On the Mechanisms of Ability Peer Effects

Alexandra de Gendre[1] Nicolás Salamanca[2]

[1]School of Economics, The University of Sydney

[2]Melbourne Institute, The University of Melbourne

ESAM
8 July 2021

Why Peer Effects In Education Are Interesting

- Economists are very excited about peer effects, including in education
  - 100+ articles in economic journals since 2009, 28 in Top 5
- Why is it interesting? Promise of **social multiplier** effects:

  "Peer effects could be harnessed to cost-effectively improve public services" (BenYishay and Mobarak, REStud 2018)

- Still, **mechanisms** remain unclear, many untested
  - limits work on theoretical underpinnings of peer effects
  - limits scope for using class assignment policies to improve outcomes

What We Know About Ability Peer Effects

- Higher-ability peer effects do exist
  - Large effects on choices: substance abuse, cheating, majors.
  - Much smaller positive effects on scores
  - Context specific: group size, settings, demographics here
  - Re-shuffling yields surprises (Carrell, Sacerdote & West ECTA 2013)
- Not many mechanisms have been tested
  - student effort provision (Todd & Wolpin JPE 2018)
  - classroom disruption (Lavy & Schlosser AEJ:AE 2011)
  - classroom dynamics (Feld & Zölitz JOLE 2017)
  - teacher effort responses (Duflo, Dupas & Kremer AER 2011)

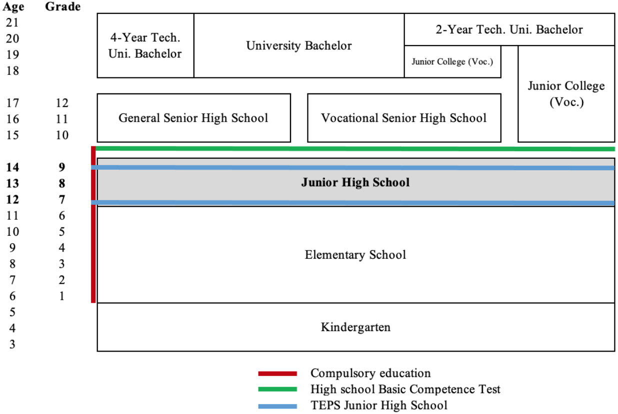
This paper

1. We estimate the effect of higher-achieving peers on test scores, in a setting where there is **a mandate to randomly assign students to classrooms** within schools
   - *Methodological innovation:* a data-driven **"fishing algorithm"** algorithm to trim schools that violate mandate of random assignment, generalizable to many other settings
2. We estimate the effect of higher-achieving peers on a **comprehensive set of 19 inputs** from students, parents and teachers
   - jointly explain 71% of the variation in test scores
3. We estimate **returns to these inputs** using cumulative VA models
4. We **decompose peer effect on test scores** in effects on educational inputs using mediation analysis

What We Find

1. Classroom peers with $+1\sigma$ higher baseline test scores $\Rightarrow$ $+0.052\sigma$ in own test scores
2. $+1\sigma$ higher baseline test scores also:
   - decrease study effort by 5.2%σ
   - increase student university aspirations by 1.6p.p. (+2.9%)
   - increase expectations about attending university by 2.0p.p. (+4.8%)
   - increase parental time investment by 8.1%σ
   - increase parental strictness by 3.6%σ
3. ...and have a precisely estimated null effect on 14 other inputs
4. Most of these inputs are productive for test scores...
5. ...yet changes in inputs **explain almost nothing** of our 5.2%σ peer effect on test scores

Setting and Data



Taiwan Education Panel Study

- 3-Stage stratified random sample of students in Taiwan
  - Schools (333)
  - Classrooms (1,244)
  - Children (20,004)
- Rich multi-party survey
  - Students
  - Parents
  - Teachers
- Repeated measures
- Behaviors, beliefs, attitudes

The Effect of Higher-Achieving Peers

- We want to estimate this equation:

$$Y_{ics2} = \beta \overline{TestScore}_{cs1}^{-i} + \delta \, TestScore_{ics1} + \theta' \boldsymbol{X}_{cs1} + \mu_s + \epsilon_{ics1}$$

  - δ_1: effect of peers' average test scores at baseline
  - $TestScore_{ics1}$: own achievement at baseline
  - $\boldsymbol{X}_{cs1}$: vector of baseline inputs and controls
  - μ_s: school-level heterogeneity, accounted for via school fixed-effects
  - ϵ_{ics1}: error term (clustered at the classroom level)

