How Far Can Inclusion Go? The Long-term Impacts of Preferential College Admissions

Michela Carlana[1] Sara Chiuri[2] Enrico Miglino[3]

Michela M. Tincani[4]

[1] Harvard Kennedy School, Bocconi University, NBER, LEAP, CEPR
[2] Bocconi University
[3] Bank of Italy
[4] University College London, IFS, LEAP, CESifo, CEPR

NBER Summer Institute
July 23rd 2025

Funded by the European Union (ERC, SOFIA, 101117537). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them.

Section 1

Motivation

The Promises of Preferential Admissions

- Disparities in human capital greatly contribute to the socioeconomic gap in higher education attainment (Bailey and Dynarski, 2011; Belley and Lochner, 2007).
- Students from disadvantaged background have fewer learning opportunities in early life (Bohren et al., 2023; Cunha and Heckman, 2007).
- Affirmative action aims at achieving diversity by enhancing admission chances of disadvantaged students.
- Preferential admissions to selective colleges have lasting benefits for well-prepared under-represented students (Black et al. 2023, Bleemer 2022).
- **But how far can inclusion go while delivering on its promises?**

The Future of College Admission

The New York Times

With End of Affirmative Action, a Push for a New Tool: Adversity Scores

To build a diverse class of students, the medical school at U.C. Davis ranks applicants by the disadvantages they have faced. Can it work nationally?

The Washington Post
Democracy Dies in Darkness

Without affirmative action, how will colleges seek racial diversity?

None of the options — from eliminating legacy preferences to reducing slots for athletes — is simple

By Nick Anderson and Susan Svriluga
Updated July 5, 2023 at 3:18 p.m. EDT | Published July 1, 2023 at 6:00 a.m. EDT

THE TEXAS TRIBUNE

UT-Austin Automatic Admissions Standard for 2017: Top 7 Percent

> Texas students who apply to UT-Austin for the fall 2017 semester will need to be in the top 7 percent of their high school class to gain automatic admission, which is tougher than the 2016 threshold.

This paper

Research Question

- What are the education and labour market impacts of affirmative action on **targeted disadvantaged students further down the academic ability distribution**?

  → Desegregating college sector requires relaxing academic requirements (Chetty, Friedman, Saez, Turner and Yagan, 2020).

This paper: preferential college admission in Chile

We evaluate the impact of preferential college admissions in Chile targeting disadvantaged high schools.

Why Chile?

- Socioeconomic gaps in college attainment similar to other industrialized countries
- A preferential admission policy, PACE, that targets students with much lower pre-college achievement than studied before
- High-quality data

Empirical methodologies:

- **RCT**: random inclusion of high schools in PACE in 2016
- **RDD**: sharp within-school cutoffs for preferential admission

Contributions to the literature

- Clear benefits for well-prepared targeted disadvantaged students (Black et al., 2023; Bleemer 2022).

  → We focus on the impact on substantially less prepared students.

- Preferential admission and the mismatch hypothesis (Sander 2004; Arcidiacono et al., 2011; Imbens et al., 2012, Badge, et al., 2016).

  → We provide a direct test of the mismatch hypothesis exploiting long-term outcomes and measures of decision-making frictions.

- RDD away from the cutoff (Angrist and Rokkanen, 2015, Palomba, 2024, Fort, Ichino, Rettore, and Zanella, 2022, Cingano et al., 2024).

  → We provide a unique connection between the RDD and RCT.

Section 2

Context

Higher Education System in Chile

Higher education institutions are classified as follows:

- **Selective colleges**, to enroll students have to:
  - take the standardized college entrance exam (PSU)
  - submit an application on the centralized platform
- **Non-selective institutions**, entrance exam not required:
  - off-platform colleges
  - vocational institutions offering 2-3 year degrees

Graduation rates by type

PACE Program

Programa de Acompañamiento y Acceso Efectivo a la Educación Superior

1. **In addition to regular process, guaranteed admissions the following year to selective colleges** for students:
   - attending a PACE high school in the last two years
   - taking the college entrance exam
   - having a GPA in the top 15% of their high school
2. **Orientation sessions** on college applications, tuition and financial aid, tips on study techniques
   - offered to all students in PACE high schools
3. **Optional tutoring** in college
   - for students offered a PACE admission to a selective college (limited take-up)

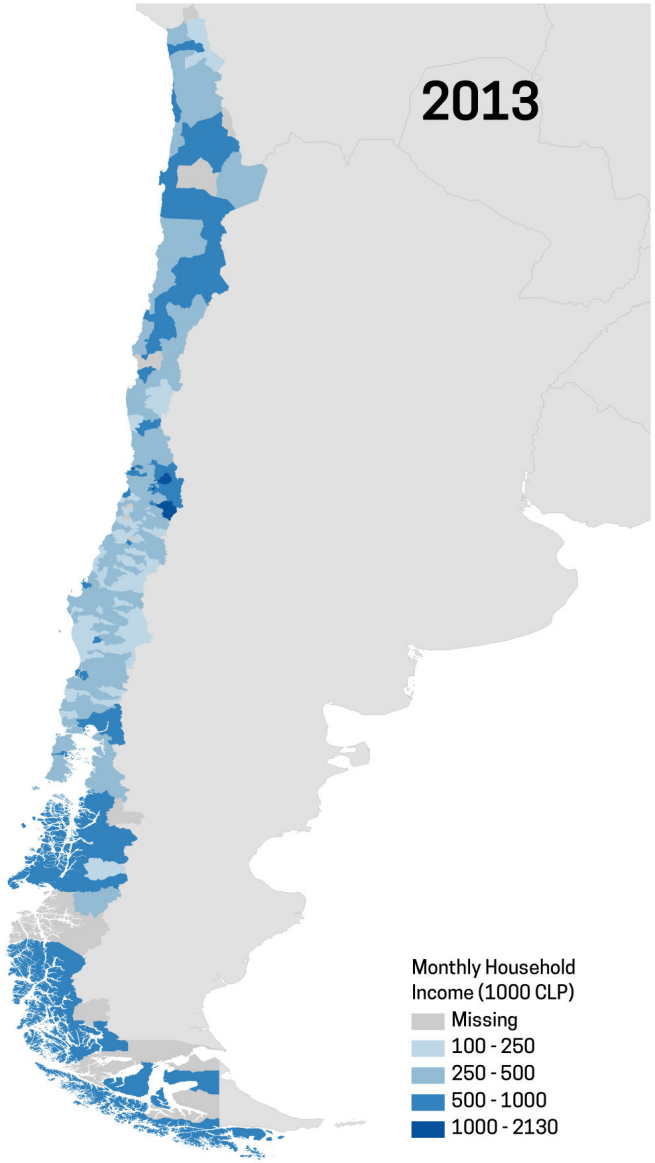
Seats

PACE Program expansion

Disadvantaged high schools:
based on a school-level vulnerability index

Expansion:
from 69 high schools in 2014
to 456 high schools in 2016

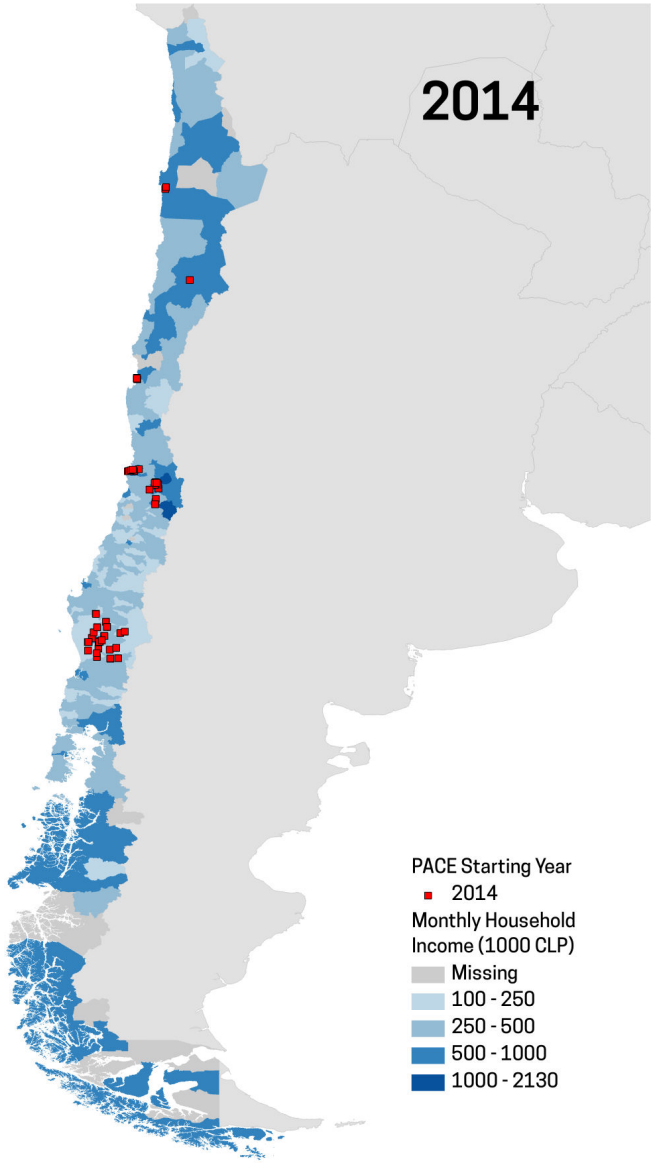
[Link to video]



PACE Program expansion

Disadvantaged high schools:
based on a school-level vulnerability index

Expansion:
from 69 high schools in 2014
to 456 high schools in 2016



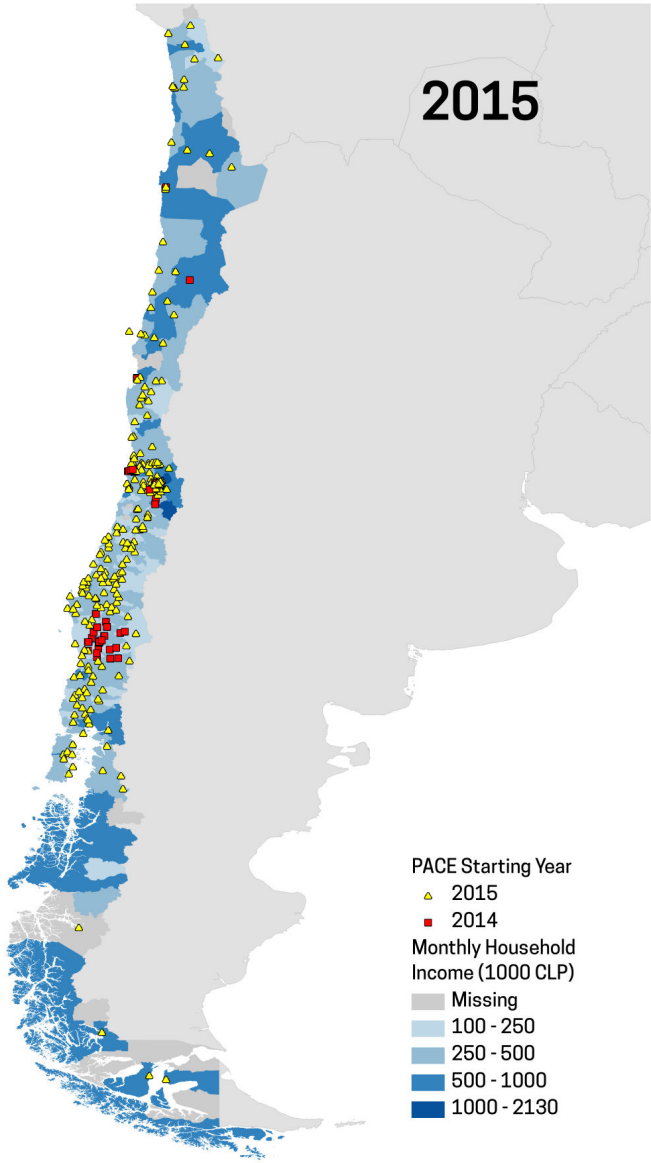
Link to video

PACE Program expansion

Disadvantaged high schools:
based on a school-level vulnerability index

Expansion:
from 69 high schools in 2014
to 456 high schools in 2016

[Link to video]

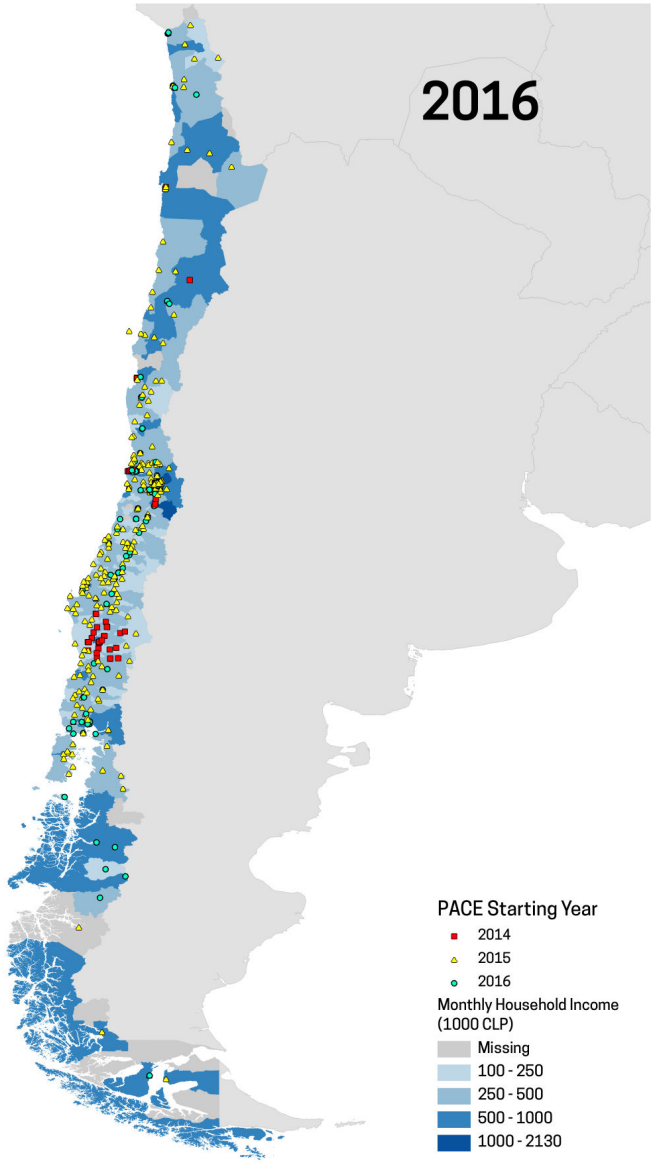


PACE Program expansion

Disadvantaged high schools:
based on a school-level vulnerability index

Expansion:
from 69 high schools in 2014
to 456 high schools in 2016

[Link to video]



Section 3

Empirical Framework

Research Design 1: Randomized Controlled Trial

In 2016, the government randomly selected 64 of the 221 newly eligible schools to receive the PACE treatment.

$$Y_{is} = \alpha + \beta PACE_s + \lambda X_i + \varepsilon_{is} \quad (1)$$

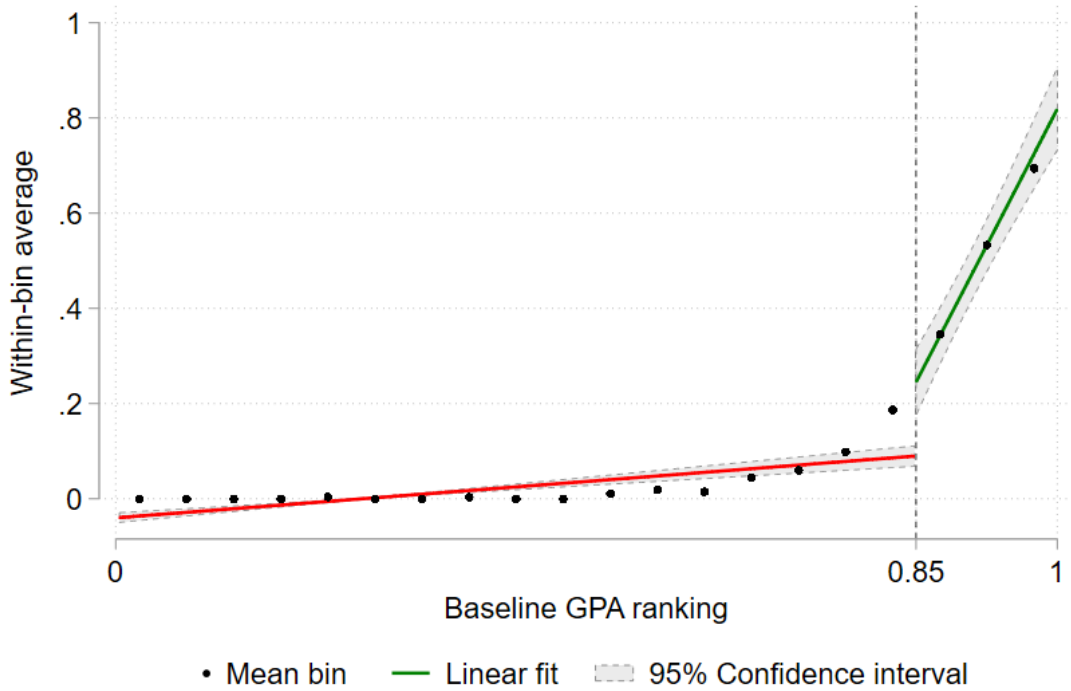
We study the effect of being in a **PACE school** for top-performing students based on baseline GPA:

- **Top 15%:** Combined effect of school treatment (orientation session, rank info, changed incentives) and individual treatment for some (college admission).

