, K–12	$1,000 per student-year, 4 years	0.03	$3,333

- Charging the full during-program $132 against the during-program 0.07 s.d.: **$184 per 0.1 s.d.**, still an order of magnitude below general spending (the grant understates teacher time and PD; tripling it leaves a wide gap)

Comparison table adapted with permission from Maranna Yoder (2026).

The program pays for itself

Marginal value of public funds, per dollar of Reading First spending

- Cost: **$132 per K–3 student-year** while funded, about **$400 per child**
- Benefit: the $509 gain at 25 (1.5% of earnings), held constant as a share of earnings; 3% discount, 18.9% tax rate

Earnings counted through age	Earnings gain per $1 (PV)	Tax revenue per $1	Net cost per $1	MVPF
25 (observed only)	$0.75	$0.14	$0.86	0.71
27	$2.39	$0.45	$0.55	3.5
29	$4.16	$0.79	$0.21	15.8
31	**$6.02**	**$1.14**	**−$0.14**	∞
35	$9.89	$1.87	−$0.87	∞
65	$32.17	$6.08	−$5.08	∞

Constant-share projection; earnings gain set to zero after the stated age.

The program pays for itself

Marginal value of public funds, per dollar of Reading First spending

- Cost: **$132 per K–3 student-year** while funded, about **$400 per child**
- Benefit: the $509 gain at 25 (1.5% of earnings), held constant as a share of earnings; 3% discount, 18.9% tax rate
- Age 25 alone: **14 cents back per dollar** (MVPF 0.71)

Earnings counted through age	Earnings gain per $1 (PV)	Tax revenue per $1	Net cost per $1	MVPF
25 (observed only)	$0.75	$0.14	$0.86	0.71
27	$2.39	$0.45	$0.55	3.5
29	$4.16	$0.79	$0.21	15.8
31	**$6.02**	**$1.14**	**−$0.14**	∞
35	$9.89	$1.87	−$0.87	∞
65	$32.17	$6.08	−$5.08	∞

Constant-share projection; earnings gain set to zero after the stated age.

The program pays for itself

Marginal value of public funds, per dollar of Reading First spending

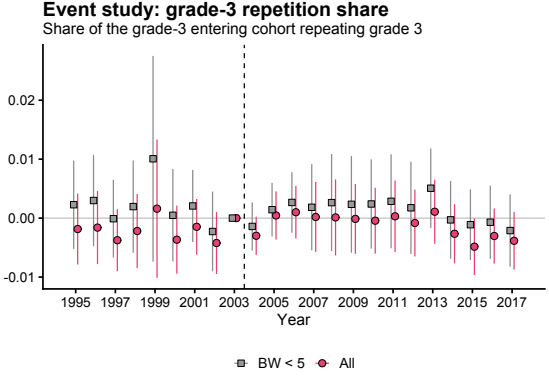
- Cost: **$132 per K–3 student-year** while funded, about **$400 per child**
- Benefit: the $509 gain at 25 (1.5% of earnings), held constant as a share of earnings; 3% discount, 18.9% tax rate
- Age 25 alone: **14 cents back per dollar** (MVPF 0.71)
- Taxes exceed the outlay by **age 31** (29 on Mountjoy's trajectory); through 65: **$32 of earnings, $6 of taxes per dollar**

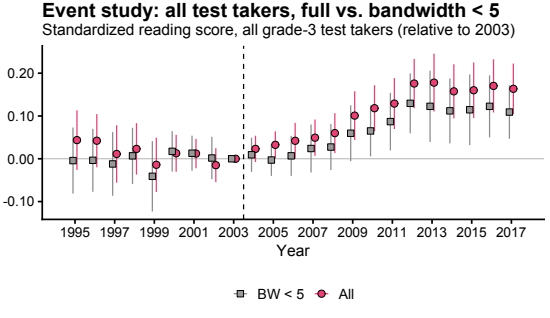
Earnings counted through age	Earnings gain per $1 (PV)	Tax revenue per $1	Net cost per $1	MVPF
25 (observed only)	$0.75	$0.14	$0.86	0.71
27	$2.39	$0.45	$0.55	3.5
29	$4.16	$0.79	$0.21	15.8
31	**$6.02**	**$1.14**	**−$0.14**	∞
35	$9.89	$1.87	−$0.87	∞
65	$32.17	$6.08	−$5.08	∞

Constant-share projection; earnings gain set to zero after the stated age.