- β is unbiased estimate of interest **if** $\overline{TestScore}_{cs1}^{-i}$ **as good as random**

Empirical Strategy

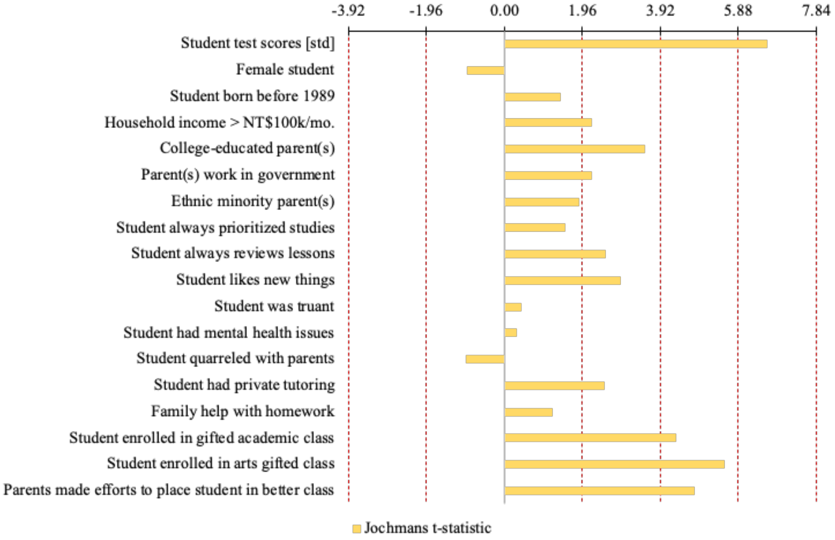
Rely on the mandate randomly assign students to classrooms within schools in Grade 7

- Article 12 in *Taiwan Primary and Junior High School Act*, 2004/2005
- Before 2004, Art. 2 in *The Implementation Guideline of Class Assignment in Junior High School in Taiwan*

But we find evidence of sorting and treatment imbalance using measures at assignment and retrospective measures way before assignment

- Sorting relates pre-assignment Y_{ics} to $\overline{Y}_{ics}^{-i}$; Balancing relates pre-assignment Y_{ics} to $\overline{TestScores}_{ics}^{-i}$
- For sorting tests, we use the standard Guryan, Notowidigdo & Kroft (AEJ:AE 2009) test and Jochmans (2020) extension

Jochmans (2020) Sorting Test



The Fishing Algorithm: Motivation

- From setting and data, we would expect most schools to comply with random assignment
  - A few non-compliers could be legitimate in context (e.g., gifted classes)
  - Some could be true violations (e.g., principals cater to pushy parents)
  - This is (partially) unobserved to us
- We could be rejecting tests because of only a few defier schools! Better to trim these than to rely on conditional exogeneity assumptions or to throw away the data

The Fishing Algorithm: 5 Steps

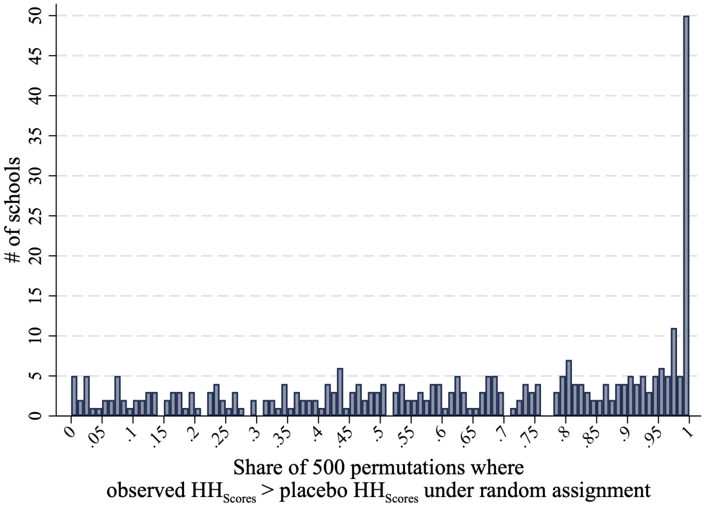
For each school s:

1. Construct H_s: a Herfindahl-Hirschman index of *ability concentration* in classrooms
2. Construct $H_s^{random,k}$, $k = 1 \ldots K$: counterfactual ability concentrations if students were randomly assigned to classrooms, simulated using permutations without replacement.
3. Construct $S_s = K^{-1} \sum_k 1[H_s > H_s^{random,k}]$: share of permutations where actual ability concentration is higher than under random assignment.

