The Long-Term Impacts of Low-Achieving Childhood Peers: Evidence from Project STAR

Jan Bietenbeck

Lund University and IZA

updated August 2018

What are the long-term impacts of school peers?

- **Large existing literature on peer effects in schools:**
  - Positive impacts from sharing a classroom with high-achieving and better-behaved peers on school performance (Hoxby 2000, ...).
  - Negative impacts from being in class with low-performing and disruptive peers (Figlio 2007, Carrell and Hoekstra 2010, ...).

- **But: almost exclusive focus on contemporaneous outcomes.**
  Example: the impact of first-grade peers on first-grade math scores.

- To judge the overall efficacy of policies that affect peer composition need to know: **what are the long-term impacts of school peers?**

Long-term spillovers from repeaters in kindergarten

- **This paper: how does sharing a kindergarten (KG) classroom with low-achieving repeaters affects non-repeating students' educational performance in the long run?**

- **Use data from Project STAR, which have three advantages:**
  1. Can identify KG repeaters as a particularly low-achieving group of peers (proven track record of failure, very low cognitive and non-cog. skills).
  2. Random assignment of teachers and students, including repeaters, to KG classes within schools $\Rightarrow$ can estimate causal spillover effects.
  3. Can follow students throughout their entire school career.

Preview of results

- **Empirical analysis compares outcomes of regular students in KG classes with and without repeaters in the same school $\Rightarrow$ identifies causal effect due to random assignment of repeaters.**

- **Students who are exposed to repeaters in kindergarten**
  1. initially score lower on standardized tests, but impact fades out quickly.
  2. show persistent improvements in non-cognitive skills such as discipline.
  3. are more likely to graduate from high school and to take a college entrance exam around the age of 18.

Related literature and contribution

- **Only few previous studies on long-term impacts of school peers:** Gould et al. (2009), Cascio and Schanzenbach (2016); Bifulco et al. (2011), Black et al. (2013).

- **This paper: provide the first evidence on long-term spillovers from low-achieving childhood peers.**
  - Both cognitive and non-cognitive skills still highly malleable in KG.
  - Effects on long-term educational outcomes likely translate more directly into changes in labor market outcomes than effects on test scores.

Talk outline

1. Introduction

2. The STAR experiment and data

3. Empirical strategy and main results

4. Mechanisms and robustness checks
   - Non-cognitive skills as a channel
   - Mechanisms underlying the non-cognitive effects
   - Additional results and robustness checks

5. Conclusion

Background on Project STAR

- Tennessee Student/Teacher Achievement Ratio (STAR) experiment: study of the effects of class size on student achievement, 1985-1989.

- **KG students/teachers randomly assigned to classes** w/in schools
  - 79 schools; 325 classes; 6,325 students, including 193 repeaters.
  - Small (ca. 15 students) or regular-sized (ca. 23 students) classes.
  - Repeaters also randomly assigned $\Rightarrow$ **observe classes with and without repeaters within the same school.**

- Experiment ended and students returned to ordinary classes after third grade. Data on (subsets of) participants also collected long after.

Definitions of treatment and treated

- **Main treatment definition: indicator for sharing a kindergarten classroom with at least one repeater.**
  - Motivation: very few classes contain more than one repeater. ▸ Details
  - Interpretation: differential exposure to repeaters during KG + 1/2 year.

- **Similar results from alternative treatment definitions** (number of repeaters; share of repeaters; indicator for any + share of repeaters).

- **Repeaters act only as treatment** and are not treated themselves.

Three sets of outcome variables

- **Cognitive skills** as measured by standardized multiple-choice tests in math and reading at the end of KG – 8^{th} grade.
- **Non-cognitive skills from teacher ratings** in 4^{th} and 8^{th} grade:
  - Effort index: completes homework, is persistent, ...
  - Initiative index: participates in classroom discussions, ...
  - Value index: appreciates the school learning environment, ...
  - Discipline index: often acts restless, distracts classmates, ...
- **Long-term educational attainment** as measured by high school grade point average (GPA) and graduation and college-test taking.