Balance

Research Design 1: Randomized Controlled Trial

Preferential college admission offers



54.2% of students in the top 15% of GPA and
2.7% of students in the bottom 85% of GPA get a PACE offer

Research Design 2: Regression Discontinuity Design

We exploit the top 15% cutoff in schools participating in PACE for the first three cohorts (2014-2016)

$$Y_{is} = \gamma + \delta A_{is} + f(p_{is} - c_s) + \theta X_i + \eta_s + \nu_{is} \quad (2)$$

and we instrument A_i using the following first-stage regression:

$$A_{is} = \zeta + \phi I(p_{is} - c_s \geq 0) + g(p_{is} - c_s) + \psi X_i + \xi_s + u_{is} \quad (3)$$

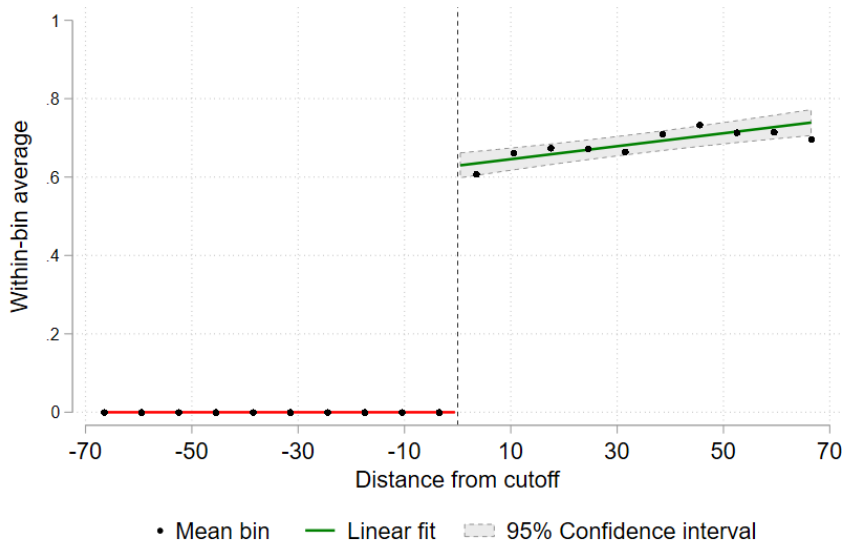
where:

- A_{is} is equal to 1 if student i from school s receives a PACE admission to selective college
- $p_{is} - c_s$ is the difference between the GPA and the cutoff

→ Local average effect of preferential college admission for marginal students in PACE high schools. McCrary Test Balance

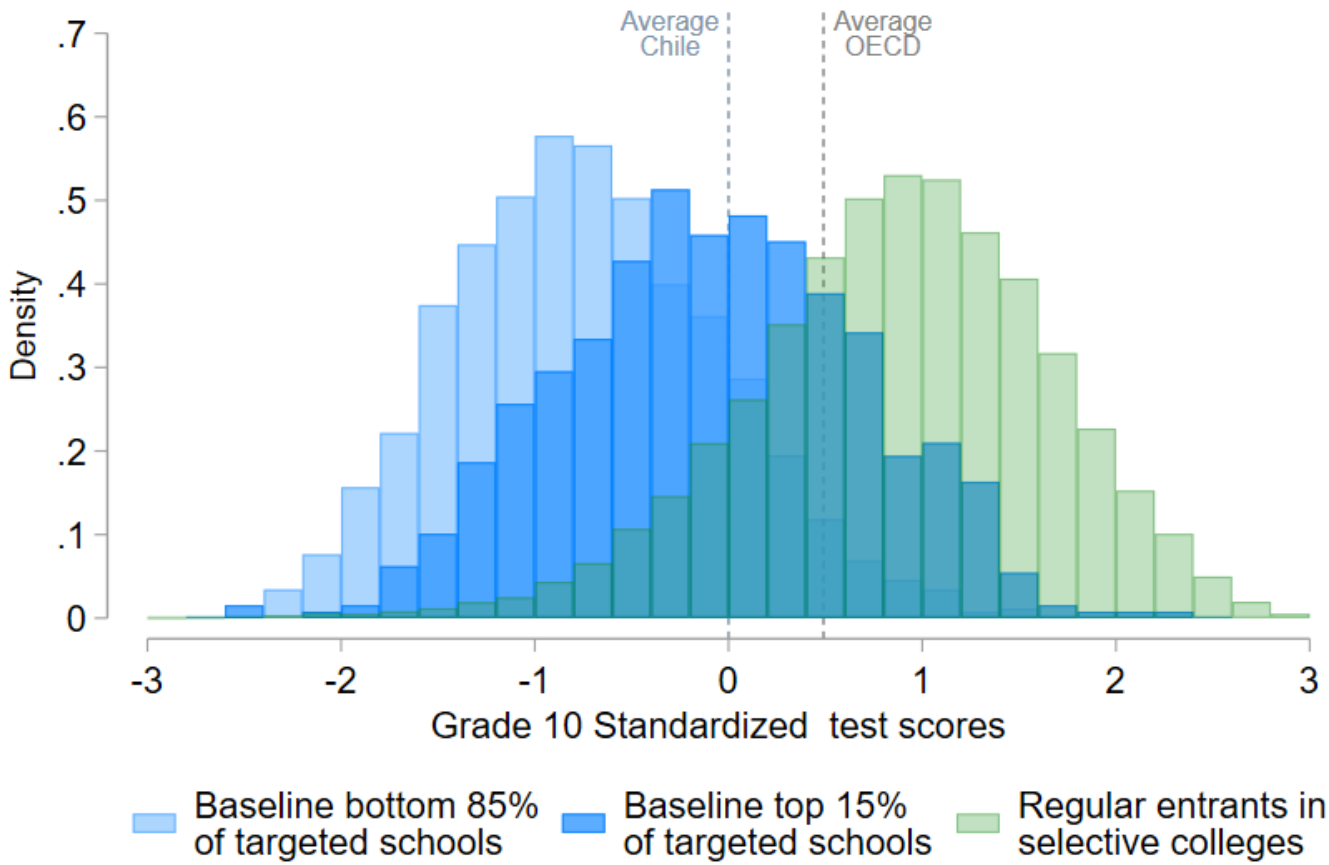
Research Design 2: RDD

First Stage: College admission offer



Optimal bandwidth (Calonico et al., 2014). S.e. clustered at school level.

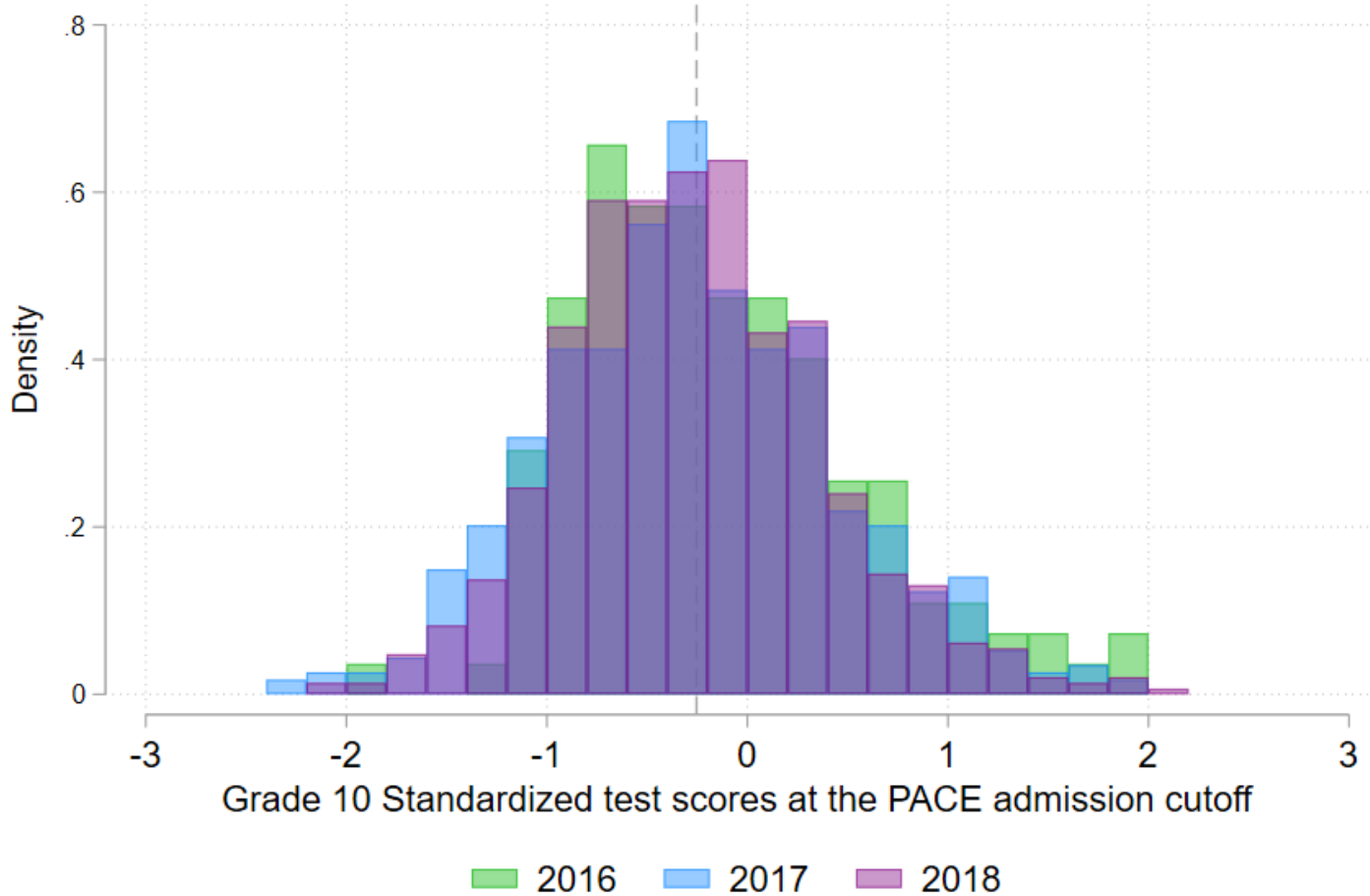
RCT: Students' academic preparation



- **PACE (Top 15%)**: 51^{th} percentile nationwide (Grade 10 test scores)
- **Texas Top Ten (Low-SES, Top 10%)**: 86^{th} percentile statewide (Grade 10 test scores, Black et al. 2023)

Other characteristics

RDD: Students' academic preparation



- **PACE (Marginal Students)**: 34^{th} percentile nationwide among entrance exam takers
- **California ELC (Marginal Students)**: 72^{nd} percentile nationwide among entrance exam takers (Bleemer, 2022)

Data Sources

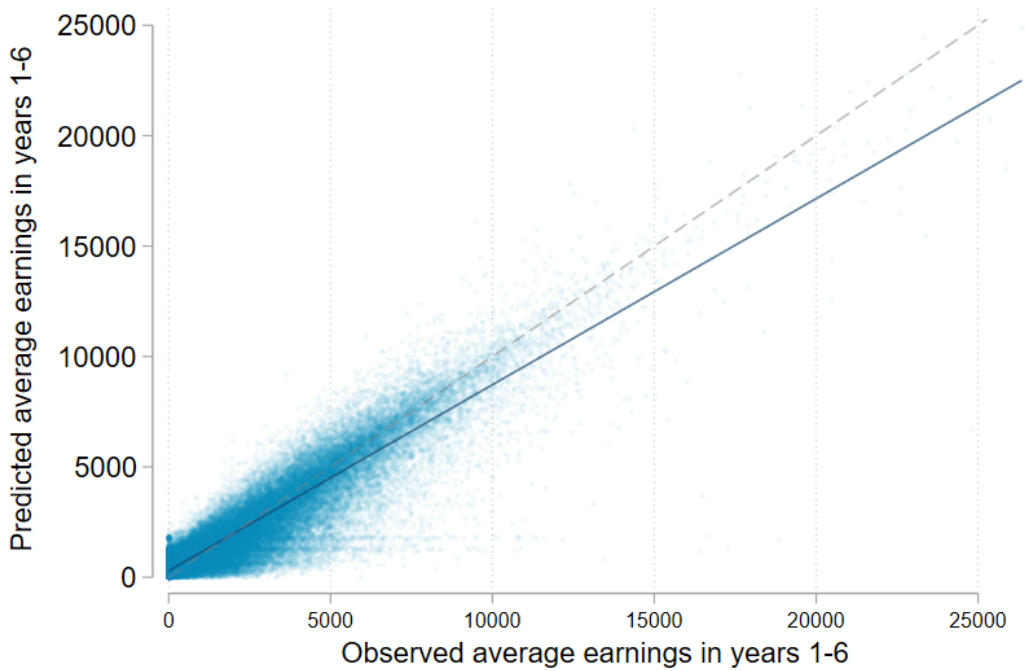
1. **Education data:** from MINEDUC, high school information, higher education enrollment, type of institution, study field, graduation.
2. **Labor market data:** unemployment insurance dataset (excluding self-employed and public sector employees).
3. **Survey data:** RCT sample of high school students in 2017, including prior beliefs on college attainment.

Study sample: 28,458 students from 606 high schools.

Surrogate index

- We predict long-term effects using a surrogate index computed from older cohorts in the same schools

  (Athey, Chetty, Imbens, and Kang, 2019)



- High correlation in year 6 between earnings predicted using the surrogate index and observed earnings

Section 4

Targeted Students

Conceptual Framework: Effect on Targeted Students

Preferential admissions can affect different margins of choice:

- **Intensive**: Induce to attend more selective programs instead of less selective ones.
- **Extensive**: Induce to attend higher education instead of working.

They can have different impacts (Arcidiacono and Lovenheim, 2016):

- **Quality effects:** large monetary inputs, high-quality professors and peers, networks $\rightarrow$ **Positive**
- **Match effects:** exams hard to pass, demanding, fast pace $\rightarrow$ **Negative (info frictions)**

We estimate the causal effect on long-term outcomes

- for those marginally eligible (RD) vs for all eligible (RCT top 15%)
- for those with different subjective expectations

Conceptual Framework: Effect on Targeted Students

Preferential admissions can affect different margins of choice:

- **Intensive**: Induce to attend more selective programs instead of less selective ones.
- **Extensive**: Induce to attend higher education instead of working.

They can have different impacts (Arcidiacono and Lovenheim, 2016):

- **Quality effects:** large monetary inputs, high-quality professors and peers, networks $\rightarrow$ **Positive**
- **Match effects:** exams hard to pass, demanding, fast pace $\rightarrow$ **Negative (info frictions)**

We estimate the causal effect on long-term outcomes

- for those marginally eligible (RD) vs for all eligible (RCT top 15%)
- for those with different subjective expectations

Conceptual Framework: Effect on Targeted Students

Preferential admissions can affect different margins of choice:

- **Intensive**: Induce to attend more selective programs instead of less selective ones.
- **Extensive**: Induce to attend higher education instead of working.