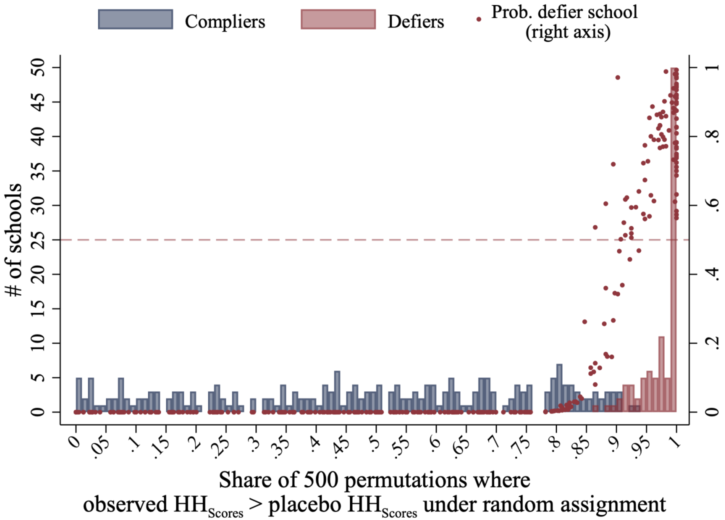
Using the shares S_s for $s = 1 \ldots S$:

4. Estimate the latent probability that school s is a "defier" school using finite mixture models (FMMs). Use school-level latent class predictors if available.
5. Flag defier school s as defier if, for its predicted posterior probabilities from the FMM, $Pr[\textit{defier class}] > Pr[\textit{complier class}]$.

Distribution of the Shares S_s



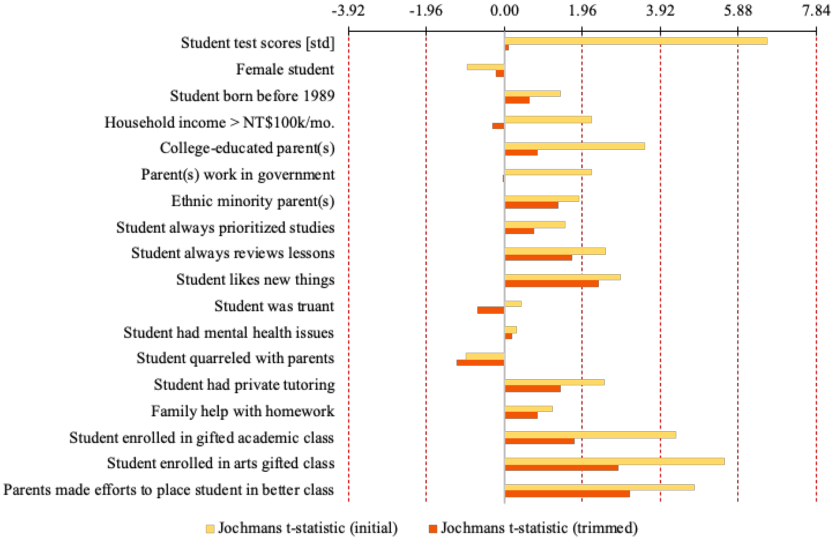
Classification of Schools by Fitting FFMs on S_s



Sample Means Before and After Trimming

Characteristics:	TEPS (1)	Trimmed (2)
Student test scores (unstandardized)	40.9	40.6
Female student	0.50	0.50
Student year of birth	1988.59	1988.59
No. of siblings of student	1.77	1.77
Responding parent is female	0.64	0.64
Ethnic minority father	0.05	0.05
Two-parent household	0.86	0.86
Father's birth year	1958.6	1958.7
Father has post-secondary education	0.12	0.12
Unemployed father	0.11	0.11
Household monthly income is		
NT$20,000 or less	0.11	0.11
NT$20,000-NT$50,000	0.41	0.41
NT$50,000-NT$100,000	0.35	0.35
More than NT$100,000	0.14	0.14
No. of students (approx.)	20,055	13,760