Descriptive statistics: non-repeaters vs. repeaters

	Non-repeaters		Repeaters	
	Mean	SD	Mean	SD
Demographic characteristics				
Male	0.51	0.50	0.70	0.46
Black	0.33	0.47	0.17	0.38
Free lunch	0.48	0.50	0.65	0.48
Age in years	5.48	0.31	6.39	0.31
Old for grade	0.03	0.17	1.00	0.00
Repeater exposure				
At least 1 repeater in class	0.39	0.49	–	–
Selected outcomes				
Kindergarten math score	0.00	1.00	-0.36	0.80
8th-grade math score	0.00	1.00	-0.88	1.09
Non-cog. skills (index)	0.00	1.00	-0.71	1.09
High school graduation	0.87	0.34	0.67	0.48
Took college entrance exam	0.41	0.49	0.12	0.32

Talk outline

1. Introduction

2. The STAR experiment and data

3. Empirical strategy and main results

4. Mechanisms and robustness checks
   - Non-cognitive skills as a channel
   - Mechanisms underlying the non-cognitive effects
   - Additional results and robustness checks

5. Conclusion

Regression specification

- **Regressions compare outcomes of students randomly assigned to KG classes with and without repeaters in the same school:**

$$y_{ics} = \alpha_s + \beta \mathrm{EXPOSURE}_{cs} + \gamma \mathrm{SMALL}_{cs} + X_{ics}\delta + \varepsilon_{ics},$$

  where i = student, c = class, s = school, EXPOSURE = repeater-exposure dummy, SMALL = small-class dummy, and X = controls.

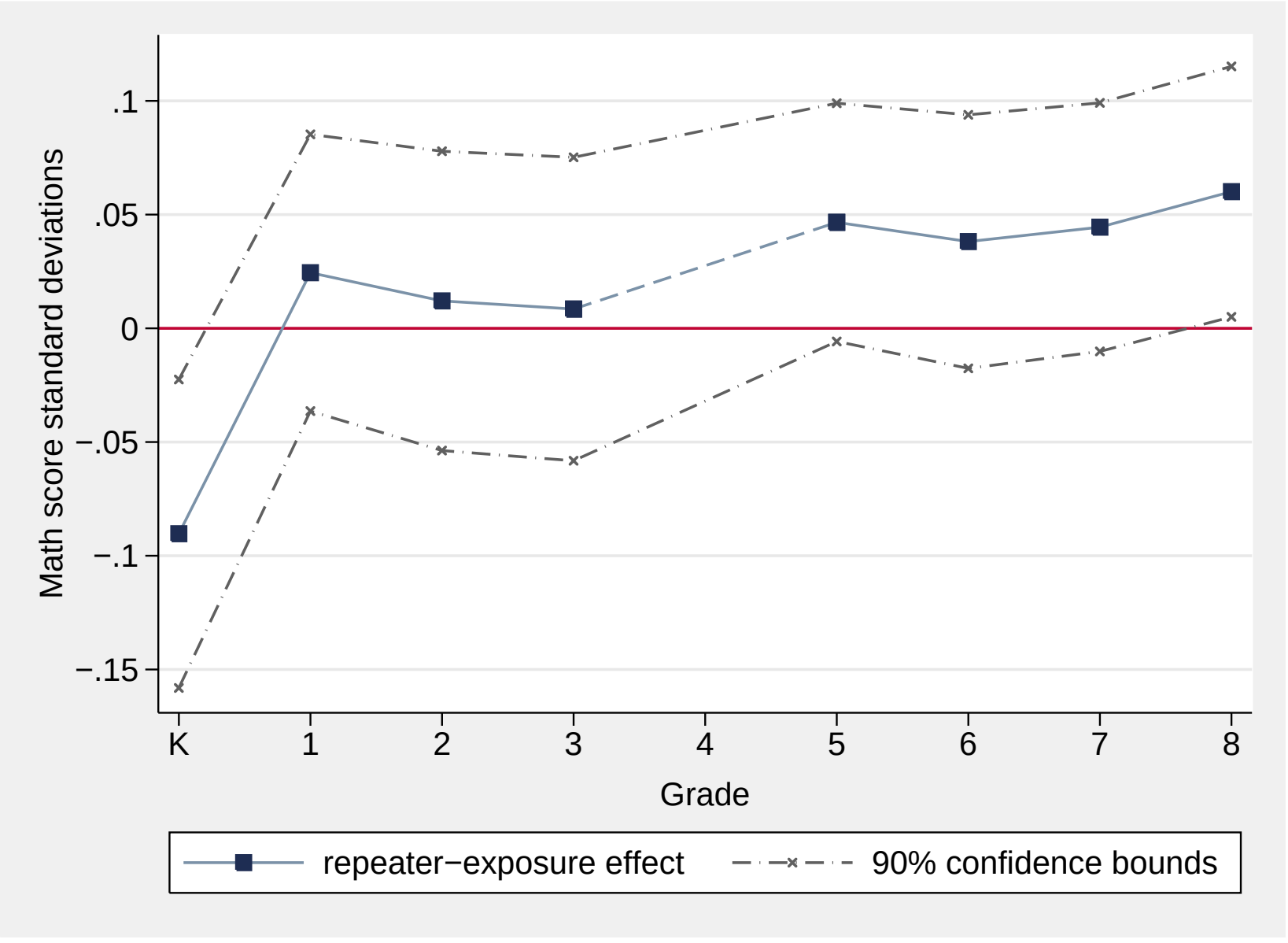
- **Identifying assumption:** classes with and without repeaters do not differ systematically in any other dimension conditional on controls $\Rightarrow$ holds due to random assignment of students and teachers in STAR.