They can have different impacts (Arcidiacono and Lovenheim, 2016):

- **Quality effects:** large monetary inputs, high-quality professors and peers, networks $\rightarrow$ **Positive**
- **Match effects:** exams hard to pass, demanding, fast pace $\rightarrow$ **Negative (info frictions)**

We estimate the causal effect on long-term outcomes

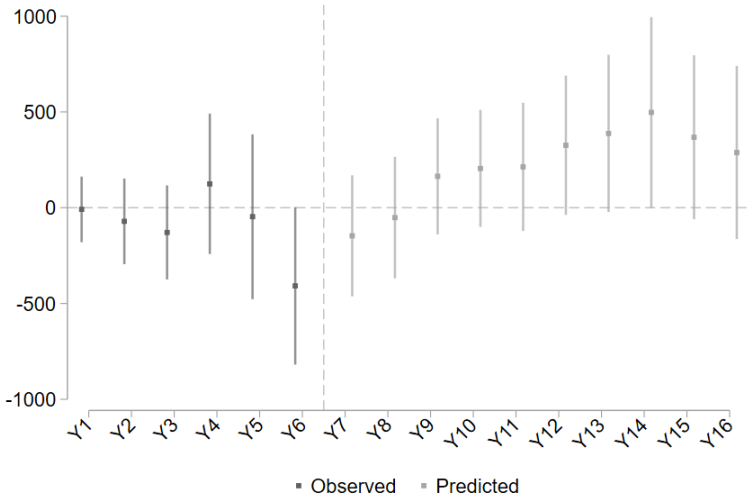
- for those marginally eligible (RD) vs for all eligible (RCT top 15%)
- for those with different subjective expectations

RCT: Impact on education for top 15%

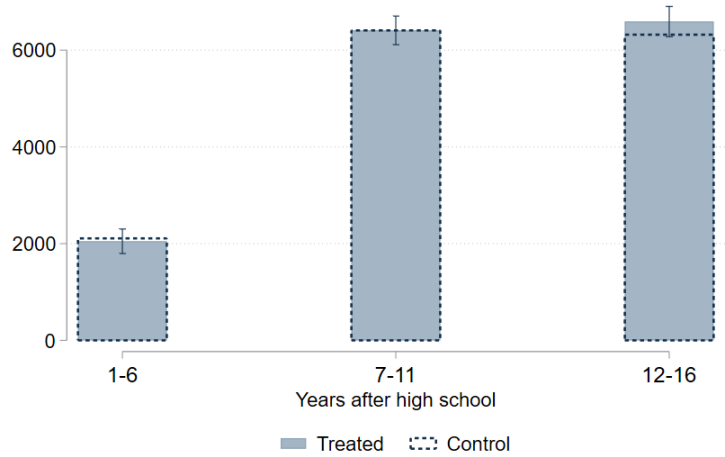
	(1)	(2)	(3)	(4)	(5)	(6)
	Selective College			Any Institution		
	Ever enrolled	Graduation	Dropout	Ever enrolled	Graduation	Dropout
Treatment	0.114***	0.072**	0.035**	0.010	0.006	0.003
	(0.032)	(0.029)	(0.014)	(0.015)	(0.021)	(0.018)
Total obs.	2437	2437	2437	2437	2437	2437
Mean	0.423	0.301	0.074	0.860	0.656	0.205

- No extensive margin effect on higher education attendance.
- Increased enrollment in selective colleges (↑ 27%), partly persisting into graduation (↑ 24%).
- Increased dropouts from selective colleges (↑ 47%), but not overall.
- PACE likely had intensive margin effects. Selectivity, ability distance, rank Majors

RCT: Impact on earnings for top 15%



(a) Earnings, in a given year



(b) Earnings, average across years

We do not observe short-term effects on earnings (-9.3% loss in year 6, 2.9% insignificant loss in years 1-6), and based on outcomes in years 1-6 we predict an insignificant positive effect on long-term earnings (+2.9% in years 12-16).

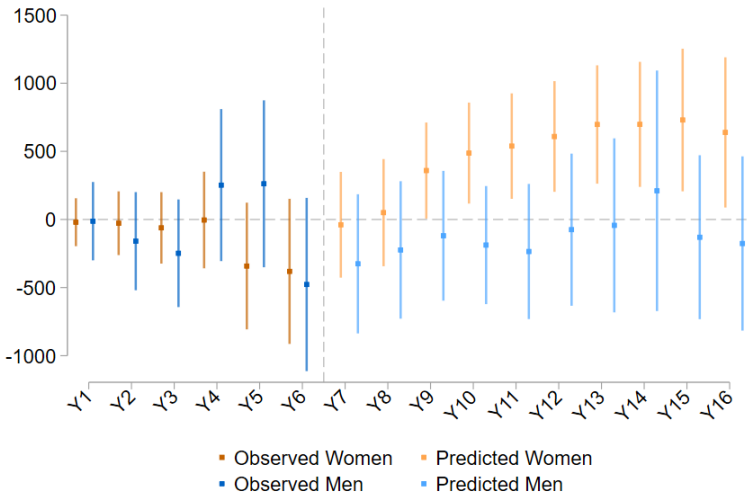
Labor force participation | Validation of surrogate index

RCT: College degree gains for top 15% driven by women

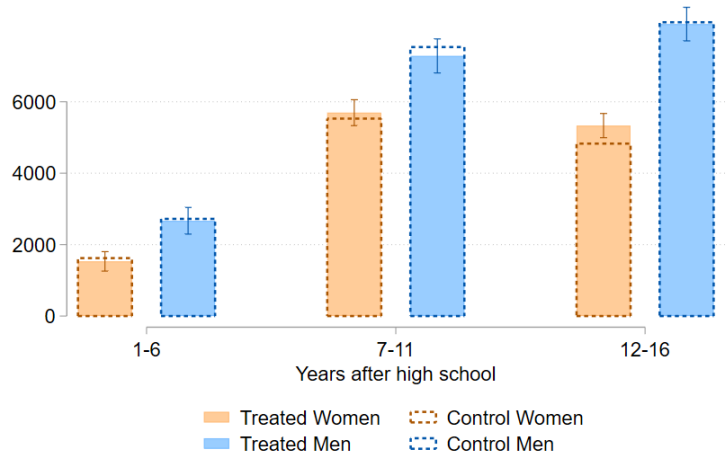
	(1)	(2)	(3)	(4)	(5)	(6)
	Selective College			Any Institution		
	Ever enrolled	Graduation	Dropout	Ever enrolled	Graduation	Dropout
	Heterogeneity by gender					
Treatment – Women	0.153***	0.120***	0.034**	0.032*	0.032	-0.001
	(0.034)	(0.033)	(0.016)	(0.017)	(0.025)	(0.020)
Mean Women	0.437	0.329	0.063	0.879	0.709	0.171
Total obs.	1369	1369	1369	1369	1369	1369
Treatment – Men	0.067	0.013	0.037*	-0.015	-0.026	0.011
	(0.044)	(0.039)	(0.022)	(0.022)	(0.029)	(0.030)
Mean Men	0.405	0.267	0.087	0.837	0.589	0.247
Total obs.	1068	1068	1068	1068	1068	1068

Selectivity, ability distance, rank Majors

RCT: Earning gains for top 15% women



(a) Earnings, in a given year



(b) Earnings, average across years

- For women, we predict a positive effect on long-term earnings (↑ 10.4% in years 12-16).
- For men, no evidence of earning impacts.
- Gender gap in long-term earnings reduced by 17%.

Labor force participation

Summary of results for RCT top 15% sample

- No extensive margin effect on higher education enrollment.
- Larger number of selective college degrees, that does not come at cost of more dropouts from higher education.
- Null labor market returns in the long term, on average.
- Substantial positive impacts on the long-term earnings of women, who are the students gaining more selective college degrees. No evidence that men take up the college opportunities nor that they benefit in the labor market.

$\rightarrow$ Earning and education gains for women, not for men.

RDD: Impact on education for marginal students

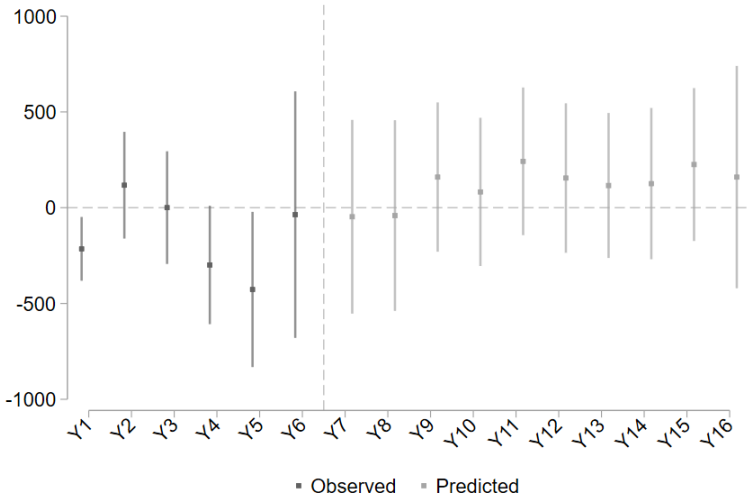
	(1)	(2)	(3)	(4)	(5)	(6)
	Selective College			Any Institution		
	Ever enrolled	Graduation	Dropout	Ever enrolled	Graduation	Dropout
Robust	0.182***	0.090***	0.100***	0.062***	-0.025	0.070***
	(0.030)	(0.023)	(0.018)	(0.020)	(0.030)	(0.025)
Bandwidth	36	57	62	60	50	54
Bandwidth obs.	7355	11428	12403	11946	9990	10731
Mean	0.401	0.275	0.071	0.851	0.658	0.195

- Extensive margin effect on higher education enrollment ($\uparrow$ 7.3%).
- Increased enrollment in selective colleges ($\uparrow$ 45%), partly persisting into graduation ($\uparrow$ 33%).
- Increased dropouts from selective colleges ($\uparrow$ 141%), and overall ($\uparrow$ 36%).
  - For every additional selective college graduate, $\sim$ 0.8 additional dropouts from higher education.
- PACE likely had intensive and extensive margin effects.

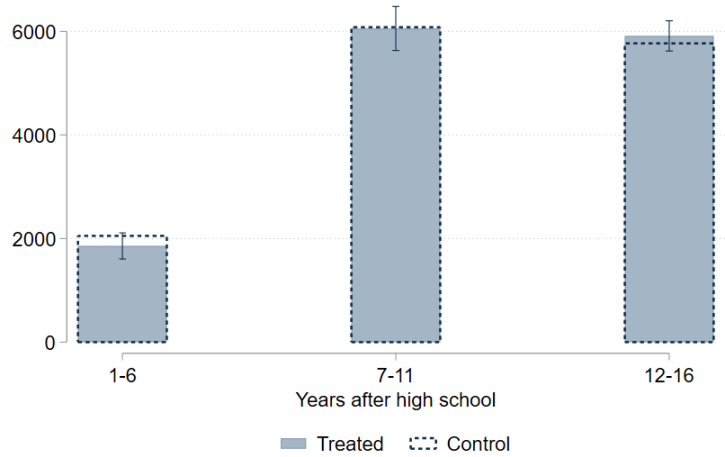
Selectivity, ability distance, rank | Majors

$\rightarrow$ Heterogeneous impacts; who benefits?

RDD: Impact on earnings for marginal students



(a) Earnings, in a given year



(b) Earnings, 6-year average

We estimate short-term negative impacts (↓ 9.5%) on marginal students' earnings, not entirely offset by year 16.
Accompanied by short-term significant reductions (↓ 8.6%) in months worked.

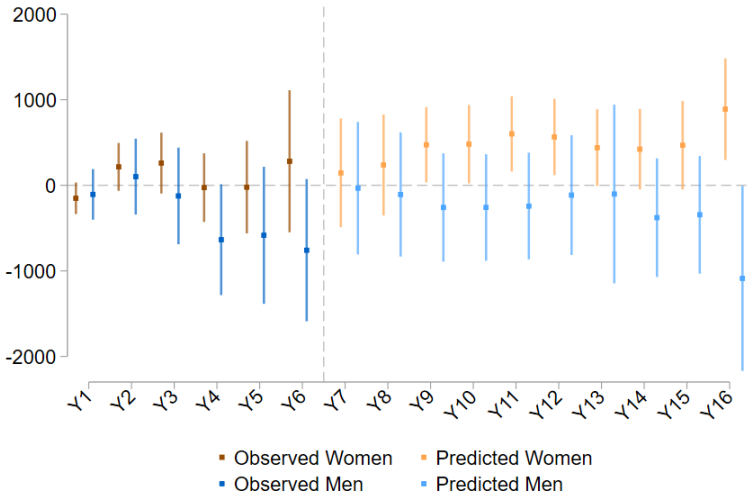
Labor force participation Validation of surrogate index

RDD: Dropouts from higher education concentrated among men

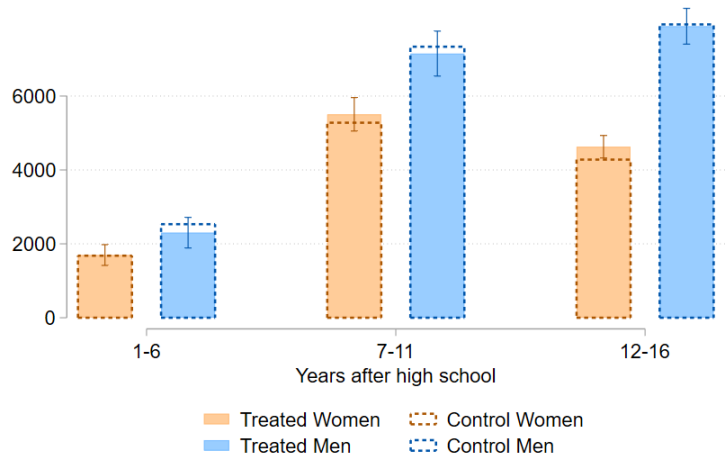
	(1)	(2)	(3)	(4)	(5)	(6)
	Selective College			Any Institution		
	Ever enrolled	Graduation	Dropout	Ever enrolled	Graduation	Dropout
	Heterogeneity by gender					
Robust - Women	0.180***	0.083***	0.085***	0.066***	-0.001	0.025
	(0.040)	(0.032)	(0.022)	(0.025)	(0.038)	(0.029)
Bandwidth	35	48	57	47	43	61
Bandwidth obs.	4261	5717	6717	5606	5172	7264
Mean	0.418	0.284	0.062	0.867	0.690	0.186
Robust - Men	0.214***	0.074**	0.136***	0.023	-0.059	0.108***
	(0.035)	(0.036)	(0.031)	(0.032)	(0.046)	(0.042)
Bandwidth	66	63	55	72	53	52
Bandwidth obs.	5370	5073	4486	5799	4291	4255
Mean	0.411	0.280	0.086	0.829	0.610	0.221
p-value Women=Men	0.549	0.906	0.263	0.289	0.395	0.154

Selectivity, ability distance, rank Majors

RDD: Earning losses driven by men; women gain



(a) Earnings, in a given year



(b) Earnings, 6-year average

- For men, we observe a 9.1% reduction in earnings in the short term, not offset by year 16.
- For women, we do not observe short-term losses, and we predict positive effects on medium-term (↑ 4.3%) and long-term earnings (↑ 8.1%).