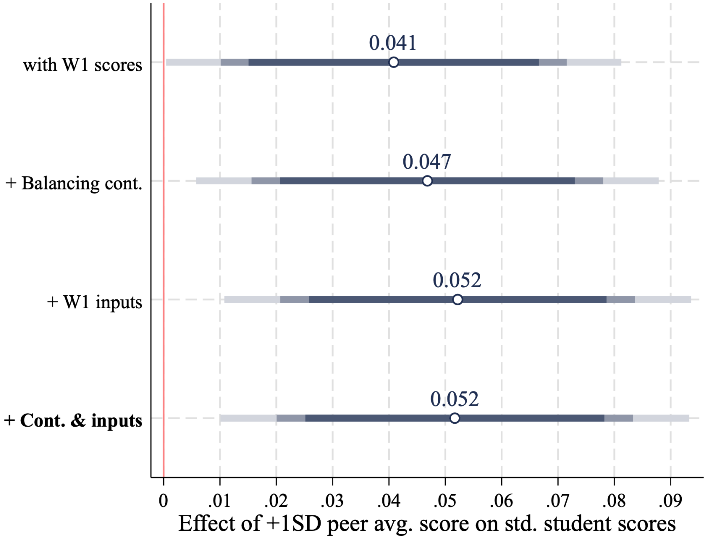
Jochmans (2020) Sorting Test Before and After Trimming



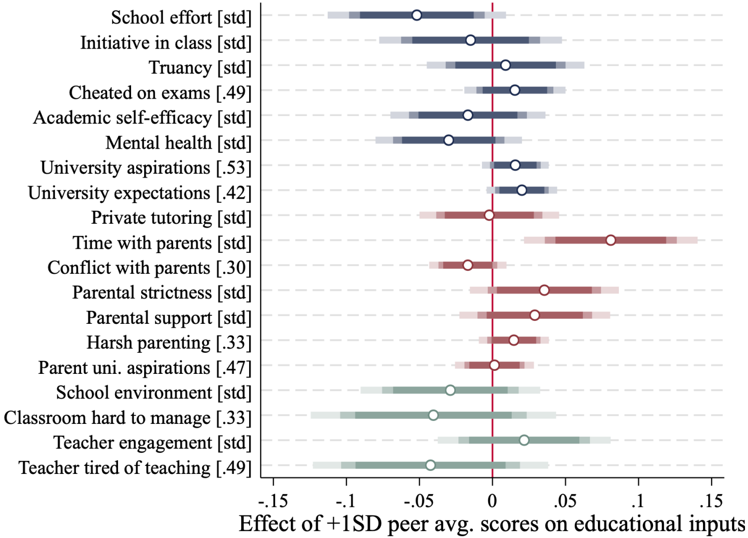
Balancing Tests After Trimming

	Peer ability leave-out-mean [std]	
Pre-assignment characteristics:	Coef. Est.	S.E.
Female student	0.008	(0.011)
Student born before 1989	−0.005	(0.010)
Monthly household income over NT$100,000	−0.019∗∗∗	(0.007)
College-educated parent(s)	0.001	(0.009)
Parent(s) work in government	0.010	(0.007)
Ethnic minority parent(s)	−0.004	(0.009)
Since primary school:		
Student always prioritized studies	−0.010	(0.009)
Student always reviews lessons	0.003	(0.008)
Student likes new things	−0.001	(0.011)
During primary school:		
Student was truant	0.000	(0.011)
Student had mental health issues	−0.004	(0.010)
Student quarreled with parents	−0.001	(0.009)
Before junior high school:		
Had private tutoring	0.004	(0.012)
Family helped with homework	−0.020∗∗	(0.008)
Student enrolled in gifted academic class	0.013	(0.008)
Student enrolled in arts gifted class	−0.013	(0.015)
Parents made efforts to place student in better class	0.035∗∗∗	(0.010)

Effects of Higher-Achieving Peers on Test Scores



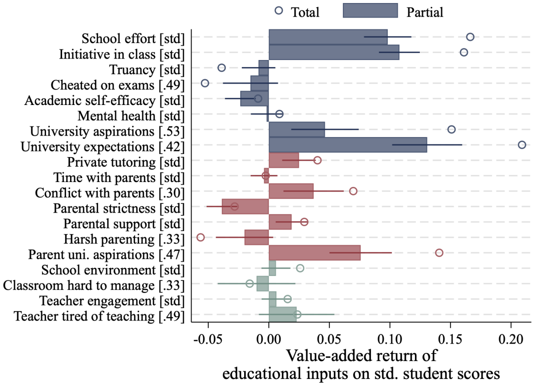
Effects of Higher-Achieving Peers on Inputs