Repeater exposure lowers end-of-kindergarten test scores

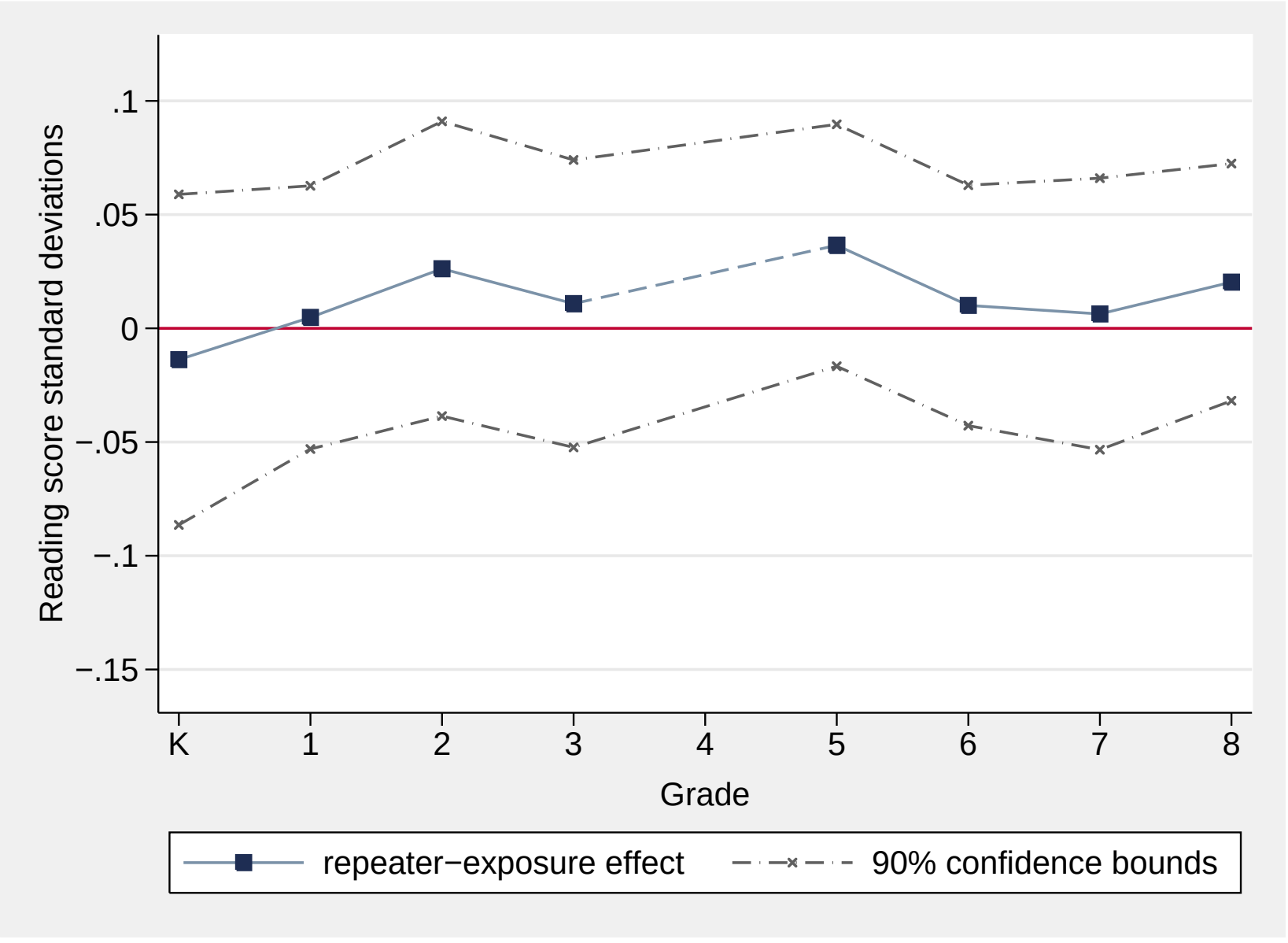
	End-of-kindergarten test scores			
	Math	Math	Reading	Reading
Repeater exposure	–0.090**	–0.090**	–0.014	–0.014
	(0.043)	(0.041)	(0.046)	(0.044)
Male		–0.144***		–0.175***
		(0.024)		(0.025)
Black		–0.355***		–0.249***
		(0.051)		(0.053)
Free lunch		–0.411***		–0.450***
		(0.029)		(0.029)
Age in years		0.550***		0.408***
		(0.044)		(0.048)
Old for grade		–0.411***		–0.346***
		(0.081)		(0.074)
Small class	0.169***	0.158***	0.194***	0.185***
	(0.045)	(0.043)	(0.043)	(0.042)
Observations	5,614	5,614	5,535	5,535