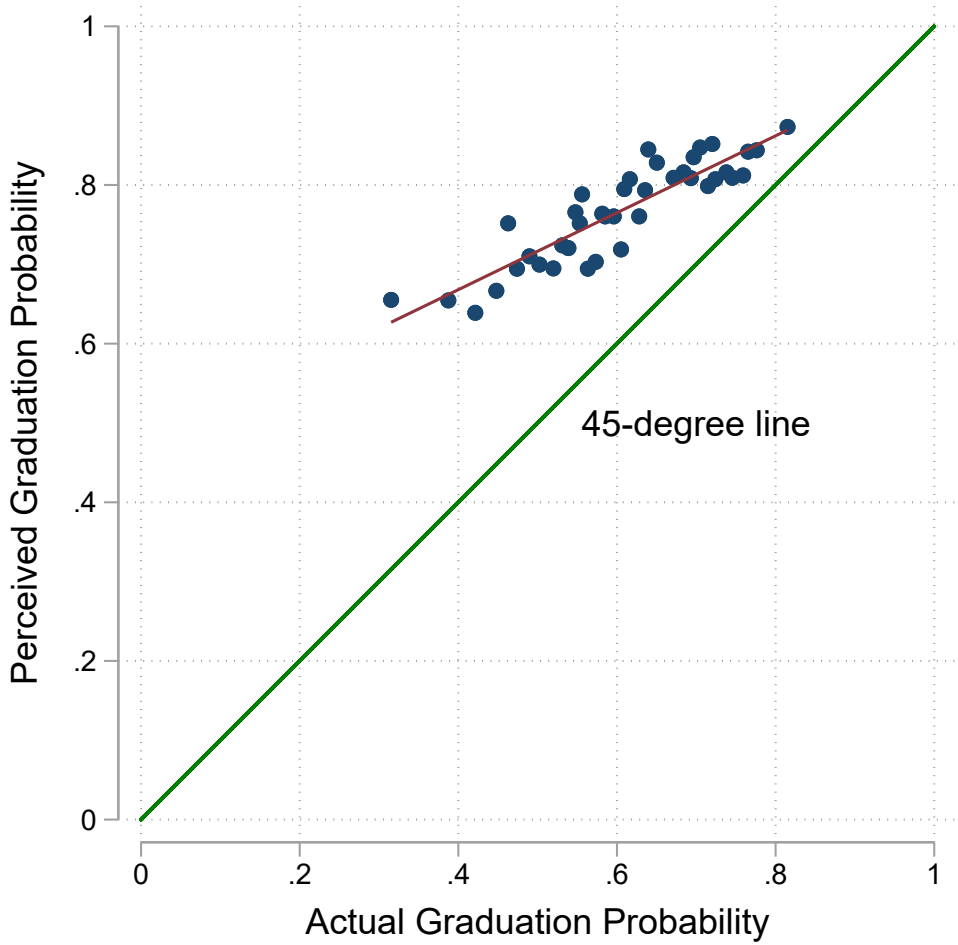
Labor force participation

Summary of results for RDD sample

- Extensive margin effect on higher education enrollment.
- Larger number of selective college degrees, that comes at cost of more dropouts from higher education.
- Earning losses in the short term, not entirely offset by year 16.
- The losses are concentrated among men, and could persist among those who drop out. Women experience earning gains in the medium and long terms.

$\rightarrow$ On average, earning gains for women and losses for men.

Why the losses? Information frictions and mismatch



...and men are around 0.36 sd more overconfident than women.

Survey answers Table

RDD: increased dropouts for most overconfident

	(1)	(2)	(3)	(4)	(5)	(6)
	Selective College			Any Institution		
	Ever enrolled	Graduation	Dropout	Ever enrolled	Graduation	Dropout
	Heterogeneity by overconfidence					
Robust - High	0.242	-0.321*	0.346***	0.005	-0.292	0.435**
	(0.190)	(0.183)	(0.118)	(0.118)	(0.201)	(0.177)
Bandwidth	64	32	42	64	45	40
Bandwidth obs.	351	188	244	352	255	230
Mean	0.511	0.502	0.056	0.930	0.797	0.075
Robust - Low	0.282	0.329*	-0.014	0.012	0.265	-0.329*
	(0.175)	(0.187)	(0.155)	(0.130)	(0.191)	(0.184)
Bandwidth	89	67	56	97	96	89
Bandwidth obs.	362	264	229	389	384	361
Mean	0.331	0.197	0.104	0.809	0.520	0.303
p-value High=Low	0.847	0.060	0.034	0.854	0.029	0.004

- Dropouts concentrated among students overoptimistic about their graduation chances from selective colleges
- More realistic students faced no negative effects when offered additional opportunities.
- Selectivity, ability distance, rank Earnings

→ Consistent with key tenet of mismatch hypothesis that information frictions are necessary to generate losses.

Section 5

Linking RCT and RDD

Getting away from the cutoff in RDD

Angrist and Rokkanen (2015) exploit additional information contained in explanatory variables other than the score to estimate treatment effects away from the cutoff.

- **Assumptions**:
  1. Conditional independence assumption (i.e., mean independence between potential outcomes and the score variable conditional on a vector of other covariates) Figures
  2. Common support Figures

In our context, the CIA is satisfied by the vector including: 1) GPA rank (9^{th} and 10^{th} grade) 2) simce 3) GPA 11^{th}

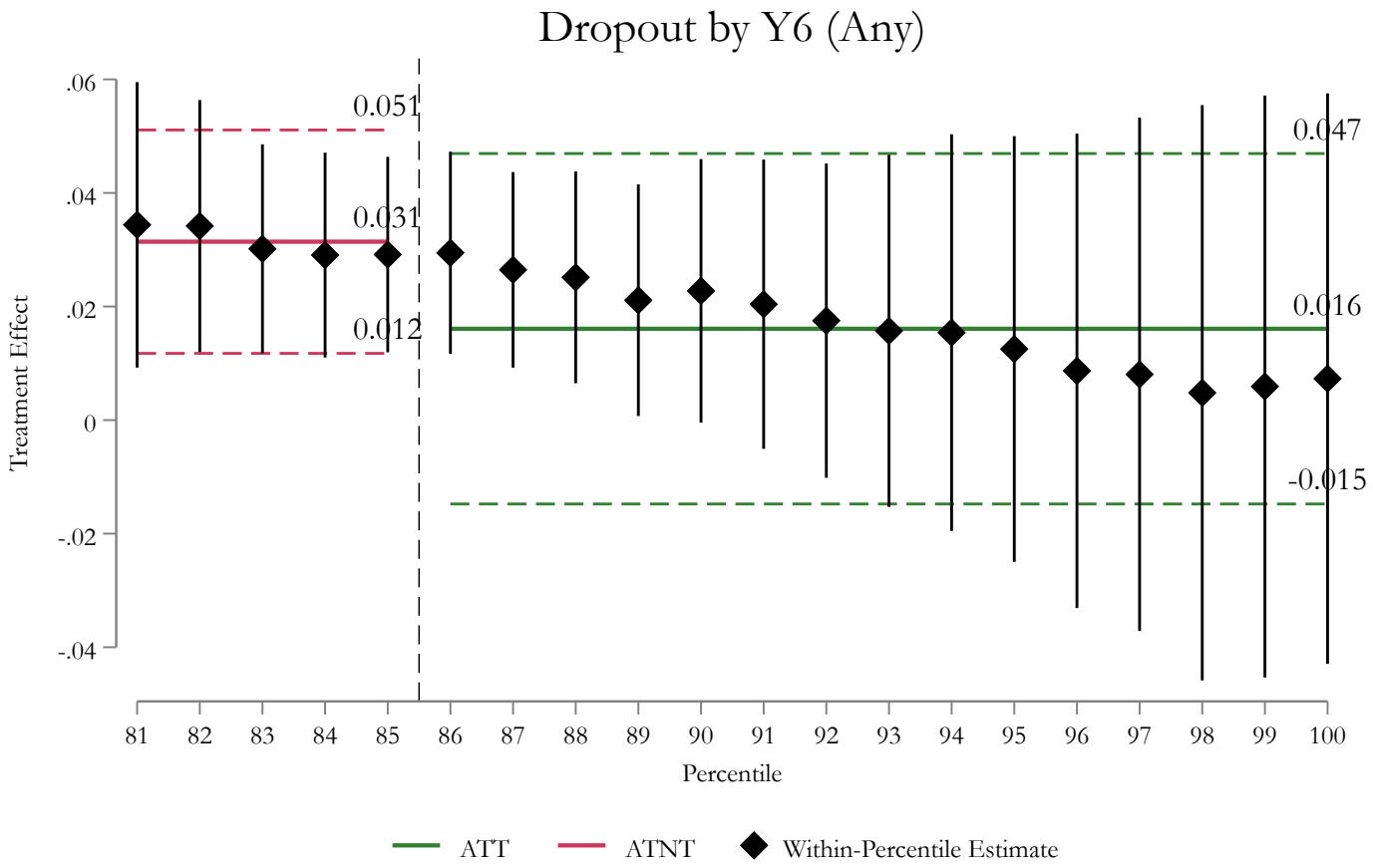
Getting away from the cutoff in RDD

Angrist and Rokkanen (2015) exploit additional information contained in explanatory variables other than the score to estimate treatment effects away from the cutoff.

- **Assumptions**:
  1. Conditional independence assumption (i.e., mean independence between potential outcomes and the score variable conditional on a vector of other covariates) Figures
  2. Common support Figures

In our context, the CIA is satisfied by the vector including: 1) GPA rank (9^{th} and 10^{th} grade) 2) simce 3) GPA 11^{th}

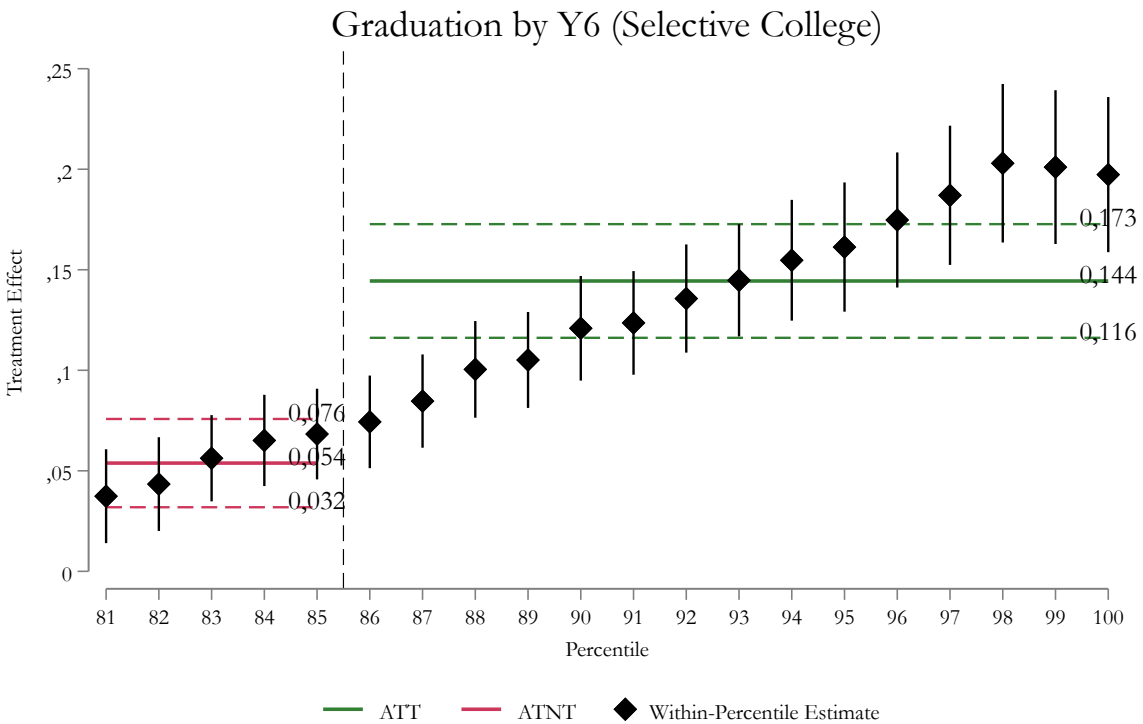
Getting away from the cutoff in RDD



How Far Can Inclusion Go? If we want to avoid the unintended consequences of dropout, a top 10% plan would achieve that goal.

Earnings

Getting away from the cutoff in RDD



However, there are positive average effects on selective college graduation even below the top 15%.

Linking RCT with RDD

Consider the potential outcomes Y_{SA} that vary depending on:

- whether you get the PACE school treatment ($S = 0, 1$, as in the RCT)
- the eligibility for preferential admission ($A = 0, 1$, as in the RDD)

Theorem (Decomposition)

$$\underbrace{\int E[Y_{1*} - Y_{00} \mid X] f(X)\, dX}_{ATE_S \rightarrow \text{ from RCT}} = \underbrace{\pi}_{\text{Fraction eligible}} \underbrace{\int E[Y_{11} - Y_{10} \mid X, A{=}1, S{=}1] f(X \mid A{=}1, S{=}1)\, dX}_{ATT_A \rightarrow \text{ from RDD getaway}}$$

$$+ \underbrace{\int E[Y_{10} - Y_{00} \mid X] f(X)\, dX}_{\text{Determined residually}}$$

Assumptions

Linking RCT with RDD: dropout from any institution

$$\underbrace{\int E[Y_{1*} - Y_{00} \mid X] f(X)\, dX}_{\text{ATE}_{S\rightarrow} \text{ from RCT}} - \underbrace{\pi}_{\text{Fraction eligible}} \underbrace{\int E[Y_{11} - Y_{10} \mid X, A{=}1, S{=}1] f(X \mid A{=}1, S{=}1)\, dX}_{\text{ATT}_{A\rightarrow} \text{ from RDD getaway}}$$

$$= \underbrace{\int E[Y_{10} - Y_{00} \mid X] f(X)\, dX}_{\text{Determined residually} = \text{-0.0006}}$$

Limited effect of the "PACE school treatment" on dropout from any institution.

Earnings

Linking RCT with RDD: graduation from selective college

$$\underbrace{\int E[Y_{1*} - Y_{00} \mid X] f(X)\, dX}_{\text{ATE}_S \rightarrow \text{ from RCT}} - \underbrace{\pi}_{\text{Fraction eligible}} \underbrace{\int E[Y_{11} - Y_{10} \mid X, A{=}1, S{=}1] f(X \mid A{=}1, S{=}1)\, dX}_{\text{ATT}_A \rightarrow \text{ from RDD getaway}}$$

$$= \underbrace{\int E[Y_{10} - Y_{00} \mid X] f(X)\, dX}_{\text{Determined residually} = 0.0030}$$

Limited effect of the "PACE school treatment" on graduation from a selective college.

Section 6

Conclusion

Summary of findings

- **Large preferential admissions can improve long-term outcomes of targeted students (RCT, average top 15% students).**
  - More selective-college degrees and higher long-term earnings for targeted disadvantaged students on average.
  - Earnings gains concentrated among women (higher take-up); men's earnings remain flat (null effects on higher education).
- **Mismatch at the margin (RDD, around top 15%).**
  - More selective college degrees for some, increased dropouts from higher education for others.
  - Earnings losses and dropouts for men and the most over-confident; women continue to experience long-term gains.
- **How far can inclusion go?**
  - Up to the **top 10%** cutoff, selective college degrees rise with no change in dropouts from higher education.
  - Beyond 10%, trade off emerges.
  - Top 10% cutoff $\sim 46^{th}$ percentile of grade 10 test scores.

Policy implications

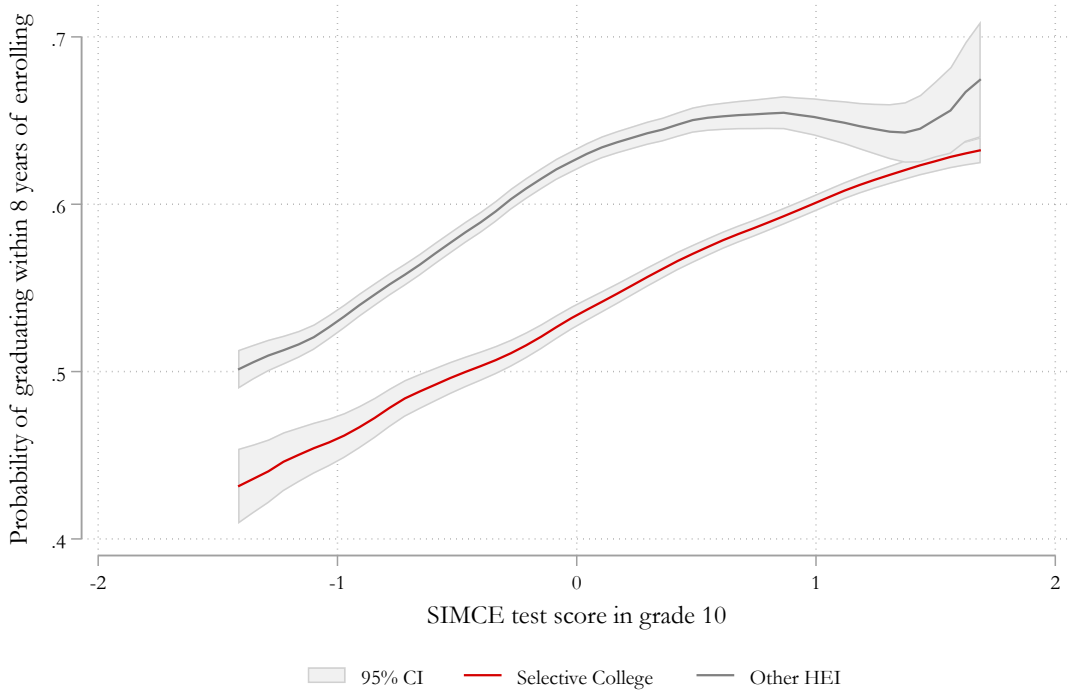
- Shift attention from *whether* to *when* mismatch in affirmative action occurs.