Returns to Educational Inputs

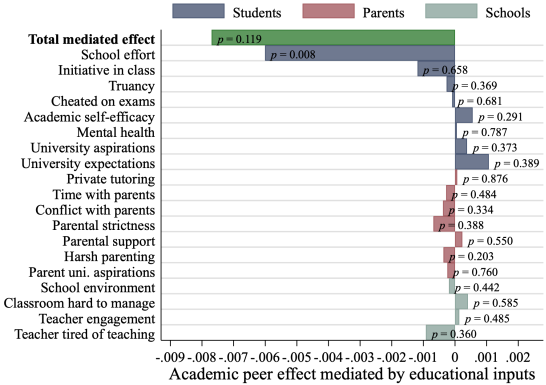
Estimated using cumulative value-added models (Todd & Wolpin, EJ 2003)

$$TestScores_{ics2} = \sum_k \beta^k y^k_{ics2} + \theta\, VaControls_{ics1} + \mu_s + \epsilon_{ics2}$$



Mediation Analysis: Test Scores through Changes in Inputs

$$\underbrace{ME}_{\substack{\text{total}\\\text{mediated}\\\text{effect}}} = \sum_k \underbrace{ME_k}_{\substack{\text{input } k\\\text{contribution}}} = \sum_k \underbrace{\frac{\partial S_{ics2}}{\partial y^k_{ics2}}}_{\substack{\text{input } k\\\text{average return}\\\text{on scores}}} \times \underbrace{\frac{\partial y^k_{ics2}}{\partial \bar{S}^{-i}_{cs1}}}_{\substack{\text{peer effect}\\\text{on input } k}}$$



Robustness of Findings

1. Identification assumptions
   - Permutation-based sorting test (Carrell & West JPE 2005) here
   - Non-parametric sorting test (Feld & Zölitz JOLE 2017) here
   - Alternative exclusion thresholds in Fishing Algorithm here
   - Proportional selection on unobservable to observable characteristics (Oster JBES 2019) here
2. Measurement error and incomplete classroom sampling
   - Ability measured with error here
   - Classical measurement error in peer ability here
   - Incomplete classroom sampling (Sojourner EJ 2013) here
3. Inference
   - Randomization inference (Young QJE 2019) here
   - Correction for multiple hypothesis testing (Romano & Wolf ECTA 2005)
4. Mediation analysis with heterogeneity
   - Heterogeneous direct and mediated effects across subgroups here

Conclusions

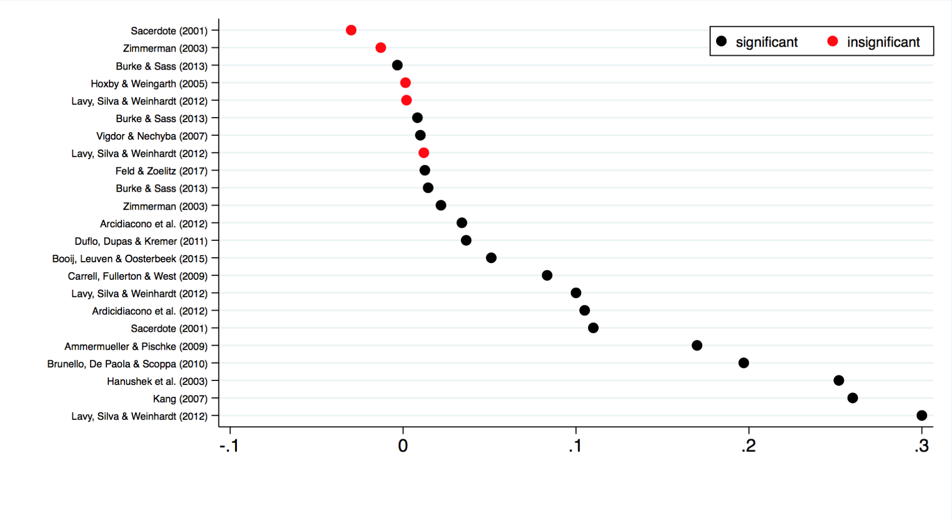
- We recover ability peer effect estimates from a natural quasi-experiment with imperfect compliance of random assignment
- Our fishing algorithm performs well and produces robust and believable peer effect estimates
- In the paper we:
  - show performance in simulated data
  - show sensitivity to model specification
  - show sensitivity to different FMM posterior probability thresholds
  - discuss how to extend to multivariate sorting
  - discuss how to generalize to many applications of cluster treatment non-compliance

Thank you!