Retention is not the story

Texas tied grade-3 promotion to the reading test from 2003; RF did not change who was retained





- Grade-3 repetition share: flat event study, **null**
- All test takers instead of first-time takers: **0.097 vs. 0.110 s.d.**
- No differential exposure to Student Success Initiative funding

Summarizing evidence

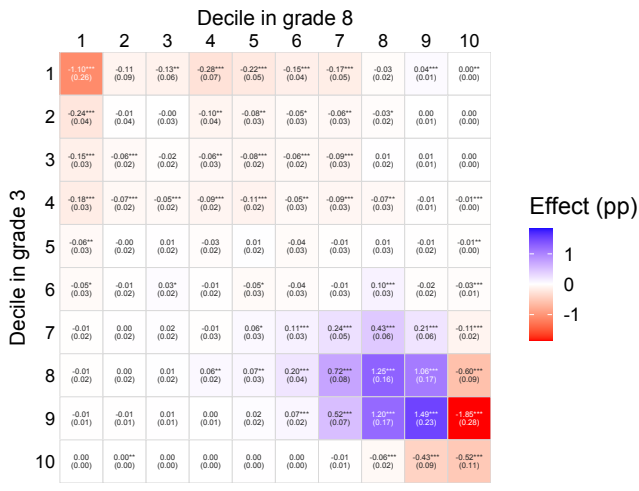
- Gains sit **where reading binds**: longer questions, easily decodable words, comprehension items
- Math moves through **word problems**, with the same per-word and decoding gradients
- The money bought **reading materials** — not library books or textbooks — and paid existing staff: no new teachers, no change in credentials or value added
- **Unconditional money does not reproduce it**: tax windfalls buy salaries, not materials, and move reading ≈ 0
- Not retention, not the exams, and for most districts **not the spending mix**

Long-run effects

The same students, grade 3 to grade 8

Transition probabilities, pp

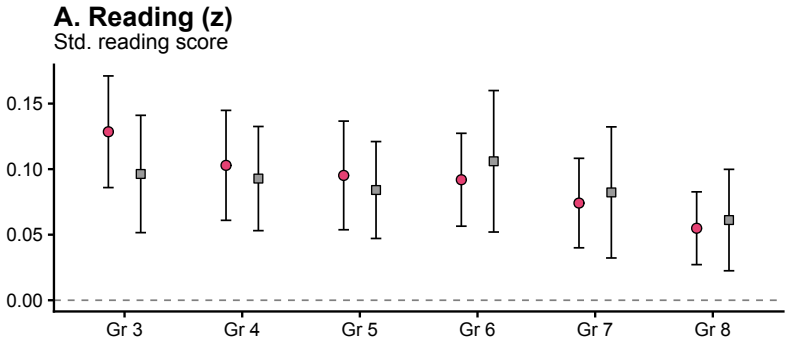
- Bottom-decile starters: **1.1 pp** less likely to still be there in grade 8
- Deciles 7–9 move **up**, not down

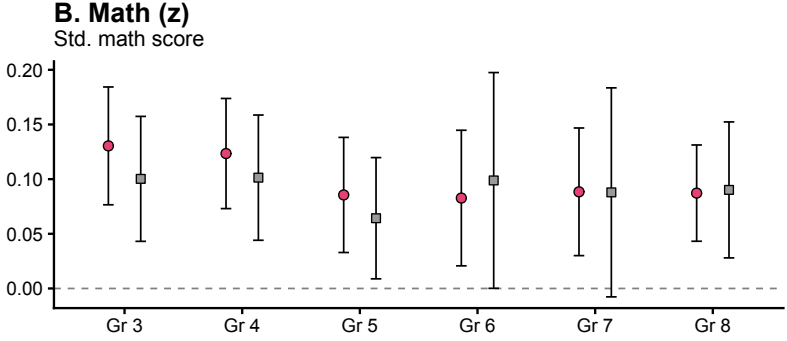


Effects persist across grades

Student level; circles all, squares BW < 5

- Reading: **0.13 s.d.** in grade 3, **0.06** in grade 8
- Math: **0.13** to **0.09**
- Most early interventions fade out; this one did not