- Promising policy avenues:
  - Information interventions to avoid mismatch in disadvantage-based affirmative action.
  - Programs aimed at improving the academic preparation and social belonging of targeted students.

Section 7

Appendix

Graduation Rates by Type of HE Institution



Presentation

Description of target population

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Top 15%			RD			Bottom 85%		
	Obs.	Mean	S.D.	Obs.	Mean	S.D.	Obs.	Mean	S.D.
Female	2437	0.56	0.50	13048	0.59	0.49	11916	0.47	0.50
Age	2437	16.32	0.64	13048	16.33	0.65	11916	16.60	0.79
Very low SES	2437	0.59	0.49	13048	0.61	0.49	11916	0.61	0.49
Mother's education	1914	9.68	3.10	8565	9.98	3.15	7754	9.53	3.14
Father's education	1795	9.50	3.15	8090	9.79	3.24	7362	9.32	3.26
Family income	1919	284.05	195.67	8554	310.16	255.54	7782	289.21	214.89
SIMCE	2432	-0.01	0.82	12929	-0.21	0.76	11875	-0.69	0.71
Never failed	2437	0.94	0.24	13048	0.92	0.27	11916	0.81	0.39
Santiago	2437	0.17	0.37	13048	0.22	0.41	11916	0.16	0.37
Rural	2437	0.04	0.19	13048	0.05	0.22	11916	0.03	0.18
Academic track	2437	0.31	0.46	13048	0.37	0.48	11916	0.26	0.44

- Average family income is **half** that of high school students, and **a third** that of regular college entrants.

Presentation

Research Design 1: Randomized Controlled Trial

Covariate Balance Test for Top 15%

Table: Sample Balance Across Treatment and Control Groups

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Age	Very Low SES	Mother education	Father education	Family income	SIMCE score	Never failed	Santiago resident	Rural school	Academic track	Female
Treatment	0.049	-0.020	0.129	-0.019	5.756	0.084	-0.014	0.041	-0.013	0.075	0.001
	(0.051)	(0.025)	(0.177)	(0.229)	(12.545)	(0.121)	(0.018)	(0.066)	(0.019)	(0.072)	(0.055)
p-value	0.340	0.418	0.468	0.935	0.647	0.487	0.444	0.533	0.497	0.297	0.979
Mean	16.303	0.596	9.642	9.508	282.134	-0.041	0.941	0.155	0.043	0.281	0.561
S.d.	0.587	0.491	3.132	3.103	198.181	0.805	0.237	0.362	0.203	0.450	0.496
N	2437	2437	1914	1795	1919	2432	2437	2437	2437	2437	2437

Notes: In this table we regress pre-determined variables on the treatment status of baseline top 15% students. Treatment is the coefficient of each regression. Standard errors clustered at the school level are shown in parentheses. The p-value is the p-value of the test of significance of the treatment coefficient. Mean is the average of the pre-determined variable in the control group. Low-SES student is a student that the Government classified as very socioeconomically vulnerable (*Prioritario*). SIMCE is a standardized achievement test taken in 10^{th} grade. Age and education are in years. Family income is the monthly family income in 1000 Chilean pesos. *** p<0.01, ** p<0.05, * p<0.10.

Presentation

Research Design 2: RDD

Covariate Balance Test

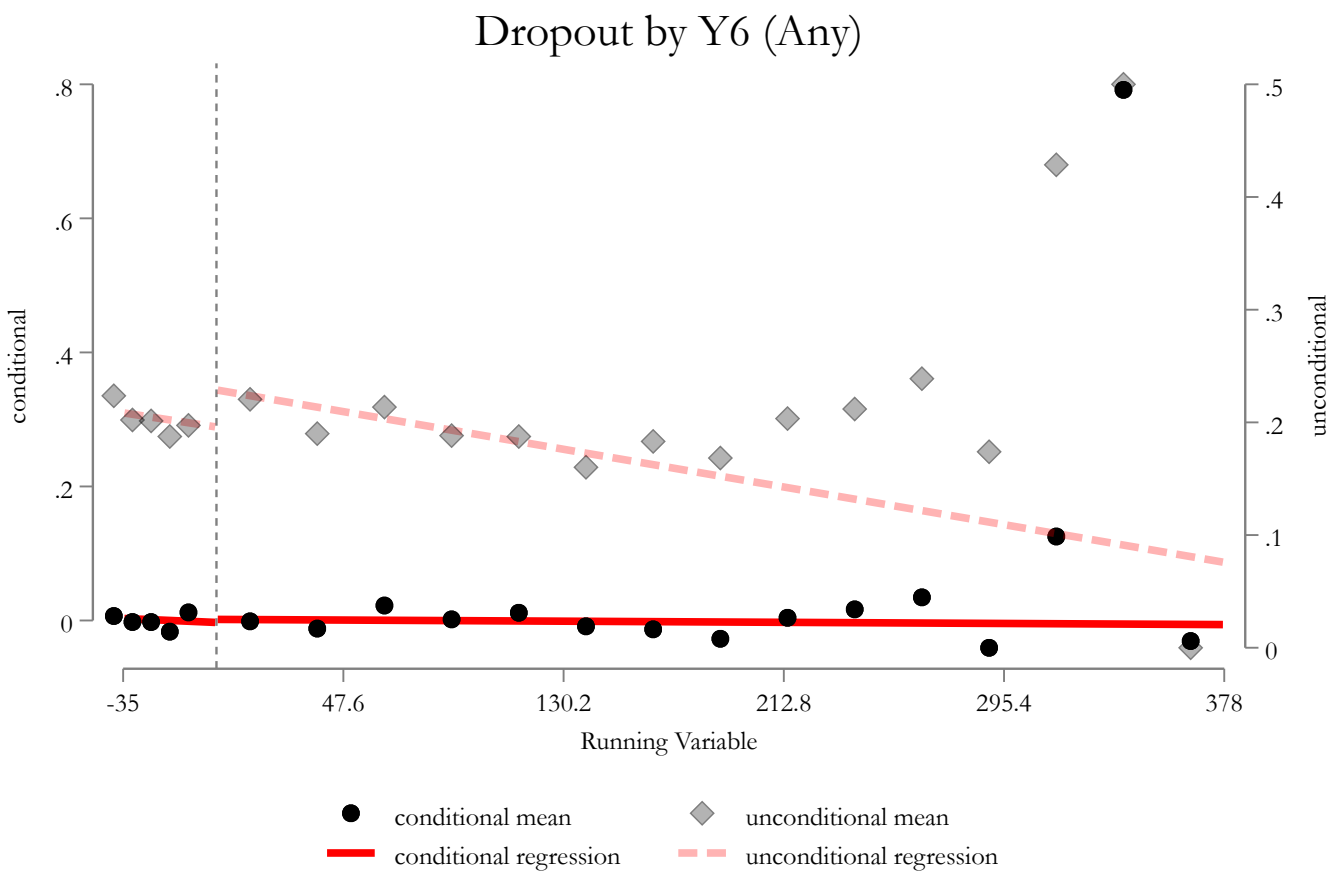
Table: Tests for discontinuities in pre-determined variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Age	Very Low SES	Mother educ.	Father educ.	Family income	SIMCE score	Never failed	Santiago resident	Rural school	Academic track	Over-conf.	Female
Conventional	0.069***	0.007	0.155	0.064	12.000	-0.013	-0.009	0.025	-0.005	-0.028	-0.020	0.005
	(0.026)	(0.022)	(0.158)	(0.169)	(14.542)	(0.055)	(0.010)	(0.037)	(0.016)	(0.042)	(0.119)	(0.026)
Robust	0.078***	0.006	0.206	0.021	14.897	-0.023	-0.009	0.036	-0.003	-0.044	-0.059	0.009
	(0.028)	(0.024)	(0.173)	(0.187)	(16.514)	(0.058)	(0.011)	(0.039)	(0.017)	(0.043)	(0.133)	(0.029)
Bandwidth	61.498	72.583	69.298	72.238	81.019	74.528	92.260	67.623	83.743	65.930	60.623	66.636
Bandwidth N	12236	14430	9015	8853	10500	14680	18217	13448	16555	13048	580	13264
R-squared	0.005	0.000	0.001	0.001	0.000	0.041	0.008	0.001	0.000	0.005	0.003	0.000
Mean	16.263	0.610	9.990	9.732	305.467	-0.176	0.929	0.209	0.051	0.391	0.110	0.595

Notes: In this table we report the estimate for coefficient ϕ in regression equation (3), using pre-determined variables as the dependent variable. Standard errors clustered at the school level are shown in parentheses. *Mean* is the mean of the outcome variable just below the cutoff. *Robust* uses the robust approach with bias-correction suggested by Calonico, Cattaneo, and Titiunik (2014). Optimal bandwidths, a linear polynomial of the ranking score and uniform kernels are used in all the specifications. Low-SES student is a student that the Government classified as very socioeconomically vulnerable (*Prioritario*). SIMCE is a standardized achievement test taken in 10^{th} grade. Age and education are in years. Family income is the monthly family income in 1000 Chilean pesos. *** p<0.01, ** p<0.05, * p<0.10.