Notes: Standard errors clustered at KG class level. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Negative spillover from repeaters on math scores fades out



Negative spillover from repeaters on reading scores fades out



Repeater exposure raises non-cognitive skills

	Effort	Initiative	Value	Discipline
		Panel A: 4th grade		
Repeater exposure	0.104*	0.025	0.124**	0.142***
	(0.054)	(0.056)	(0.053)	(0.054)
Observations	1,628	1,628	1,628	1,628
		Panel B: 8th grade		
Repeater exposure	0.169***	0.105*	0.160***	0.194***
	(0.054)	(0.056)	(0.051)	(0.052)
Observations	1,731	1,731	1,731	1,731
		Panel C: summary index		
Repeater exposure		0.117***		
		(0.041)		
Observations		2,589		

Notes: Standard errors clustered at KG class level. * p<0.10, ** p<0.05, *** p<0.01.

Repeater exposure raises long-term educational attainment

	HS GPA	HS graduate	Took ACT/SAT	Summary index
Repeater exposure	0.552*	0.021*	0.033**	0.074***
	(0.308)	(0.013)	(0.015)	(0.028)
Observations	2,438	2,955	6,039	6,039

Notes: Standard errors clustered at KG class level. * p<0.10, ** p<0.05, *** p<0.01.

Talk outline

1. Introduction

2. The STAR experiment and data

3. Empirical strategy and main results

4. Mechanisms and robustness checks
   - Non-cognitive skills as a channel
   - Mechanisms underlying the non-cognitive effects
   - Additional results and robustness checks

5. Conclusion

Making sense of the results

- Fading negative effect on test scores, but lasting positive impact on non-cognitive skills $\Rightarrow$ **improved non-cognitive skills as the main driver of positive long-term effects?**

- Idea in line with previous evidence that early-life non-cognitive skills determine later educational success (e.g. Heckman et. al 2006).

- **Roadmap for studying mechanisms:**
  1. Evidence in support of the non-cognitive skills channel.
  2. How does repeater exposure raise non-cognitive skills?

Lecture outline

1. Introduction

2. The STAR experiment and data

3. Empirical strategy and main results

4. Mechanisms and robustness checks
   - Non-cognitive skills as a channel
   - Mechanisms underlying the non-cognitive effects
   - Additional results and robustness checks

5. Conclusion

Suggestive evidence: non-cognitive skills as a channel

	Summary index of long-term attainment			Difference
	(1)	(2)	(3)	[(2)-(3)]
Repeater exposure		0.060	0.012	0.048***
		(0.042)	(0.038)	[p=0.004]
Non-cog. index	0.408***		0.408***	
	(0.019)		(0.019)	
Observations	2,589	2,589	2,589	

Notes: Standard errors clustered at KG class level. * p<0.10, ** p<0.05, *** p<0.01.