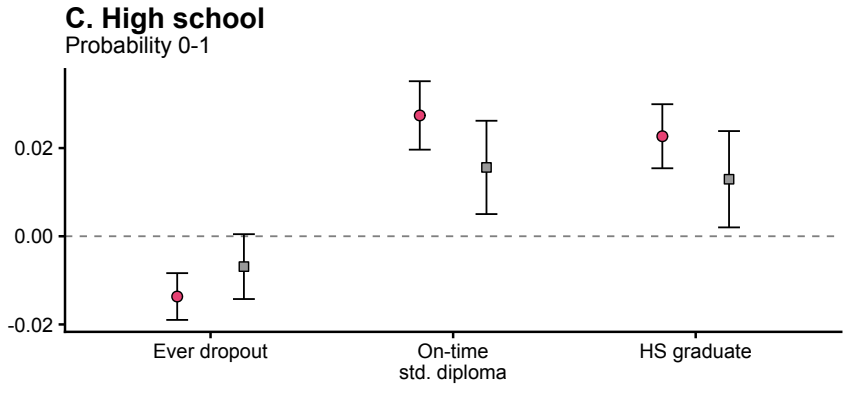




High school

Student level; cohorts through 2014; circles all, squares BW < 5

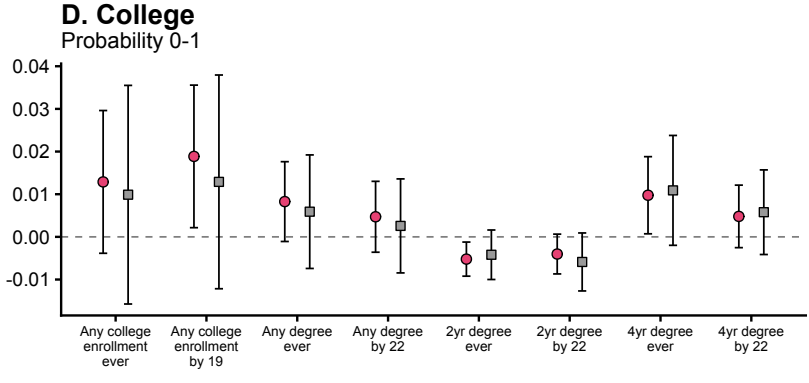
- Ever dropout: -1.4 pp
- On-time diploma: +2.7 pp (base 90.7%)
- Ever graduate: +2.3 pp (base 92.9%)



College

Enrollment through the 2014 cohort, degrees through 2011

- Enroll by 19: **+1.9 pp** (base 56.8%)
- Four-year degree: **+1.0 pp**; two-year: −0.5 pp
- Substitution toward four-year colleges



Long-run effects

Student level; HS and enrollment for cohorts through 2014, degrees through 2011

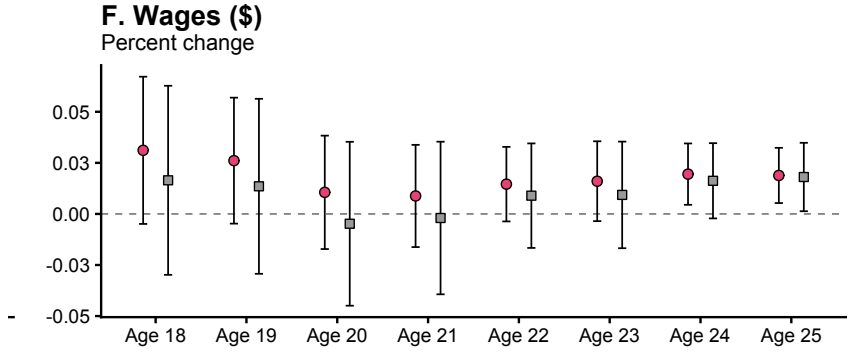
	Full sample	Pre-RF mean
Ever dropout	−1.4 pp***	
On-time standard diploma	+2.7 pp***	90.7%
HS completion (ever)	+2.3 pp***	92.9%
College enrollment by age 19	+1.9 pp**	56.8%
Four-year degree (ever)	+1.0 pp**	29.9%
Two-year degree (ever)	−0.5 pp**	11.6%
Earnings at age 25 (2025 $)	+**$509****	$33,713

Earnings ≈1.5% of the untreated mean (Poisson: 1.9 log points, ≈$641); age 25 is early in the lifecycle and RF-induced schooling depresses it, so this likely understates. 2SLS per $20M: +1.1 pp on-time diploma, +1.7 pp enrollment by 19, +0.9 pp four-year degree.