Appendix

Peer Effects in Education: Estimated Effect Sizes

Back



Source: Zoelitz and Isphording, "The Value of A Peer", 2019

TEPS Measures and Index Construction Back

Measure	Description	Wave 1 Items \ Values	Wave 2 Items \ Values
Students:			
Test scores	Comprehensive Analytical Ability standardized test, measure of cognitive ability	70 questions	70 questions
School effort	Study effort in English, Chinese, Math class; homework on time in English, Chinese, Math class	7 \ 23	7 \ 25
Initiative in class	Initiative to ask and answer questions in English, Chinese and Math class	3 \ 12	3 \ 12
Cheating in exams	How often cheating in exams	1 \ 2	1 \ 2
Mental health	Feeling troubled, depressed, suicidal, nervous, issues focusing, pressured, irritated, isolated, guilty	6 \ 19	12 \ 22
Truancy	Skipping class, fighting, watching porn, drinking alcohol, stealing, running away from home	6 \ 19	4 \ 10
Academic self-efficacy	Focus, diligence, conscientiousness, initiative, eloquence, organization, cooperation, curiosity	7 \ 22	10 \ 19
University aspirations	Student wants to go to university	1 \ 2	1 \ 2
University expectations	Student expects to go to university	1 \ 2	1 \ 2
School environment	Study ethos, campus safety, fairness, engagement of school administrators	5 \ 16	5 \ 16
Parents:			
Money investments	Amount paid on out-of-school tutoring, how much tutoring	2 \ 10	3 \ 10
Time investments	Frequency going to bookstores and cultural events together	2 \ 7	2 \ 11
Parent-child conflict	Ever quarrel with father, ever quarrel with mother	2 \ 2	2 \ 2
Parental strictness	Father and mother's strict discipline	2 \ 7	2 \ 17
Parental support	Discuss future, listens carefully, worries and gives advice, accepts unconditionally	8 \ 25	8 \ 7
Harsh parenting	Parents use punishment	1 \ 2	1 \ 2
University aspirations	Parents want student to go to university	1 \ 2	1 \ 2
Teacher:			
Teacher engagement	knows names of students, encourages students who work hard, uses several different teaching materials, gives homework, cares about students, reviews questions after exams	6 \ 19	6 \ 36

- Confirm strong first component via PCAs

TEPS Indices Construction

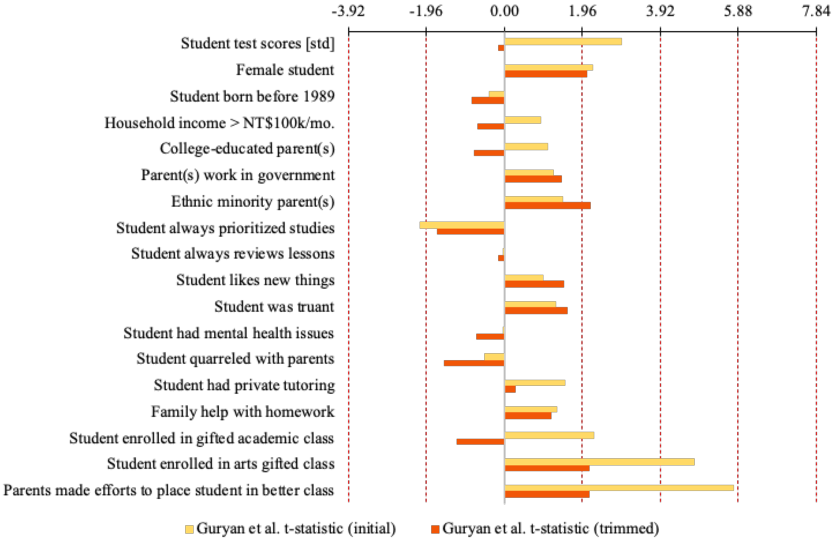
Back

Table: First Factor Eigenvalues

	Wave 1	Wave 2
Study Effort	3.46	3.62
Initiative in class	0.86	1.16
Mental Health	3.09	2.83
Truancy	2.77	0.79
Self-Efficacy	2.83	2.01
Parental Money	0.73	1.22
Parental Time	0.55	0.41
Parental Strictness	0.53	2.55
Parental Support	2.33	1.27
School Environment	1.86	1.34
Teacher Engagement	1.4	2.56

Guryan et al. (2009) Sorting Test

Back



Permutation tests (Lehmann and Romano, 2005) Back

For each pre-assignment characteristic and outcome of interest,

- Construct 10,000 simulated classrooms respecting school, number and size of classes.
- For each classroom, calculate classroom-mean of characteristic.
- Construct empirical p-value = share of simulated classrooms with lower mean than the mean in the realized classroom.
- Test school-by-school if empirical p-values are uniformly distributed using Kolmogorov-Smirnov and Chi-Square goodness of fit tests.