- Predicted impact of repeater exposure on long-term outcomes via non-cognitive skills: $0.117 \times 0.408 = 0.048 = 2/3$ of main estimate.
- Estimated impact of repeater exposure on long-term outcomes drops significantly once intermediate non-cognitive skills are controlled for.

Lecture outline

1. Introduction

2. The STAR experiment and data

3. Empirical strategy and main results

4. Mechanisms and robustness checks
   - Non-cognitive skills as a channel
   - Mechanisms underlying the non-cognitive effects
   - Additional results and robustness checks

5. Conclusion

How does repeater exposure raise non-cognitive skills?

- **Three broad classes of explanations for non-cog. skills results:**
  1. **Selection** into classes or schools, selection out of the sample.
  2. **Additional resources** in classes containing repeaters.
  3. **Behavioral adjustments:** e.g. disruption by repeaters $\Rightarrow$ teachers shift their focus from teaching curriculum towards teaching behavioral skills.

- **Present evidence against the first two explanations,** then discuss the plausibility of the third explanation.

Evidence against the selection channel

- Repeaters were indeed randomly assigned to KG classes. ▸ Details
- Selection into classes between KG and first measurement of non-cog. skills (4^{th} grade)? Main way to do this is to leave the experiment. Checked: repeater-exposed students are not more likely to leave.
- Selective attrition ("healthy survivor effect"): no evidence of selection out of the sample based on repeater exposure. ▸ Details

Evidence against the resources channel

- Well-designed and well-implemented experiment: no assignment of additional teachers etc. to classes containing repeaters.

- But experiment did not control (pull-out) special education programs, which may foster non-cognitive skills. However, unlikely mechanism:
  - Repeater-exposed students not more likely to participate in programs.
  - Results are robust to excluding classes with participating repeaters.

Behavioral responses as a likely channel

- **Idea: teachers, students, or parents react to the presence of repeaters in a way that promotes non-cognitive skills.** E.g.:
  - Repeaters disrupt the class by misbehaving (cf. low non-cog. skills).
  - Teachers focus on teaching behavioral skills rather than curriculum.

- **Unfortunately, no data on teaching practices or lesson content. But explanation receives some support from previous studies:**
  - Teachers adjust their teaching practices to students' cognitive ability and behavior (Nurmi 2012) and are more likely to establish explicit rules for behavior when facing low-ability students (Pakarinen et al. 2011).

Lecture outline

1. Introduction

2. The STAR experiment and data

3. Empirical strategy and main results

4. Mechanisms and robustness checks
   - Non-cognitive skills as a channel
   - Mechanisms underlying the non-cognitive effects
   - Additional results and robustness checks

5. Conclusion

Additional results and robustness checks

- Heterogeneity by demographic background and class size. ▸ Details
- Controlling for classmates' demographic characteristics. ▸ Details
- Alternative measures of repeater exposure. ▸ Details

Talk outline

1. Introduction
2. The STAR experiment and data
3. Empirical strategy and main results
4. Mechanisms and robustness checks
   - Non-cognitive skills as a channel
   - Mechanisms underlying the non-cognitive effects
   - Additional results and robustness checks
5. Conclusion

Conclusion (I)

- **Provide some of the first evidence of the long-term impacts of early childhood peers.**

- **Impacts from sharing a KG classroom with repeaters:**
  - Negative but rapidly fading effect on standardized test scores.
  - Positive effect on non-cognitive skills, which persists over time.
  - Positive effect on long-term educational attainment.