Wages by age

Percent change (Poisson); circles all, squares BW < 5

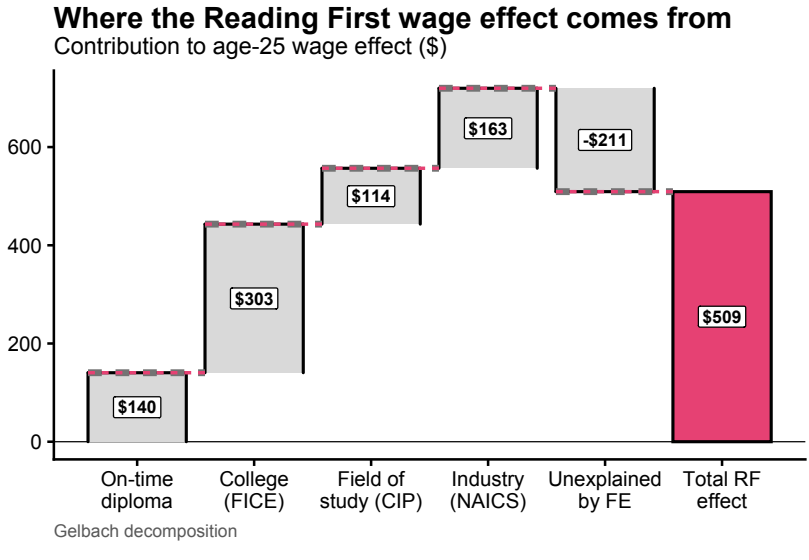
- Positive at **every age 18–25**
- At 25: **+$509** (1.5%), about one year with a one-s.d. better teacher
- A floor: college returns double by age 30 (Mountjoy 2026)



Where the earnings effect comes from

Gelbach decomposition

- College attended **$303**; industry **$163**
- Diploma **$140**; field of study **$114**
- Descriptive, not causal: more schooling, better jobs



Wrapping up

Takeaways

- A conditional, curriculum-linked grant moved reading and math **0.11 s.d.** for **$75–$132 per K–3 student-year**, about $400 per child
- Gains **persist through grade 8** and follow cohorts into **diplomas, four-year degrees, and $509 more at age 25**
- Unconditional money (tax elections) does not reproduce it: ***how* education dollars are spent can matter far more than how much**
- Cheap and self-financing: **$70–$184 per 0.1 s.d.**; pays for itself in tax revenue by age 31
- Open: one state; earnings only to 25; materials, coaching, and assessments cannot be separated within the bundle

Thank you!

hollingsworth.126@osu.edu

Paper · **Non-technical summary** · **These slides**

hollina.github.io

Backup

Backup: reading, main table

	Full sample	Bandwidth < 5
	(1)	(2)
A. DID		
1(RF Eligible) $\times$ 1(year$>$2003)	0.110***	0.084***
	(0.022)	(0.025)
B. DID, separate periods		
1(RF Eligible) $\times$ 1(2004–2011)	0.072***	0.047***
	(0.017)	(0.017)
1(RF Eligible) $\times$ 1(2012 and after)	0.159***	0.130***
	(0.031)	(0.038)
C. 2SLS, cumulative RF $		
Cumulative RF $ per $20 million (2SLS)	0.059***	0.045***
	(0.019)	(0.008)
First-stage F stat	1,711	863
Observations	60,619	34,240
Campus FE	Yes	Yes
Year FE	Yes	Yes

Backup: math, main table

	Full sample	Bandwidth < 5
	(1)	(2)
A. DID		
1(RF Eligible) $\times$ 1(year$>$2003)	0.109***	0.084***
	(0.028)	(0.032)
B. DID, separate periods		
1(RF Eligible) $\times$ 1(2004–2011)	0.067***	0.040
	(0.024)	(0.025)
1(RF Eligible) $\times$ 1(2012 and after)	0.163***	0.140***
	(0.036)	(0.045)
C. 2SLS, cumulative RF $		
Cumulative RF $ per $20 million (2SLS)	0.069***	0.056***
	(0.017)	(0.008)
First-stage F stat	1,574	848
Observations	60,619	34,240
Campus FE	Yes	Yes
Year FE	Yes	Yes