Permutation tests (Lehmann and Romano, 2005) Back

Pre-assignment characteristics:	Share of classes with empirical p-val. under 0.10 (1)	0.05 (2)	0.01 (3)	Avg. p-value (4)
Student test scores	0.10	0.06	0.02	0.486
Female student	0.06	0.04	0.02	0.562
Student born before 1989	0.10	0.05	0.01	0.490
Monthly household income over NT$100,000	0.09	0.04	0.01	0.491
College-educated parent(s)	0.09	0.06	0.02	0.485
Parent(s) work in government	0.08	0.04	0.01	0.487
Ethnic minority parent(s)	0.08	0.04	0.01	0.494
Since primary school:				
Student always prioritized studies	0.12	0.06	0.01	0.491
Student always reviews lessons	0.12	0.06	0.01	0.478
Student likes new things	0.13	0.07	0.02	0.465
During primary school:				
Student was truant	0.08	0.04	0.01	0.498
Student had mental health issues	0.10	0.06	0.02	0.495
Student quarreled with parents	0.10	0.05	0.01	0.503
Before junior high school:				
Student had private tutoring	0.11	0.06	0.01	0.479
Family help with homework	0.09	0.05	0.01	0.496
Student enrolled in gifted academic class	0.09	0.05	0.02	0.466
Student enrolled in arts gifted class	0.12	0.08	0.03	0.447
Parents made efforts to place student in better class	0.13	0.07	0.02	0.465

Non-Parametric Sorting Test

Back

- For each school, run a regression of each characteristic on classroom dummies, then F-tests of joint significance
- Calculate share of F-test p-values falling under nominal level of test
- Under perfect balancing, we should find a uniform distribution of the share of p-values falling under each significance level

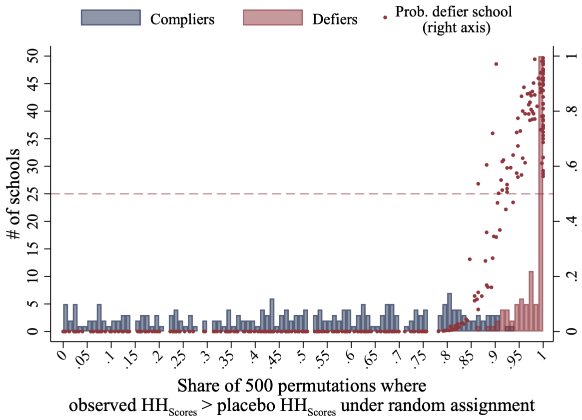
Non-Parametric Sorting Test

Back

Pre-assignment characteristics:	No. school-reg.	Share of class-dummy joint significance test p-val. under: 0.10	0.05	0.01
Student test scores	227	0.06	0.04	0.03
Female student	216	0.05	0.02	0.02
Student born before 1989	227	0.12	0.03	0.01
Monthly household income over NT$100,000	208	0.09	0.04	0.00
College-educated parent(s)	204	0.13	0.07	0.02
Parent(s) work in government	205	0.06	0.02	0.01
Ethnic minority parent(s)	179	0.06	0.02	0.01
Since primary school:				
Student always prioritized studies	227	0.12	0.06	0.01
Student always reviews lessons	227	0.10	0.06	0.02
Student likes new things	227	0.14	0.10	0.02
During primary school:				
Student was truant	227	0.10	0.03	0.01
Student had mental health issues	227	0.12	0.07	0.01
Student quarreled with parents	227	0.10	0.04	0.00
Before junior high school:				
Had private tutoring	227	0.13	0.08	0.02
Family help with homework	226	0.08	0.06	0.02
Student enrolled in gifted academic class	206	0.11	0.05	0.02
Student enrolled in arts gifted class	186	0.15	0.09	0.07
Parents made efforts to place student in better class	225	0.14	0.10	0.04

Fishing Algorithm: Alternative Thresholds

Back



- We classified schools as "Latent Defier" if $P(\text{defier}) > P(\text{complier})$
- An alternative way is to assign based on $P(\text{defier} > T)$, where T is a decision to make.
- Results stable for $T \in [0.5, 1[$

Oster (2019) Proportional Selection on Unobservables

Back

Outcomes	Degree of selection required to explain effect of peer test scores on outcomes
Test scores	-0.20
School effort	-0.10
Initiative in class	-3.00
Truancy	2.40
Cheated on exams	-0.50
Academic self-efficacy	-5.70
Mental health	-0.70
University aspirations	-1.00
University expectations	-1.10
Private tutoring	0.10
Time with parents	-0.40
Conflict with parents	-3.40
Parental strictness	-1.60
Parental support	-0.50
Harsh parenting	-4.80
Parent uni. aspirations	-0.00
School environment	-0.70
Classroom hard to manage	-0.10
Teacher engagement	-0.90
Teacher tired of teaching	-0.00
Selection proportional to:	
Balancing controls	Y
W1 inputs	Y