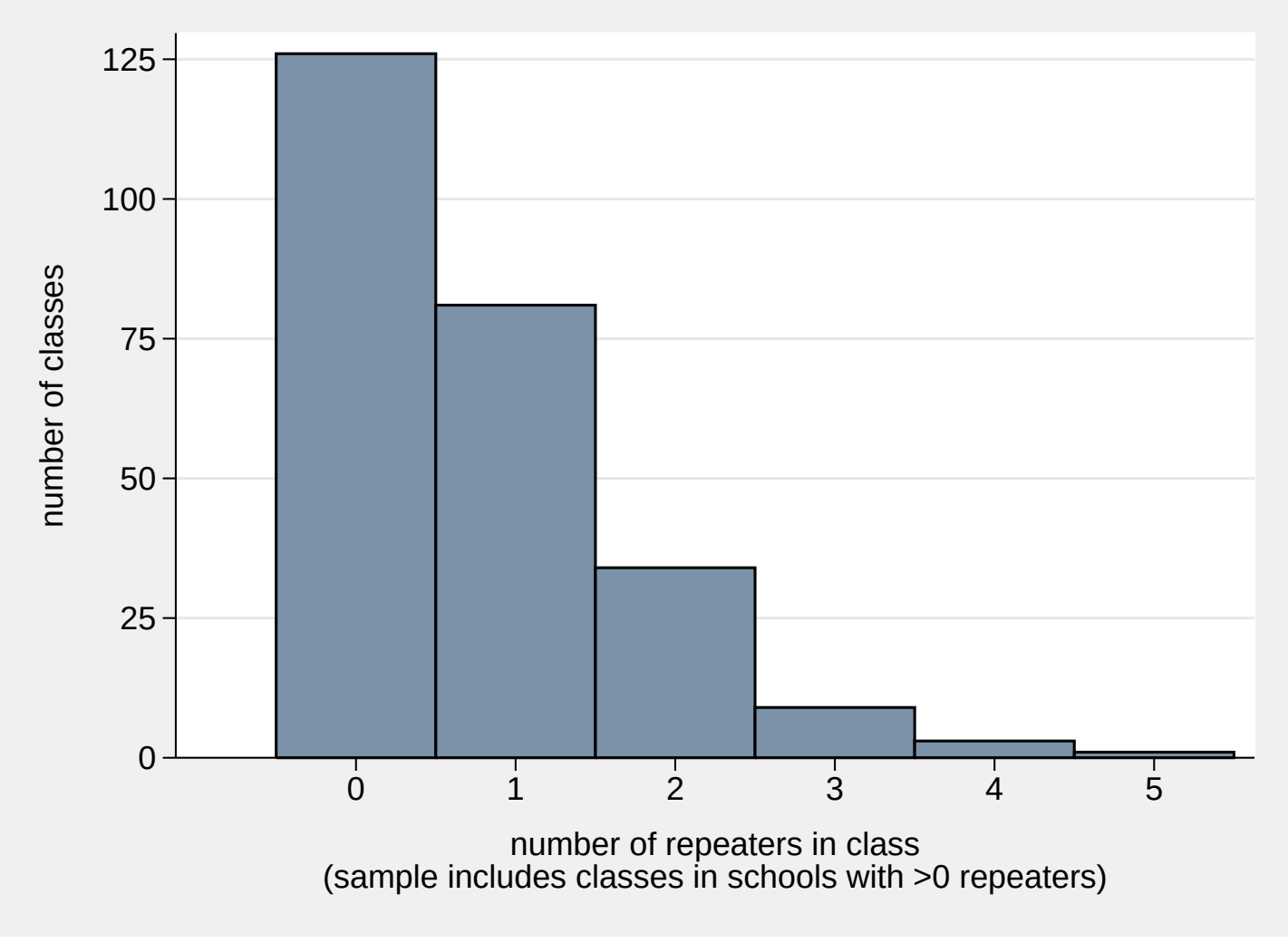
- **Non-cognitive skills as the likely channel for the impact of repeater exposure on long-term outcomes.** Evidence consistent with the idea that the differential accumulation of such skills is due to changed behavior by teachers.

Conclusion (II)

- **Policy implication #1: should we separate low-achieving and high-achieving students at an early age?**
  - Looking at short- and long-term outcomes gives opposite answers $\Rightarrow$ shows importance of looking at long-term effects.

- **Policy implication #2: should we teach kindergarten students primarily non-cognitive skills?**
  - Results in this paper provide fresh evidence of the importance of non-cognitive skills learned early in life.