Backup: question-level results (Table 3)

Word count, basic-phonics share, vocabulary, comprehension, all-number options

	Pooled		Reading		Math	
	All	BW < 5	All	BW < 5	All	BW < 5
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Reading First effect						
1(RF Eligible) × 1(year>2003)	2.308***	1.715***	2.323***	1.665***	2.259***	1.713**
	(0.540)	(0.618)	(0.464)	(0.539)	(0.613)	(0.710)
Panel B. Effect by question feature: 1(RF Eligible) × 1(year>2003) × feature						
Word count (per word)	0.041***	0.016***	0.038***	0.021***	0.039***	0.011
	(0.005)	(0.005)	(0.005)	(0.006)	(0.008)	(0.008)
Share of basic-phonics words (per pct. pt.)	0.029***	0.020***	0.023***	0.009	0.033***	0.035***
	(0.005)	(0.007)	(0.005)	(0.006)	(0.009)	(0.013)
Vocabulary question (= 1)			-0.412***	-0.157		
			(0.158)	(0.159)		
Comprehension question (= 1)			1.307***	0.916***		
			(0.148)	(0.185)		
Answer options are all numbers (= 1)					0.273	0.027
					(0.185)	(0.231)
Observations	3,053,732	1,722,818	1,443,812	814,724	1,609,920	908,094
Campus FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Backup: where the money went

Broad expenditure categories, per K–3 student

	All tracked expenditures		Salary, wages & benefits		Travel		Prof. purchased services		Supplies & materials		Misc. operating	
	All	BW < 5	All	BW < 5	All	BW < 5	All	BW < 5	All	BW < 5	All	BW < 5
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
A. DID												
1(RF Eligible) × 1(year>2003)	74.35***	58.87***	41.12***	33.87***	1.50***	0.99***	4.79***	2.95***	26.41***	22.06***	0.55	-1.00
	(8.44)	(11.40)	(4.30)	(5.93)	(0.36)	(0.35)	(1.05)	(0.70)	(3.57)	(4.61)	(1.72)	(2.43)
Observations	23,831	8,628	23,831	8,628	23,831	8,628	23,831	8,628	23,831	8,628	23,831	8,628
B. DID, separate periods												
1(RF Eligible) × 1(2004–2011)	131.47***	105.38***	72.50***	60.28***	2.68***	1.81***	8.35***	5.12***	46.78***	39.49***	1.16	-1.31
	(14.64)	(20.04)	(7.48)	(10.52)	(0.62)	(0.62)	(1.81)	(1.25)	(6.23)	(8.17)	(2.73)	(3.86)
1(RF Eligible) × 1(2012 and after)	1.68	0.23	1.18	0.57	-0.01	-0.03	0.25	0.21	0.48	0.08	-0.22	-0.60
	(1.47)	(1.69)	(0.78)	(0.78)	(0.05)	(0.04)	(0.17)	(0.18)	(0.53)	(0.60)	(0.45)	(0.66)
Observations	23,831	8,628	23,831	8,628	23,831	8,628	23,831	8,628	23,831	8,628	23,831	8,628
C. 2SLS												
Cumulative RF $ per $20 million (2SLS)	2.13***	1.76**	2.90**	1.95***	0.01	-0.11**	-0.96	-0.12***	0.48	0.46	-0.31	-0.43
	(0.75)	(0.75)	(1.17)	(0.53)	(0.14)	(0.05)	(0.82)	(0.04)	(0.82)	(0.41)	(0.20)	(0.31)
First-stage F stat	1687.92	1158.48	1687.92	1158.48	1687.92	1158.48	1687.92	1158.48	1687.92	1158.48	1687.92	1158.48
Observations	23,831	8,628	23,831	8,628	23,831	8,628	23,831	8,628	23,831	8,628	23,831	8,628
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Weights	K–3 students	K–3 students	K–3 students	K–3 students	K–3 students	K–3 students	K–3 students	K–3 students	K–3 students	K–3 students	K–3 students	K–3 students