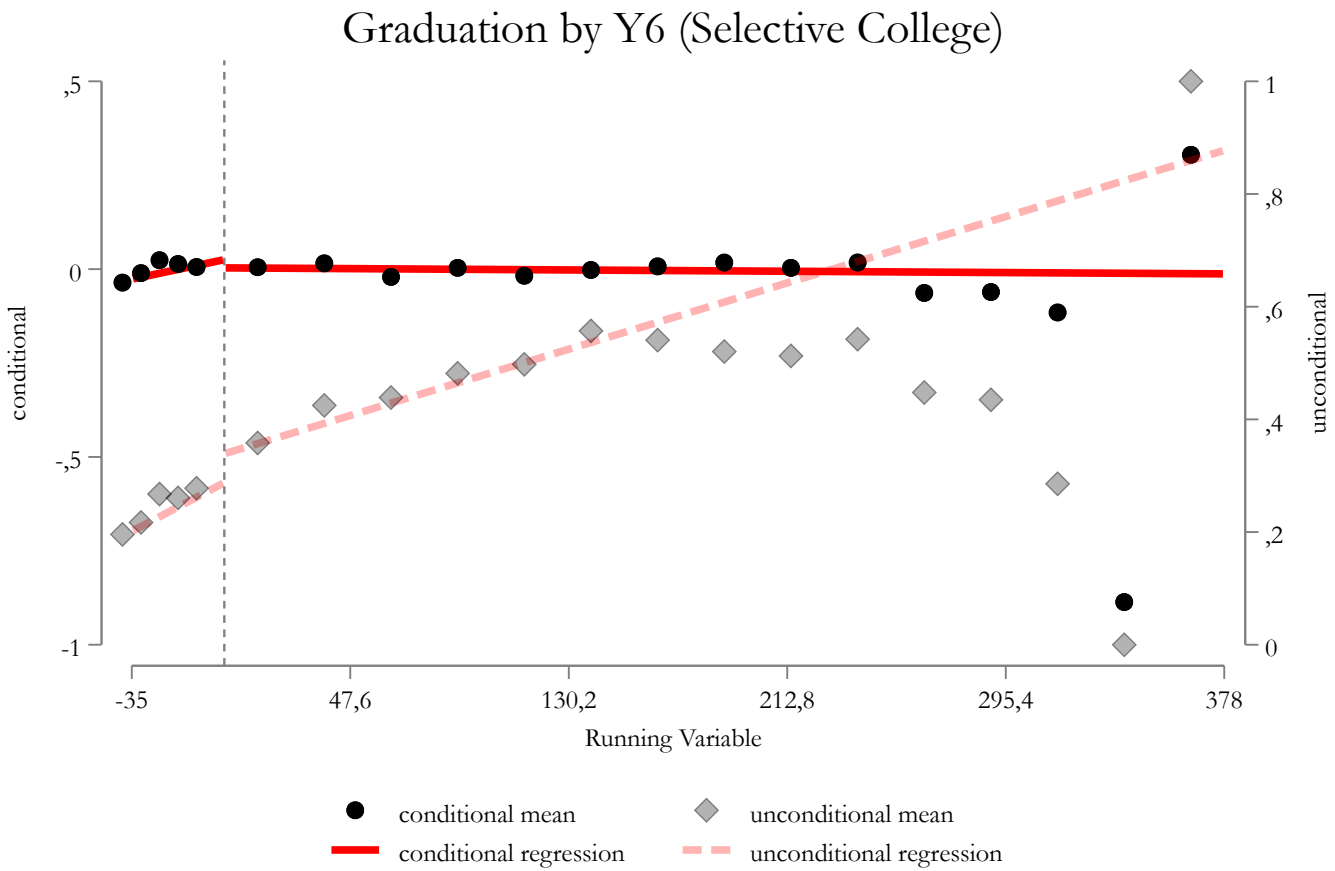
Presentation

CIA



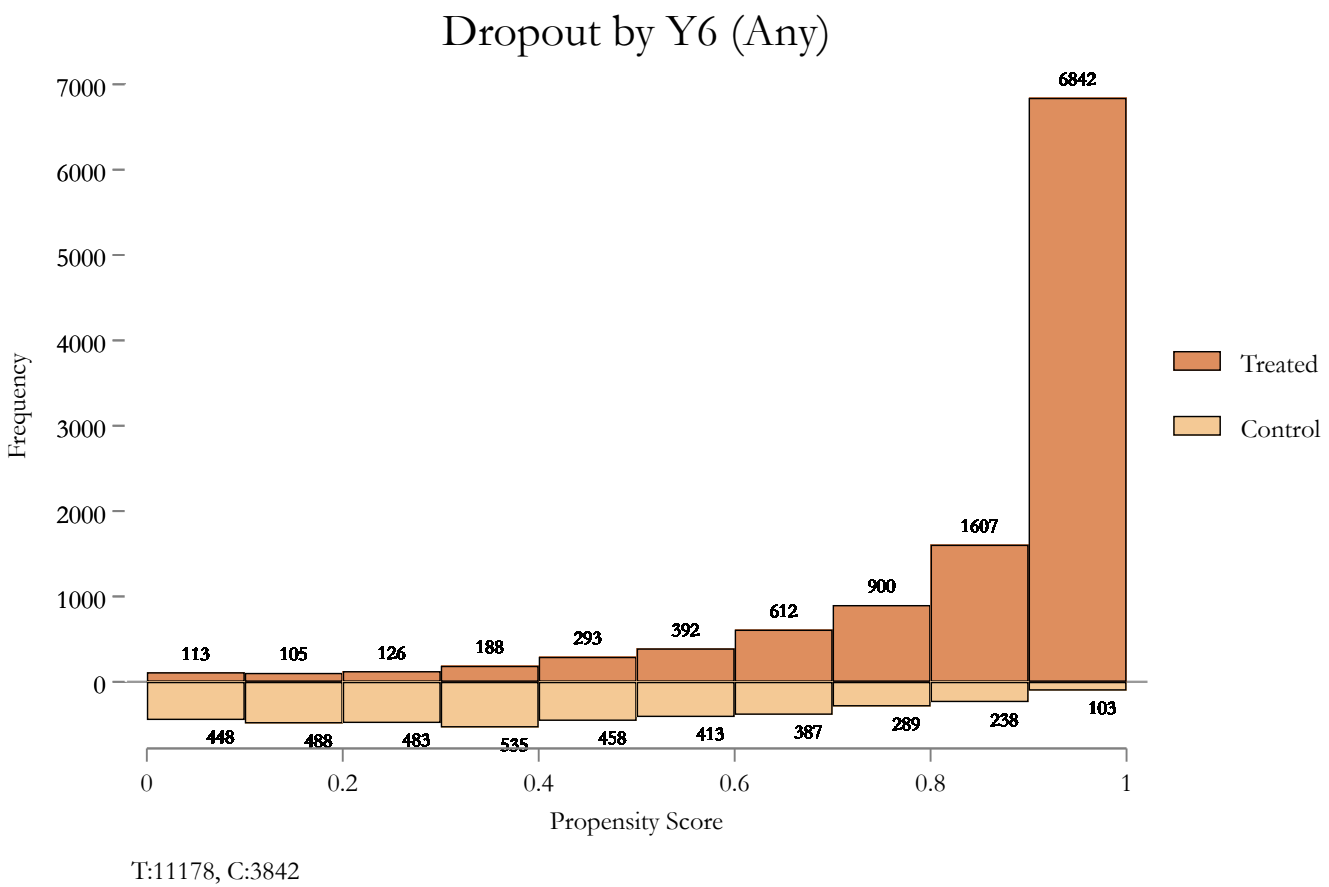
Presentation

CIA



Presentation

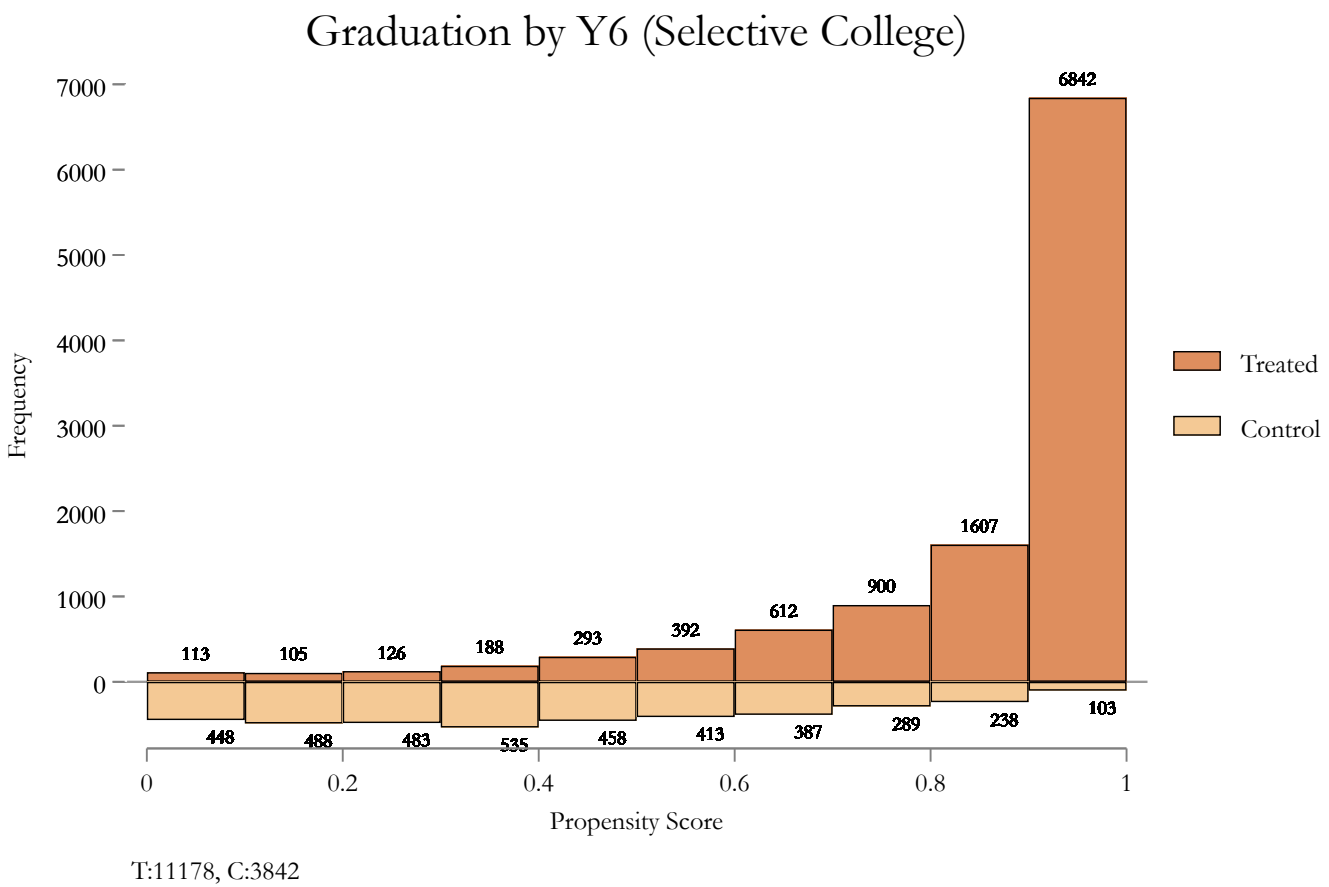
Common Support



T:11178, C:3842

Presentation

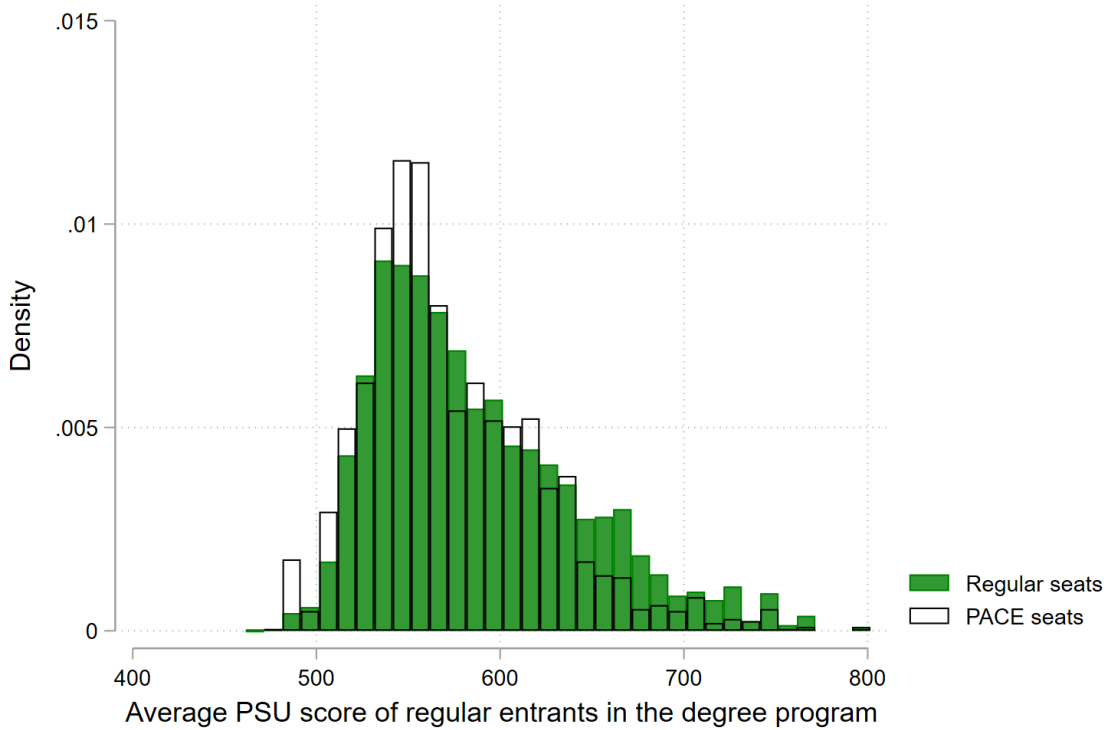
Common Support



T:11178, C:3842

Presentation

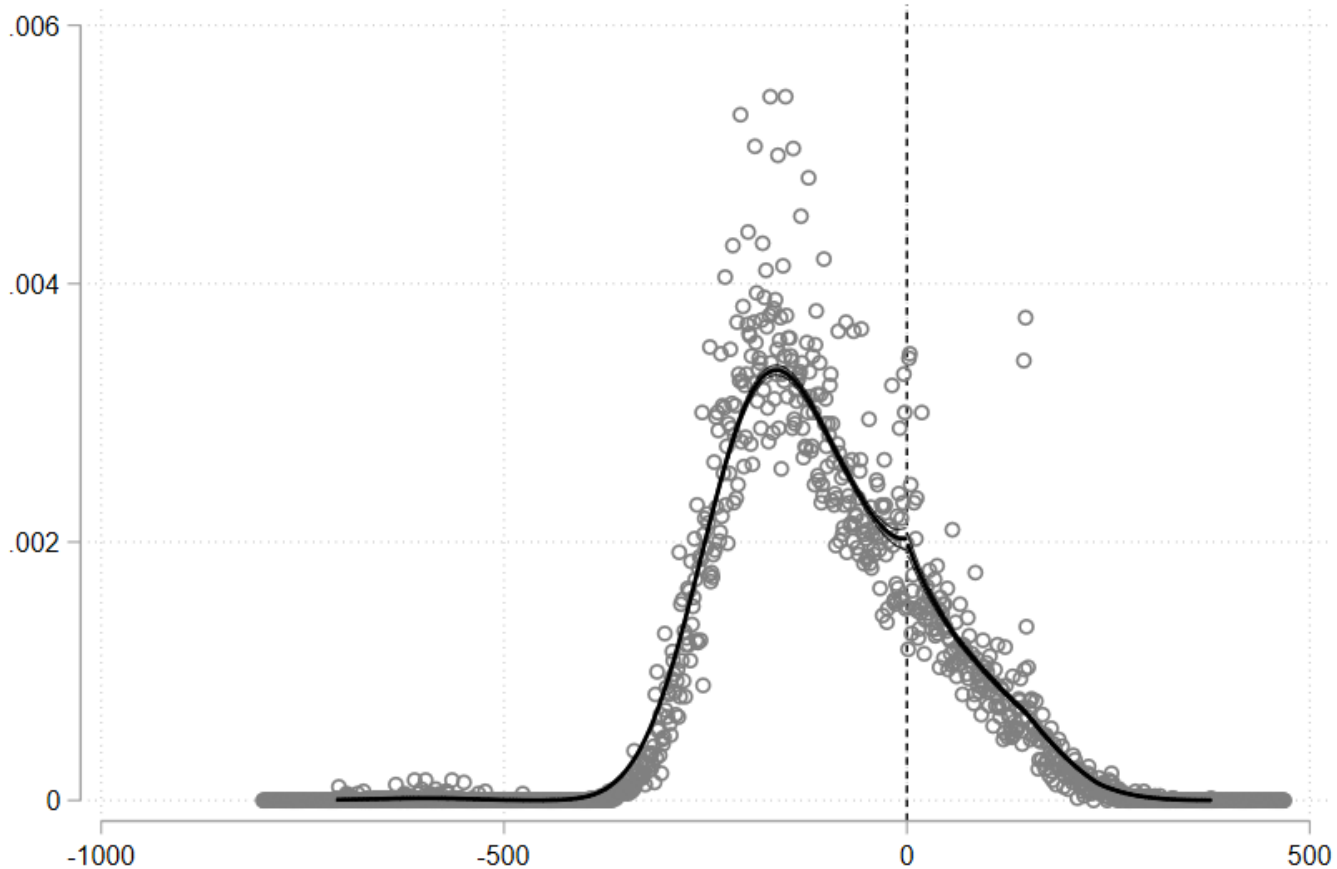
Selectivity of regular and PACE college seats



Notes: Nationwide distribution of college seat selectivity (average entrance exam score of regular entrants, 2018).

Presentation

McCrary test



Presentation

Surrogate index

Estimate following regression on older student cohorts:

$$Y_{ist} = \beta_{0t} + \beta_{1t} Z_i + \beta_{2t} X_{is} + u_{it}$$

and predict surrogate index $\hat{Y}_{ist}$ for study sample.

Key assumptions:

1. *Comparability:* conditional distribution of long-term outcomes given surrogate index is the same in the prediction and main samples.
2. *Surrogacy:* long-term outcomes independent of treatment conditional on the surrogate index.

Surrogate index

Using five older cohorts, we use these predictors Z_i for each year from 1 to 6:

- Enrollment in any HEI and in a selective college
- Graduation from any HEI and from a selective college
- Interactions of previous predictors with major area
- Enrollment followed by dropout from any institution and from selective college
- Extensive margin LFP (i.e., working in the private sector, attending higher education, neither or both) and yearly months employed
- LFP*months, LFP*sector, months*sector, LFP*months*sector
- Yearly earnings
- Earnings*sector, LFP*earnings, LFP*earnings*sector
- Z_i interacted with gender

Additional controls X_{is}:

- Individual: Gender, age, indicator for never failed a grade, high school track (academic or vocational), high school GPA
- School: rural, Santiago, tracks offered, cohort size
- Youth unemployment rate (gender-specific) the year after graduating from high school
- X_{is} interacted with gender

Presentation

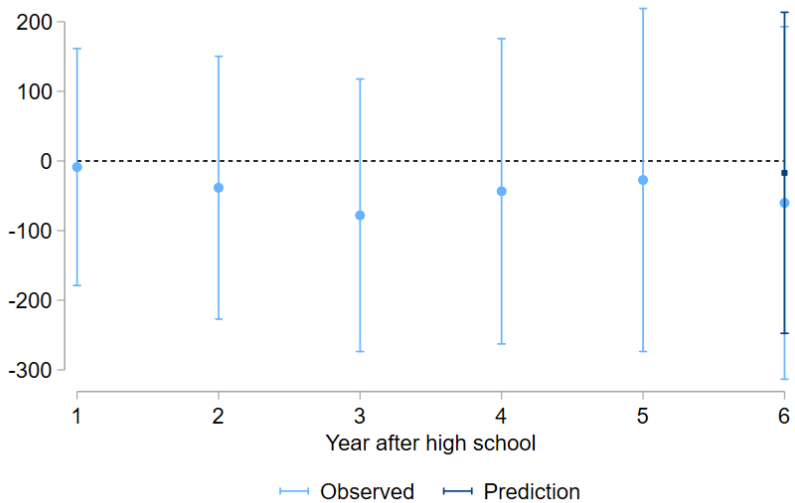
RCT: Impact on months worked for top 15%

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	In a given year						**Average across years**		
	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	Years 1-6	Years 7-11	Years 12-16
Treatment	-0.157	-0.083	-0.172	0.157	-0.329	-0.493*	-0.127	-0.149	0.052
	(0.182)	(0.221)	(0.200)	(0.245)	(0.293)	(0.260)	(0.191)	(0.118)	(0.069)
Mean	2.052	2.716	3.000	4.233	5.628	6.452	3.799	7.468	5.539
Total obs.	2,007	2,172	2,140	2,095	2,041	1,934	2,384	2,437	2,437

Presentation

RCT top 15%: validation of earning effect predictions

Figure: Earnings, 6-year average



Presentation

RCT top 15%: Impact on selectivity of HE program

Table: Lee bounds on selectivity and rank effects in higher-education course (RCT, baseline top 15%)

	(1)	(2)	(3)	(4)	(5)	(6)
	Selectivity		Ability distance		Rank	
	Lower bound	Upper bound	Lower bound	Upper bound	Lower bound	Upper bound
Treatment	0.177***	0.235***	0.070*	0.156***	-0.053***	-0.032**
	(0.036)	(0.034)	(0.041)	(0.041)	(0.017)	(0.015)
Total obs.	2437	2437	2437	2437	2437	2437
Selected obs.	2123	2123	2123	2123	2120	2120
Mean	0.029	0.029	0.011	0.011	0.458	0.458

Presentation

RCT top 15%: Impacts on STEM/Non-STEM

	(1)	(2)	(3)	(4)	(5)	(6)
	Selective College			Any Institution		
	Ever enrolled	Graduation	Dropout	Ever enrolled	Graduation	Dropout
STEM						
Treatment	0.073***	0.037*	0.014	0.045*	0.010	0.012
	(0.025)	(0.021)	(0.009)	(0.025)	(0.023)	(0.013)
Mean	0.232	0.141	0.039	0.461	0.312	0.100
Total obs.	2437	2437	2437	2437	2437	2437
Non-STEM						
Treatment	0.047**	0.024	0.019*	-0.023	-0.008	-0.010
	(0.024)	(0.020)	(0.011)	(0.028)	(0.024)	(0.014)
Mean	0.218	0.137	0.039	0.467	0.297	0.116
Total obs.	2437	2437	2437	2437	2437	2437

Presentation

RCT: Impact on months worked for top 15% by gender

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	In a given year						**Average across years**		
	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	Years 1-6	Years 7-11	Years 12-16
Treatment – Women	-0.188	-0.119	-0.131	0.001	-0.678*	-0.505	-0.188	-0.114	0.080
	(0.203)	(0.263)	(0.241)	(0.273)	(0.346)	(0.356)	(0.217)	(0.148)	(0.072)
Mean Women	1.723	2.272	2.332	3.573	5.056	5.794	3.239	7.007	4.660
Total obs.	1,119	1,222	1,205	1,167	1,124	1,039	1,335	1,369	1,369
Treatment – Men	-0.139	-0.080	-0.256	0.334	0.074	-0.474	-0.067	-0.211	0.006
	(0.301)	(0.328)	(0.306)	(0.355)	(0.366)	(0.346)	(0.264)	(0.184)	(0.114)
Mean Men	2.455	3.284	3.838	5.037	6.306	7.196	4.504	8.057	6.664
Total obs.	888	950	935	928	917	895	1,049	1,068	1,068
p-value Women=Men	0.891	0.921	0.733	0.403	0.077	0.948	0.680	0.675	0.554

Presentation

RCT top 15%: Impact on selectivity of HE program, females

Table: Lee bounds on selectivity and rank effects in higher-education course (RCT, Top 15 % female students)

	(1)	(2)	(3)	(4)	(5)	(6)
	Selectivity		Ability distance		Rank	
	Lower bound	Upper bound	Lower bound	Upper bound	Lower bound	Upper bound
Treatment - Women	0.208***	0.320***	0.089*	0.233***	-0.085***	-0.038*
	(0.043)	(0.043)	(0.049)	(0.050)	(0.021)	(0.020)
Total obs.	1369	1369	1369	1369	1369	1369
Selected obs.	1228	1228	1228	1228	1227	1227
Mean	0.021	0.021	-0.001	-0.001	0.468	0.468

Notes: In this table we report the Lee bounds for the estimate of the coefficient β in regression equation (1). Standard errors clustered at the school level are shown in parentheses. *Total obs.* are the number of observations before the trimming procedure. *Selected obs.* are the number of observations after the trimming procedure. *Selectivity* represents the average baseline ability of peers in the first degree program a student enrolls in. Ability distance is the difference between selectivity and own baseline ability (in these regressions we do not control for own baseline ability). Rank denotes a student's relative ability among these peers: 0 if the student is the lowest-ability one and 1 if the student is the highest-ability one. *Mean* is the mean of the outcome variable just below the cutoff in the untrimmed sample. *** $p<0.01$, ** $p<0.05$, * $p<0.10$.