Distribution of repeaters across classes



Back to Project STAR

Repeaters were randomly assigned to KG classes

Regressions of regular students' demographic characteristics on treatment:

	Male	Black	Free lunch	Age in years	Old for grade
	Panel A: controlling for school fixed effects				
Repeater exposure	–0.005	–0.001	0.004	0.001	–0.003
	(0.015)	(0.007)	(0.015)	(0.009)	(0.005)
	Panel B: controlling for school fixed effects and class size				
Repeater exposure	–0.006	–0.002	0.001	0.004	–0.001
	(0.015)	(0.007)	(0.015)	(0.009)	(0.005)
Obs. (both panels)	6,039	6,039	6,039	6,039	6,039

Notes: Standard errors clustered at KG class level. * p<0.10, ** p<0.05, *** p<0.01.

▸ Back to selection

No attrition from the sample based on repeater exposure

	KG math	G4 dpln	G8 dpln	HS grad	ACT/SAT
	Panel A: outcome = 1(observed with variable in column head)				
Exposed	–0.011	–0.012	–0.020	0.007	
	(0.008)	(0.015)	(0.013)	(0.014)	
	Panel B: like panel A + interactions with demographic controls				
Joint *p*	0.47	0.40	0.60	0.48	
	Panel C: outcomes are the variables in the column heads, sample is restricted to non-attritors (*N*=2,100)				
Exposed	–0.081	0.160**	0.228***	0.010	0.038*
	(0.057)	(0.072)	(0.055)	(0.012)	(0.022)

Notes: Standard errors clustered at KG class level. * p<0.10, ** p<0.05, *** p<0.01.

Back to selection

Heterogeneous effects

- Some evidence that traditionally low-achieving students (males, free-lunchers, blacks) profit less from repeater exposure in the long term.
- No consistent differences in effects by class size:

	KG math	G8 math	NC index	LT index
Repeater exposure	–0.078	0.089**	0.119**	0.065**
	(0.048)	(0.039)	(0.050)	(0.032)
× small class	–0.043	–0.099	–0.007	0.031
	(0.095)	(0.064)	(0.084)	(0.056)
Observations	5,614	4,353	2,589	6,039

Notes: Standard errors clustered at KG class level. * p<0.10, ** p<0.05, *** p<0.01.

▸ Back to robustness

Controlling for classmates' demographic characteristics

	KG math	G8 math	NC index	LT index
	Panel A: controlling for the share of male classmates			
Repeater exposure	–0.087**	0.060*	0.118***	0.075***
	(0.041)	(0.033)	(0.041)	(0.028)
	Panel B: controlling for the share of black classmates			
Repeater exposure	–0.090**	0.061*	0.118***	0.074***
	(0.041)	(0.033)	(0.041)	(0.028)
	Panel C: controlling for the average age of classmates			
Repeater exposure	–0.116***	0.043	0.095**	0.064**
	(0.044)	(0.034)	(0.045)	(0.030)
	Panel D: controlling for the share of old-for-grade classmates			
Repeater exposure	–0.086	0.075*	0.136***	0.080**
	(0.053)	(0.038)	(0.051)	(0.032)

Notes: Standard errors clustered at KG class level. * p<0.10, ** p<0.05, *** p<0.01.

▸ Back to robustness

Alternative measures of repeater exposure

	KG math	G8 math	NC index	LT index
	Panel A: indicators for different numbers of repeaters			
1 repeater in class	−0.096**	0.070*	0.120***	0.072**
	(0.046)	(0.036)	(0.046)	(0.031)
2 repeaters in class	−0.092	0.039	0.134**	0.093**
	(0.070)	(0.053)	(0.065)	(0.041)
3-5 rep's in class	−0.021	0.019	0.025	0.030
	(0.103)	(0.090)	(0.090)	(0.063)
	Panel B: linear share of repeaters in class			
Share of repeaters	−0.601	0.370	1.045**	0.781**
	(0.483)	(0.406)	(0.445)	(0.310)
	Panel C: exposure dummy and linear share of repeaters			
Rep. exposure	−0.135**	0.096*	0.123*	0.056
	(0.068)	(0.057)	(0.068)	(0.045)
Share of repeaters	0.659	−0.525	−0.080	0.256
	(0.802)	(0.677)	(0.739)	(0.484)

Notes: Standard errors clustered at KG class level. * p<0.10, ** p<0.05, *** p<0.01.

Back to robustness