Alternative Measures of Ability

Back

			IRT Bayesian posterior mean of:		
	Analytical	Mathematical	General	Analytical	Mathematical
	(1)	(2)	(3)	(4)	(5)
Peer ability [std]	0.042**	0.046***	0.048***	0.043**	0.047***
	(0.018)	(0.016)	(0.017)	(0.019)	(0.017)
Own ability [std]	0.389***	0.542***	0.606***	0.396***	0.558***
	(0.010)	(0.009)	(0.009)	(0.010)	(0.009)
R^2	0.46	0.61	0.70	0.49	0.64

Classical Measurement Error in Peer Ability: Mixed IV

Back

	Measure of Ability used:		
	Analytical (1)	Mathematical (2)	Mixed (3)
Peer ability [std]	0.054*	0.042*	0.068*
	(0.029)	(0.025)	(0.029)
Instrument used:	Mathematical	Analytical	Alt. mixed
t-statistic of first-stage coefficient	30.53	28.25	27.01

Mixed IV: Effect of Peer Ability on Educational Inputs

Back

	Mixed IV effect of peer ability [std]	
	Coef.Est.	Std. err.
Outcomes	(1)	(2)
School effort	−0.072**	(0.034)
Initiative in class	−0.002	(0.038)
Truancy	0.019	(0.034)
Cheated on exams	0.013	(0.022)
Academic self-efficacy	−0.014	(0.034)
Mental health	−0.052	(0.034)
University aspirations	0.022	(0.014)
University expectations	0.021	(0.016)
Private tutoring	−0.006	(0.031)
Time with parents	0.113***	(0.038)
Conflict with parents	−0.016	(0.017)
Parental strictness	0.057*	(0.031)
Parental support	0.037	(0.031)
Harsh parenting	0.025*	(0.014)
Parent uni. aspirations	0.011	(0.017)
School environment	−0.008	(0.037)
Classroom hard to manage	−0.026	(0.050)
Teacher engagement	0.022	(0.035)
Teacher tired of teaching	−0.066	(0.048)

Sojourner (2013): Incomplete Classroom Sampling Back

- Only partial picture of classrooms due to sampling design.
- Sojourner (2013): If students are randomly assigned to classrooms,
  - Missing at random $\Rightarrow$ attenuated peer effect estimates - akin to classical measurement error.
- Solution: Weight estimates by classroom sampling rate...
  - interacted with school-level sampling rate (preferred but restrictive)
  - interacted with school-level sampling rate in pre-determined clusters
- Our results using Sojourner's correction:
  - Test scores: 8.9 to 13.3%SD
  - Study effort: no significant effect
  - Aspirations and expectations: from 2.8 to 5.0 p.p.
  - Parental time investment: 9.8 to 16.8%SD
  - Classroom hard to manage: 11.3 to 15.3 p.p.
  - Teacher tired of teaching: 9.1 to 16.6 p.p.
- Consistent with random assignment and missing at random.

Corrected P-Values: Randomization Inference & Multiple Hypotheses Testing

Back

Outcomes	Young (2019) Randomization-t inference (1)	Romano & Wolf (2005) Step-down procedure (2)
Test scores	**0.008**	**0.036**
School effort	**0.058**	0.382
Initiative in class	0.589	0.948
Truancy	0.673	0.956
Cheated on exams	0.287	0.860
Academic self-efficacy	0.487	0.914
Mental health	0.176	0.774
University aspirations	0.122	0.680
University expectations	**0.078**	0.394
Private tutoring	0.928	0.982
Time with parents	**0.000**	**0.014**
Conflict with parents	0.146	0.742
Parental strictness	0.136	0.670
Parental support	0.220	0.800
Harsh parenting	0.192	0.774
Parent uni. Aspirations	0.916	0.982
School environment	0.305	0.860
Classroom hard to manage	0.248	0.858
Teacher engagement	0.421	0.890
Teacher tired of teaching	0.216	0.838

Mediation with heterogeneity Back

- Mediation results could suffer from heterogeneity in 3 ways:
  1. Heterogeneity of main effects (e.g. Lavy, Silva and Weinhardt, 2012)
     - Middle and higher ability students
     - Less experienced homeroom teachers
     - No heterogeneity in parental income, education, and gender
  2. Heterogeneity in drivers of academic peer effects
  3. Heterogeneity in returns to educational inputs
- Overall little heterogeneity, and still no mediation.