Presentation

RCT top 15%: Impact on selectivity of HE program, males

Table: Lee bounds on selectivity and rank effects in higher-education course (RCT, Top 15 % male students)

	(1)	(2)	(3)	(4)	(5)	(6)
	Selectivity		Ability distance		Rank	
	Lower bound	Upper bound	Lower bound	Upper bound	Lower bound	Upper bound
Treatment - Men	0.117**	0.144**	0.025	0.069	-0.026	-0.008
	(0.053)	(0.064)	(0.075)	(0.071)	(0.025)	(0.027)
Total obs.	1068	1068	1068	1068	1068	1068
Selected obs.	895	895	895	895	893	893
Mean	0.041	0.041	0.028	0.028	0.446	0.446

Notes: In this table we report the Lee bounds for the estimate of the coefficient β in regression equation (1). Standard errors clustered at the school level are shown in parentheses. *Total obs.* are the number of observations before the trimming procedure. *Selected obs.* are the number of observations after the trimming procedure. *Selectivity* represents the average baseline ability of peers in the first degree program a student enrolls in. Ability distance is the difference between selectivity and own baseline ability (in these regressions we do not control for own baseline ability). Rank denotes a student's relative ability among these peers: 0 if the student is the lowest-ability one and 1 if the student is the highest-ability one. *Mean* is the mean of the outcome variable just below the cutoff in the untrimmed sample. *** p<0.01, ** p<0.05, * p<0.10.

Presentation

RCT top 15%: Impacts on STEM by gender

	(1)	(2)	(3)	(4)	(5)	(6)
	Selective College			Any Institution		
	Ever enrolled	Graduation	Dropout	Ever enrolled	Graduation	Dropout
	STEM					
Treatment – Women	0.083***	0.052**	0.014	0.067*	0.030	0.017
	(0.026)	(0.026)	(0.009)	(0.035)	(0.031)	(0.015)
Mean Women	0.211	0.136	0.026	0.390	0.287	0.056
Total obs.	1369	1369	1369	1369	1369	1369
Treatment – Men	0.064	0.018	0.017	0.017	-0.017	0.008
	(0.040)	(0.031)	(0.015)	(0.036)	(0.032)	(0.024)
Mean Men	0.260	0.147	0.054	0.552	0.345	0.156
Total obs.	1068	1068	1068	1068	1068	1068
p-value Treat Women=Men	0.645	0.384	0.861	0.323	0.280	0.752

Presentation

RCT top 15%: Impacts on Non-STEM by gender

	(1)	(2)	(3)	(4)	(5)	(6)
	Selective College			Any Institution		
	Ever enrolled	Graduation	Dropout	Ever enrolled	Graduation	Dropout
	Non-STEM					
Treatment – Women	0.078**	0.054*	0.020	-0.024	-0.003	-0.016
	(0.032)	(0.029)	(0.015)	(0.037)	(0.034)	(0.020)
Mean Women	0.253	0.160	0.038	0.558	0.368	0.122
Total obs.	1369	1369	1369	1369	1369	1369
Treatment – Men	0.008	-0.011	0.017	-0.015	-0.011	-0.003
	(0.030)	(0.022)	(0.015)	(0.038)	(0.028)	(0.019)
Mean Men	0.173	0.108	0.041	0.351	0.205	0.109
Total obs.	1068	1068	1068	1068	1068	1068
p-value Treat Women=Men	0.097	0.047	0.896	0.850	0.839	0.621

Presentation

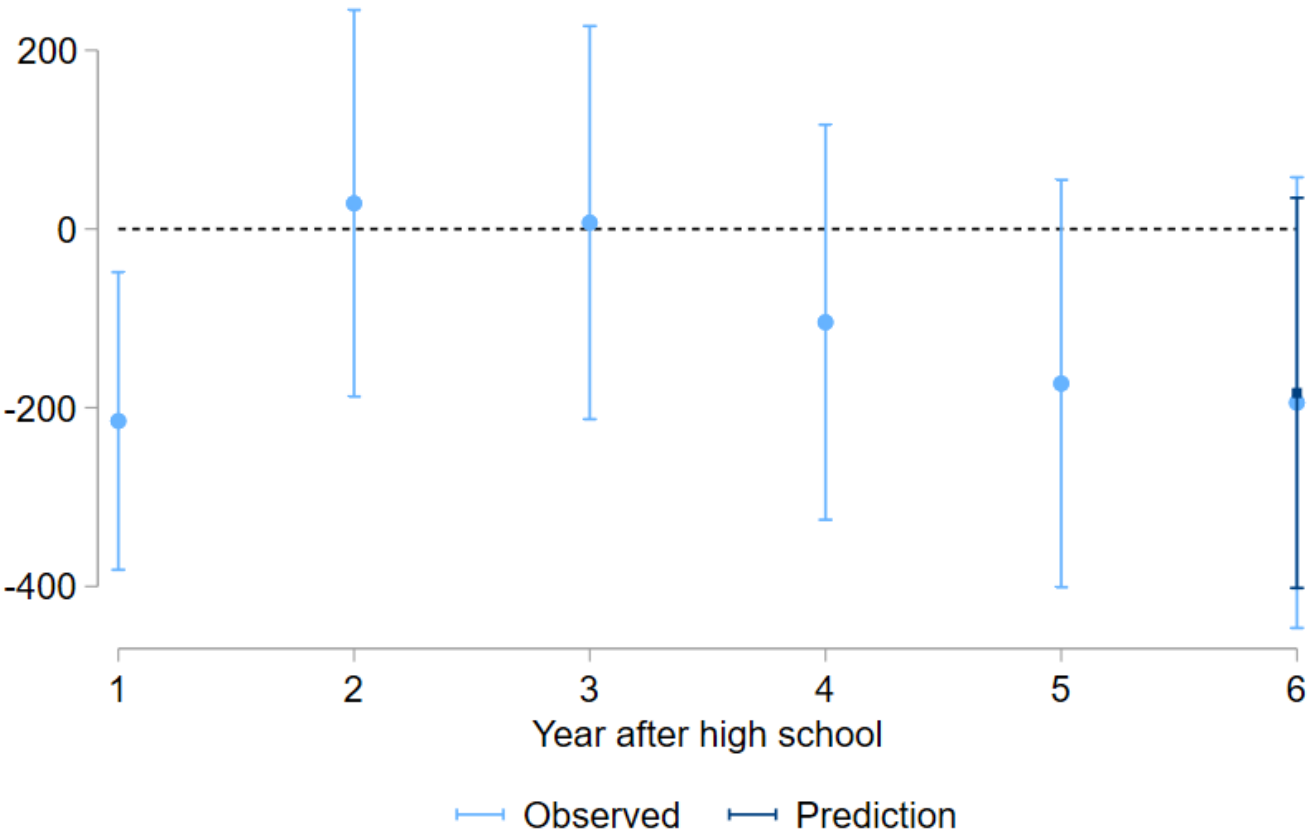
RDD: Impact on months worked

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	In a given year						**Average across years**		
	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	Years 1-6	Years 7-11	Years 12-16
Robust	-0.706***	0.074	-0.183	-0.571**	-0.455*	0.091	-0.317*	0.015	-0.029
	(0.161)	(0.255)	(0.232)	(0.239)	(0.250)	(0.340)	(0.184)	(0.163)	(0.068)
Bandwidth	81.835	44.874	67.821	72.052	79.248	49.223	62.062	43.766	69.614
Bandwidth obs.	13,353	8,118	11,697	12,126	12,994	7,758	12,125	8,853	13,838
Mean	2.203	2.896	3.150	4.125	5.063	6.307	3.702	7.267	5.172

Presentation

RDD: validation of earning effect predictions

Figure: Earnings, 6-year average



RDD: Impact on selectivity of HE program

Table: Bounds on selectivity and rank effects in higher education course (RD, all students)

	(1)	(2)	(3)	(4)	(5)	(6)
	Selectivity		Ability distance		Rank	
	Lower bound	Upper bound	Lower bound	Upper bound	Lower bound	Upper bound
Robust	0.129***	0.263***	0.034	0.346***	-0.125***	-0.026
	(0.037)	(0.035)	(0.050)	(0.047)	(0.019)	(0.022)
Bandwidth	58	58	59	59	59	59
Observations	9812	9813	9755	9756	9755	9756
Mean	-0.059	-0.059	0.055	0.055	0.444	0.444

Notes: In this table we report the bounds on the estimate for coefficient δ in regression equation (2). Panel A shows the estimates for all students who belong to the RD sample of students. Panel B displays the estimates for the same sample divided by gender. Standard errors clustered at the school level are shown in parentheses. *Robust* uses the robust approach with bias-correction suggested by **?**. Optimal bandwidths, a linear polynomial of the ranking score and uniform kernels are used in all the specifications. Selectivity is average baseline ability of student peers in the same degree program. Ability distance is the difference between selectivity and own baseline ability. Rank is the position in the baseline-ability ranking from 0 to 1 in the same degree program. *Mean* is the mean of the outcome variable just below the cutoff in the untrimmed sample.

Presentation

RDD: Impacts on STEM/Non-STEM

	(1)	(2)	(3)	(4)	(5)	(6)
	Selective College			Any Institution		
	Ever enrolled	Graduation	Dropout	Ever enrolled	Graduation	Dropout
STEM						
Robust	0.034	-0.002	0.022*	-0.019	-0.014	0.004
	(0.029)	(0.018)	(0.013)	(0.034)	(0.026)	(0.017)
Bandwidth	33	55	66	41	56	67
Bandwidth obs.	6869	11003	13048	8383	11099	13349
Mean	0.213	0.123	0.040	0.433	0.283	0.106
Non-STEM						
Robust	0.175***	0.056**	0.079***	0.078***	-0.025	0.079***
	(0.027)	(0.022)	(0.015)	(0.025)	(0.030)	(0.023)
Bandwidth	58	58	43	72	54	39
Bandwidth obs.	11518	11640	8783	14328	10731	8009
Mean	0.217	0.137	0.033	0.483	0.343	0.097

Presentation

RDD: Impact on months worked by gender

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	In a given year						**Average across years**		
	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	Years 1-6	Years 7-11	Years 12-16
Robust - Women	-0.627***	-0.136	-0.103	-0.691**	-0.645*	0.109	-0.380	-0.007	-0.013
	(0.204)	(0.297)	(0.321)	(0.323)	(0.352)	(0.466)	(0.231)	(0.206)	(0.077)
Bandwidth	71.373	56.587	54.100	53.915	61.432	46.574	53.088	42.545	70.788
Bandwidth obs.	6,726	5,938	5,566	5,303	5,872	4,255	6,186	5,172	8,294
Mean	2.104	2.680	2.837	3.755	4.804	5.782	3.485	6.948	4.365
Robust - Men	-0.623**	-0.025	-0.236	-0.770*	0.140	0.294	-0.114	0.231	0.160
	(0.292)	(0.373)	(0.436)	(0.460)	(0.424)	(0.434)	(0.296)	(0.219)	(0.225)
Bandwidth	60.329	66.482	54.469	49.067	62.426	61.712	66.428	50.552	46.410
Bandwidth obs.	4,220	4,850	3,894	3,490	4,350	4,122	5,227	4,152	3,789
Mean	2.413	3.137	3.675	4.651	5.448	6.559	4.119	7.681	6.372
p-value Women=Men	0.974	0.911	0.804	0.934	0.244	0.864	0.611	0.527	0.682

Presentation

RDD: Impact on selectivity of HE program, females

Table: Bounds on selectivity and rank effects in higher education course (RD, female students)

	(1)	(2)	(3)	(4)	(5)	(6)
	Selectivity		Ability distance		Rank	
	Lower bound	Upper bound	Lower bound	Upper bound	Lower bound	Upper bound
Robust - Women	0.111**	0.270***	0.008	0.330***	-0.119***	-0.008
	(0.050)	(0.047)	(0.067)	(0.064)	(0.026)	(0.029)
Bandwidth	50	50	49	49	49	49
Observations	5133	5134	4938	4939	4938	4939
Mean	-0.096	-0.096	0.071	0.071	0.443	0.443

Notes: In this table we report the bounds on the estimate for coefficient δ in regression equation (2). Panel A shows the estimates for all students who belong to the RD sample of students. Panel B displays the estimates for the same sample divided by gender. Standard errors clustered at the school level are shown in parentheses. *Robust* uses the robust approach with bias-correction suggested by Calonico et al., 2014. Optimal bandwidths, a linear polynomial of the ranking score and uniform kernels are used in all the specifications. Selectivity is average baseline ability of student peers in the same degree program. Ability distance is the difference between selectivity and own baseline ability. Rank is the position in the baseline-ability ranking from 0 to 1 in the same degree program. *Mean* is the mean of the outcome variable just below the cutoff in the untrimmed sample.

Presentation

RDD: Impact on selectivity of HE program, males

Table: Bounds on selectivity and rank effects in higher education course (RD, male students)

	(1)	(2)	(3)	(4)	(5)	(6)
	Selectivity		Ability distance		Rank	
	Lower bound	Upper bound	Lower bound	Upper bound	Lower bound	Upper bound
Robust - Men	0.138**	0.183***	0.064	0.274***	-0.117***	-0.049
	(0.055)	(0.054)	(0.070)	(0.063)	(0.026)	(0.030)
Bandwidth	71	71	78	78	78	78
Observations	4824	4825	5123	5124	5123	5124
Mean	0.008	0.008	0.027	0.027	0.446	0.446

Notes: In this table we report the bounds on the estimate for coefficient δ in regression equation (2). Panel A shows the estimates for all students who belong to the RD sample of students. Panel B displays the estimates for the same sample divided by gender. Standard errors clustered at the school level are shown in parentheses. *Robust* uses the robust approach with bias-correction suggested by Calonico et al., 2014. Optimal bandwidths, a linear polynomial of the ranking score and uniform kernels are used in all the specifications. Selectivity is average baseline ability of student peers in the same degree program. Ability distance is the difference between selectivity and own baseline ability. Rank is the position in the baseline-ability ranking from 0 to 1 in the same degree program. *Mean* is the mean of the outcome variable just below the cutoff in the untrimmed sample.