Backup: long-run outcomes, full table

	High School		Any College Enrollment		College Degree						Age 25	
					Any		2-year		4-year			
	On-Time Diploma	Any Completion	Ever	By 19	Ever	By 22	Ever	By 22	Ever	By 22	Employed	Earnings ($)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
A. DID (FEOLS)												
1(RF Eligible) × 1(3rd grade Year ≥ 2004)	2.74***	2.27***	1.29	1.89**	0.83*	0.47	-0.52**	-0.40*	0.98**	0.48	-0.24	$509**
	(0.40)	(0.37)	(0.85)	(0.85)	(0.48)	(0.42)	(0.20)	(0.24)	(0.46)	(0.37)	(0.17)	($213)
B. 2SLS: Cumulative RF $ (per $20 million)												
Cumulative RF $ per $20 million (2SLS)	1.14***	0.95***	1.48***	1.70***	0.79***	0.12	-0.07	-0.31***	0.90***	0.24	-0.47***	$441*
	(0.21)	(0.14)	(0.31)	(0.24)	(0.17)	(0.21)	(0.13)	(0.09)	(0.18)	(0.22)	(0.13)	($245)
First-stage F stat	1,368	1,368	1,390	1,390	1,475	1,475	1,475	1,475	1,475	1,475	1,366	1,366
C. Poisson (Earnings Only)												
Poisson coefficient (log points)	–	–	–	–	–	–	–	–	–	–	–	0.019***
												(0.007)
Implied $ effect at mean	–	–	–	–	–	–	–	–	–	–	–	$641
Observations	4,913,997	4,915,214	5,875,509	5,875,509	4,825,091	4,825,091	4,825,091	4,825,091	4,825,091	4,825,091	3,214,226	3,214,546
Last third-grade cohort	2014	2014	2014	2014	2011	2011	2011	2011	2011	2011	2014	2014
Pre-RF Cohort Mean	90.7	92.9	62.0	56.8	35.9	25.4	11.6	6.9	29.9	20.1	78.7	$33,713
Campus FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Third-Grade-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Backup: tax elections, revenue and spending

Sun–Abraham ATT, 580 treated districts, 2000–2019

	Tax election
A. Revenue ($ per student)	
Total revenue	427.7***
	(125.8)
Local revenue	320.9*
	(164.6)
B. Expenditure ($ per student)	
All tracked expenditure	339.6***
	(130.4)
Salary, wages & benefits	189.5***
	(53.0)
Travel	7.0***
	(2.2)
Prof. purchased services	125.5
	(111.2)
Supplies & materials	24.7
	(27.2)
General supplies	16.4
	(24.1)
Testing materials	−2.0***
	(0.7)
Books (total)	−4.2
	(5.5)
Library books	0.3
	(1.7)
Textbooks	−3.1
	(4.4)
Reading materials	−1.5
	(3.9)
Misc. operating	−7.1
	(13.9)
Observations	24,112
Districts	1,343
Passed a tax election	580
District FE	Yes
Year FE	Yes
Weights	Enrollment

Backup: MVPF by projection horizon

Earnings projected through age	Constant share of the age-earnings profile					Mountjoy (2026) trajectory	
	PV of earnings gain per $1 (1)	Additional tax revenue per $1 (2)	Net cost to government per $1 (3)	Beneficiary WTP per $1 (4)	MVPF (5)	Net cost to government per $1 (6)	MVPF (7)
25 (no projection)	$0.75	$0.14	$0.86	$0.61	0.71	$0.86	0.71
26	$1.54	$0.29	$0.71	$1.25	1.77	$0.66	2.16
27	$2.39	$0.45	$0.55	$1.93	3.52	$0.48	4.71
28	$3.26	$0.62	$0.38	$2.64	6.87	$0.22	14.83
29	$4.16	$0.79	$0.21	$3.38	15.83	**–$0.09**	∞
30	$5.09	$0.96	$0.04	$4.12	105.98	–$0.46	∞
31	**$6.02**	**$1.14**	**–$0.14**	**$4.88**	∞	–$0.83	∞
35	$9.89	$1.87	–$0.87	$8.02	∞	–$2.37	∞
45	$19.67	$3.72	–$2.72	$15.95	∞	–$6.25	∞
55	$27.55	$5.21	–$4.21	$22.34	∞	–$9.38	∞
65 (baseline)	$32.17	$6.08	–$5.08	$26.09	∞	–$11.22	∞

Backup: the full life course

Figure 6, all six panels