Presentation

RDD: Impacts on STEM by gender

	(1)	(2)	(3)	(4)	(5)	(6)
	Selective College			Any Institution		
	Ever enrolled	Graduation	Dropout	Ever enrolled	Graduation	Dropout
	STEM					
Robust - Women	0.010	-0.008	0.002	-0.001	0.010	-0.015
	(0.036)	(0.028)	(0.013)	(0.042)	(0.038)	(0.019)
Bandwidth	36	41	62	43	43	56
Bandwidth obs.	4452	4959	7362	5172	5233	6579
Mean	0.197	0.117	0.034	0.354	0.246	0.077
Robust - Men	0.013	-0.013	0.069***	-0.086	-0.087**	0.032
	(0.044)	(0.034)	(0.023)	(0.054)	(0.038)	(0.030)
Bandwidth	39	44	60	44	61	75
Bandwidth obs.	3177	3628	4856	3628	4972	6051
Mean	0.234	0.132	0.052	0.534	0.322	0.144
p-value Women=Men	0.190	0.991	0.021	0.195	0.082	0.146

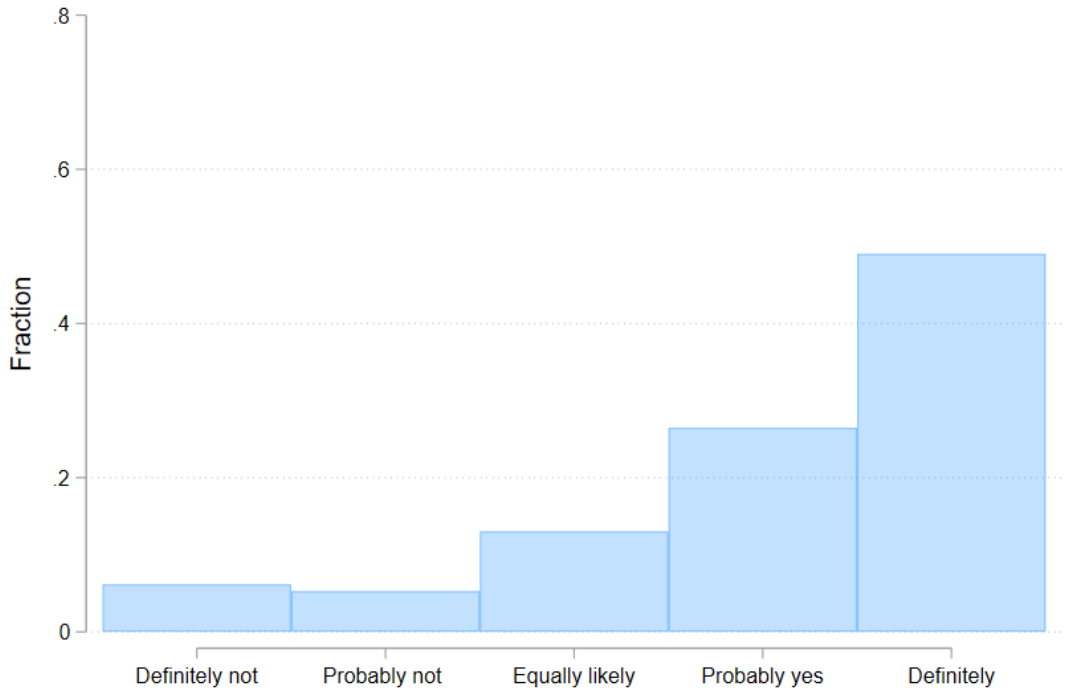
Presentation

RDD: Impacts on Non-STEM by gender

	(1)	(2)	(3)	(4)	(5)	(6)
	Selective College			Any Institution		
	Ever enrolled	Graduation	Dropout	Ever enrolled	Graduation	Dropout
	Non-STEM					
Robust - Women	0.198***	0.071**	0.089***	0.078**	-0.019	0.081***
	(0.032)	(0.028)	(0.019)	(0.039)	(0.035)	(0.030)
Bandwidth	62	52	55	49	57	40
Bandwidth obs.	7362	6155	6517	5815	6777	4845
Mean	0.230	0.144	0.029	0.540	0.389	0.105
Robust - Men	0.167***	0.021	0.074***	0.070*	-0.005	0.076**
	(0.037)	(0.037)	(0.023)	(0.041)	(0.044)	(0.033)
Bandwidth	64	43	46	67	55	40
Bandwidth obs.	5199	3509	3789	5409	4451	3258
Mean	0.196	0.124	0.043	0.382	0.253	0.092
p-value Women=Men	0.653	0.435	0.388	0.795	0.475	0.902

Presentation

Perceived graduation probability



Presentation

Raw survey answers, perceived graduation chances

Survey Answer	Actual % Graduates (1)
Chances of Graduating <= 50%	56.49
Will Probably Graduate	64.40
Will Certainly Graduate	69.47
Any Survey Answer	66.93

Notes: The table uses the sample of students who enrolled in a selective college during the six years after high school and who were surveyed on their beliefs regarding selective college graduation conditional on enrolling. Beliefs were collected through a survey in the last high school year (2017). Information on actual college performance comes from linked administrative records for the same students six years after leaving high school (2023). Each row restricts the sample according to students' survey answers, and shows among the students who gave each answer what percentage have graduated or are on track to graduate from a selective college six years later.

Presentation

Gender gap in overconfidence

	(1)	(2)
	Overconfidence	Overconfidence
Female	-0.306***	-0.296***
	(0.031)	(0.032)
Controls	NO	YES
Obs.	5770	5770
Mean	0.141	0.141

Notes: In this table we regress overconfidence on the gender dummy, in the sample of survey respondents. Overconfidence is the difference between the perceived and the actual likelihood of graduating from a selective college (see Appendix **??**), standardized to have mean zero and variance one. Standard errors clustered at the school level are shown in parentheses. Mean refers to average overconfidence among male students. The regression in column (2) includes the following controls: age, indicator for very-low-SES student, baseline SIMCE test score, never failed a grade, and high school track (academic or vocational). *** p<0.01, ** p<0.05, * p<0.10.

Presentation

RDD: Impact of selectivity of HE program, most overconfident

Table: Bounds on selectivity and rank effects in higher education course (RD) by overconfidence

	(1)	(2)	(3)	(4)	(5)	(6)
	Selectivity		Ability distance		Rank	
	Lower bound	Upper bound	Lower bound	Upper bound	Lower bound	Upper bound
Robust - High	0.090	0.203	0.980**	1.277***	-0.242	-0.193
	(0.272)	(0.267)	(0.456)	(0.435)	(0.167)	(0.161)
Bandwidth	65	65	65	65	65	65
Observations	309	308	309	308	309	308
Mean	-0.028	-0.028	0.073	0.073	0.426	0.426

Notes: In this table we report the bounds on the estimate for coefficient δ in regression equation (2). Standard errors clustered at the school level are shown in parentheses. *Robust* uses the robust approach with bias-correction suggested by Calonico, Cattaneo, and Titiunik (2014). Optimal bandwidths, a linear polynomial of the ranking score and uniform kernels are used in all the specifications. Selectivity is average baseline ability of student peers in the same degree program. Ability distance is the difference between selectivity and own baseline ability. Rank is the position in the baseline-ability ranking from 0 to 1 in the same degree program. *Mean* is the mean of the outcome variable just below the cutoff in the untrimmed sample.

Presentation

RDD: Impact of selectivity of HE program, least overconfident

Table: Bounds on selectivity and rank effects in higher education course (RD) by overconfidence

	(1)	(2)	(3)	(4)	(5)	(6)
	Selectivity		Ability distance		Rank	
	Lower bound	Upper bound	Lower bound	Upper bound	Lower bound	Upper bound
Robust - Low	0.311	0.413	-0.230	0.219	-0.079	0.091
	(0.289)	(0.274)	(0.306)	(0.342)	(0.122)	(0.126)
Bandwidth	97	97	97	97	97	97
Observations	311	312	311	312	311	312
Mean	-0.098	-0.098	-0.021	-0.021	0.485	0.485

Notes: In this table we report the bounds on the estimate for coefficient δ in regression equation (2). Standard errors clustered at the school level are shown in parentheses. *Robust* uses the robust approach with bias-correction suggested by Calonico, Cattaneo, and Titiunik (2014). Optimal bandwidths, a linear polynomial of the ranking score and uniform kernels are used in all the specifications. Selectivity is average baseline ability of student peers in the same degree program. Ability distance is the difference between selectivity and own baseline ability. Rank is the position in the baseline-ability ranking from 0 to 1 in the same degree program. *Mean* is the mean of the outcome variable just below the cutoff in the untrimmed sample.

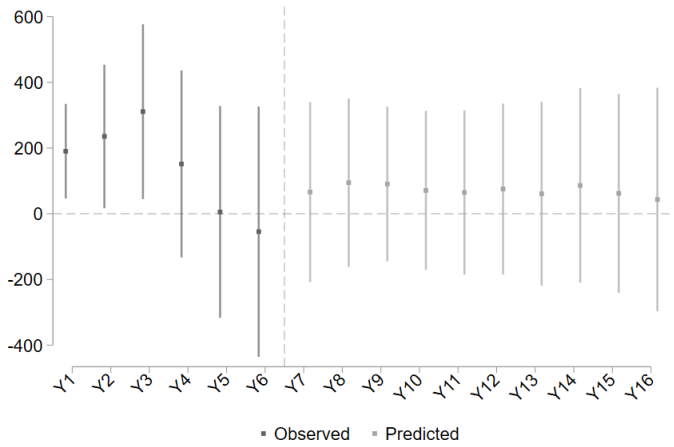
Presentation

RDD: Earning effects by overconfidence

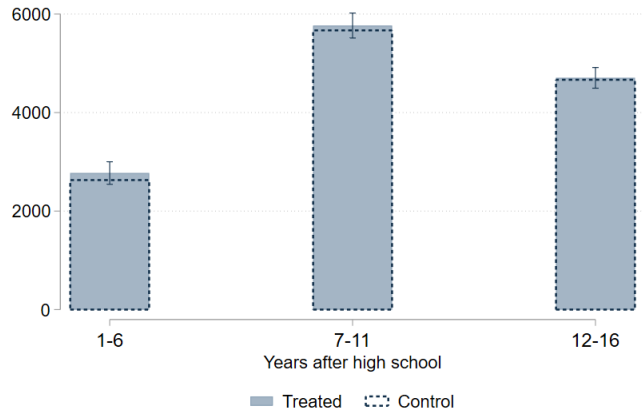
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	In a given year						Average across years		
	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	Years 1-6	Years 7-11	Years 12-16
Robust - High	-52.555	-76.956	830.815	1606.291	109.469	-4.709	103.083	-236.323	-1090.371
	(425.721)	(832.804)	(1682.171)	(2394.818)	(1382.605)	(1554.168)	(1160.737)	(1023.102)	(1105.424)
Bandwidth	41.529	39.529	41.269	36.190	40.020	52.611	36.428	68.298	36.576
Bandwidth obs.	192	208	215	185	204	250	208	373	217
Mean	662.012	496.777	944.577	2,138.712	2,918.167	4,066.827	1,806.012	6,125.628	6,430.049
Robust - Low	-1516.938**	-874.235	-1156.478	-308.721	-2540.748*	-4563.825**	-1651.310**	-1908.334*	-266.315
	(713.175)	(892.652)	(946.603)	(1187.595)	(1426.411)	(2009.059)	(802.285)	(1115.746)	(1025.777)
Bandwidth	53.128	55.483	70.363	97.239	75.937	80.805	69.747	60.563	67.685
Bandwidth obs.	180	188	241	337	252	243	269	243	265
Mean	1,590.028	1,594.653	1,677.043	2,992.074	4,203.526	5,017.841	2,490.686	5,828.689	4,898.594
p-value High=Low	0.025	0.211	0.271	0.436	0.152	0.073	0.140	0.298	0.727

Presentation

RCT bottom 85%: Earnings



(a) Earnings, in a given year



(b) Earnings, 6-year average

- No gender differences.

Presentation

Mechanism/1: No impacts on educational attainment

	(1)	(2)	(3)	(4)	(5)	(6)
	Selective College			Any Institution		
	Ever enrolled	Graduation	Dropout	Ever enrolled	Graduation	Dropout
Treatment	0.021	0.010	0.007	-0.009	-0.002	-0.007
	(0.017)	(0.012)	(0.007)	(0.017)	(0.019)	(0.013)
Total obs.	11916	11916	11916	11916	11916	11916
Mean	0.138	0.089	0.034	0.672	0.414	0.258

- No evidence of peer effects on education outcomes.

Presentation

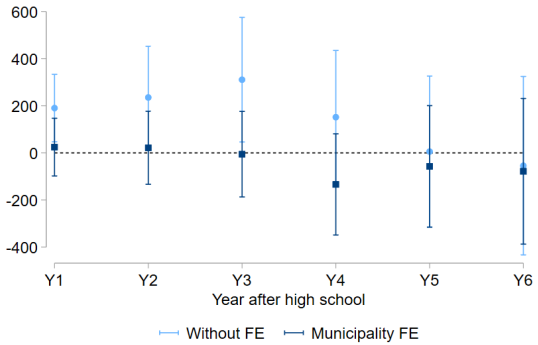
Mechanism/2: No disengagement from college sector

	(1)	(2)	(3)	(4)	(5)	(6)
	PSU	PSU	PSU	PSU	PSU	PSU
	2018	2019	2020	2021	2022	2023
Treatment	-0.017	-0.002	0.001	0.000	0.001	0.001
	(0.027)	(0.012)	(0.005)	(0.003)	(0.002)	(0.002)
Total obs.	11916	11916	11916	11916	11916	11916
Mean	0.764	0.096	0.034	0.017	0.012	0.010

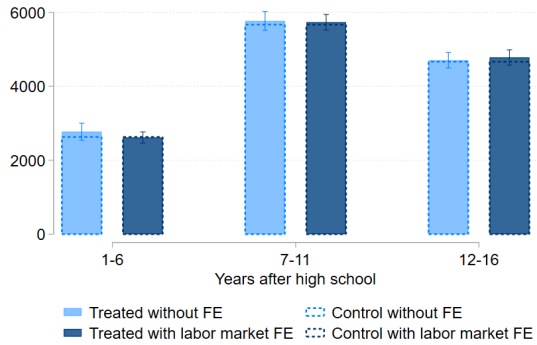
- No discouragement from taking college entrance exam, in current or future rounds

Presentation

Mechanism/3: Spillover effects in local labor markets



(a) Earnings, in a given year



(b) Earnings, average until a given year

- Positive effects vanish when comparing students in treated schools to students in control schools in the same labor market (municipality, robust to other definitions)
- Consistent with less competition from top 15% students while entering local labor markets

→ Preferential admissions can affect untargeted disadvantaged students through equilibrium effects in local labor markets.

Presentation

Getting away from the cutoff in RDD

How Far Can Inclusion Go? If we want to avoid the unintended consequences on earnings, an around top 10% plan would achieve that goal. Presentation

Assumptions

Assumption (Conditional balance of potential outcomes)

Conditional mean independence of the PACE school assignment:
$E[Y_{00}|X] = E[Y_{00}|X, S = 1] = E[Y_{00}|X, S = 0]$,
$E[Y_{10}|X] = E[Y_{10}|X, S = 1] = E[Y_{10}|X, S = 0]$,
$E[Y_{11}|X] = E[Y_{11}|X, S = 1] = E[Y_{11}|X, S = 0]$.
Potential admission eligibility likelihood: $Pr(a = 1|X) = Pr(a = 1|X, S)$, $S = 0, 1$.

Assumption (Conditional balance across eligibility status)

$E[Y_{00}|X, S = 1] = E[Y_{00}|X, S = 1, A = 1] = E[Y_{00}|X, S = 1, A = 0]$

Assumption (Covariate balance)

$f(X) = f(X|S = 1) = f(X|S = 0)\ \forall X \in \mathcal{X}$.

Presentation

Linking RCT with RDD: earnings

$$\underbrace{\int E[Y_{1*} - Y_{00} \mid X] f(X)\, dX}_{\text{ATE}_S \rightarrow \text{ from RCT}} - \underbrace{\pi}_{\text{Fraction eligible}} \underbrace{\int E[Y_{11} - Y_{10} \mid X, A{=}1, S{=}1] f(X \mid A{=}1, S{=}1)\, dX}_{\text{ATT}_A \rightarrow \text{ from RDD getaway}}$$

$$= \underbrace{\int E[Y_{10} - Y_{00} \mid X] f(X)\, dX}_{\text{Determined residually} = 225.827}$$

Limited effect of the "PACE school treatment" on yearly earnings.

Presentation