

Income Segregation and Intergenerational Mobility Across Colleges in the United States

Raj Chetty, Harvard University
John N. Friedman, Brown University
Emmanuel Saez, UC-Berkeley
Nicholas Turner, Federal Reserve Board
Danny Yagan, UC-Berkeley

February 2020

The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors and do not necessarily represent the views of the U.S. Department of Treasury, the Federal Reserve Board of Governors, or the College Board.

Introduction

- What role do colleges play in intergenerational income mobility?
 - Large returns to college attendance suggest that higher education can be an important pathway to upward mobility
 - But inequality in access between high- and low-income families may limit (or even reverse) this effect

This Paper

- We analyze how changes in the colleges that students attend would affect rates of intergenerational mobility in the U.S. by estimating three sets of parameters:
 1. [Segregation] Parental income distributions by college
 2. [Outcomes] Students' earnings outcomes conditional on parental income by college
 3. [Causal share] Portion of variation in students' earnings outcomes that is due to colleges' causal effects ("value-added")
- Construct publicly available statistics on the first two elements using data on all college students from 1999-2013 (previously released as "Mobility Report Cards" in 2017)
- Combine those statistics with data on SAT and ACT scores and estimates of colleges' causal effects based on Dale-Krueger (2002) method

Three Main Findings

1. Colleges are highly segregated by parent income, both across and within tiers
 - Students from the top 1% of family income are 77 times more likely to attend an Ivy-plus institution than those from the poorest quintile of families
 - Colleges are as segregated as neighborhoods in the average American city by parental income

Three Main Findings

2. Gaps in earnings between children from rich and poor families who attend the same college are relatively small
 - Much of the gap in earnings between children from low vs. high income families is accounted for by variation between colleges
 - Most colleges with high levels of student earnings have few students from low-income families
 - Certain “high mobility” colleges, such as the City University of New York, have both a high share of low-income students and excellent outcomes

Three Main Findings

3. Changing allocations of students across colleges could increase intergenerational mobility in the U.S. substantially
 - Currently, rich students are much more likely to attend highly selective colleges than students from low-income families even holding fixed SAT/ACT scores
 - Parental income segregation would thus fall substantially if rich and poor students attended similar colleges conditional on test scores
 - At elite private (Ivy-plus) colleges, middle-class children are under-represented relative to their test scores, but children from the bottom quintile are *not*
 - Giving low-income students a preference similar to that given to legacy students would equate parental income distributions across all colleges
 - Income-neutral allocations would reduce the intergenerational persistence of income by 15%; need-affirmative allocations would reduce it by 25%

Outline

1. Segregation: Parents' Marginal Income Distributions by College
2. Outcomes: Distributions of Students' Earnings by College
3. Intergenerational Mobility: Counterfactual Student Allocations

Data

- Data source: de-identified data from 1996-2014 income tax returns
 - Includes data on income of non-filers through information returns filed by employers (W-2 forms)
- Primary sample: all children in 1980-82 birth cohorts claimed as dependents by tax filers in the U.S.
 - Earliest cohorts where we can link almost all children to parents
 - Approximately 11 million children

Measuring College Attendance

- All Title IV institutions report student attendance to IRS on Form 1098-T
 - 1098-T data cover 95% of enrolled students; students who pay no tuition sometimes not covered
 - Use Dept. of Ed data (NSLDS) on students receiving Pell grants to identify these students
- Baseline: define college attendance as most-attended college between ages 19-22
 - Similar results obtained with alternative definitions (e.g., college attended at age 20)
- Following established disclosure standards, all college-specific numbers are estimates (approx. +/- 1% measurement error)

Measuring Test Scores

- Link data on SAT scores from the College Board and ACT scores to tax data to measure test scores at point of college application
 - Map ACT scores to SAT scores using a standard concordance table

Part 1

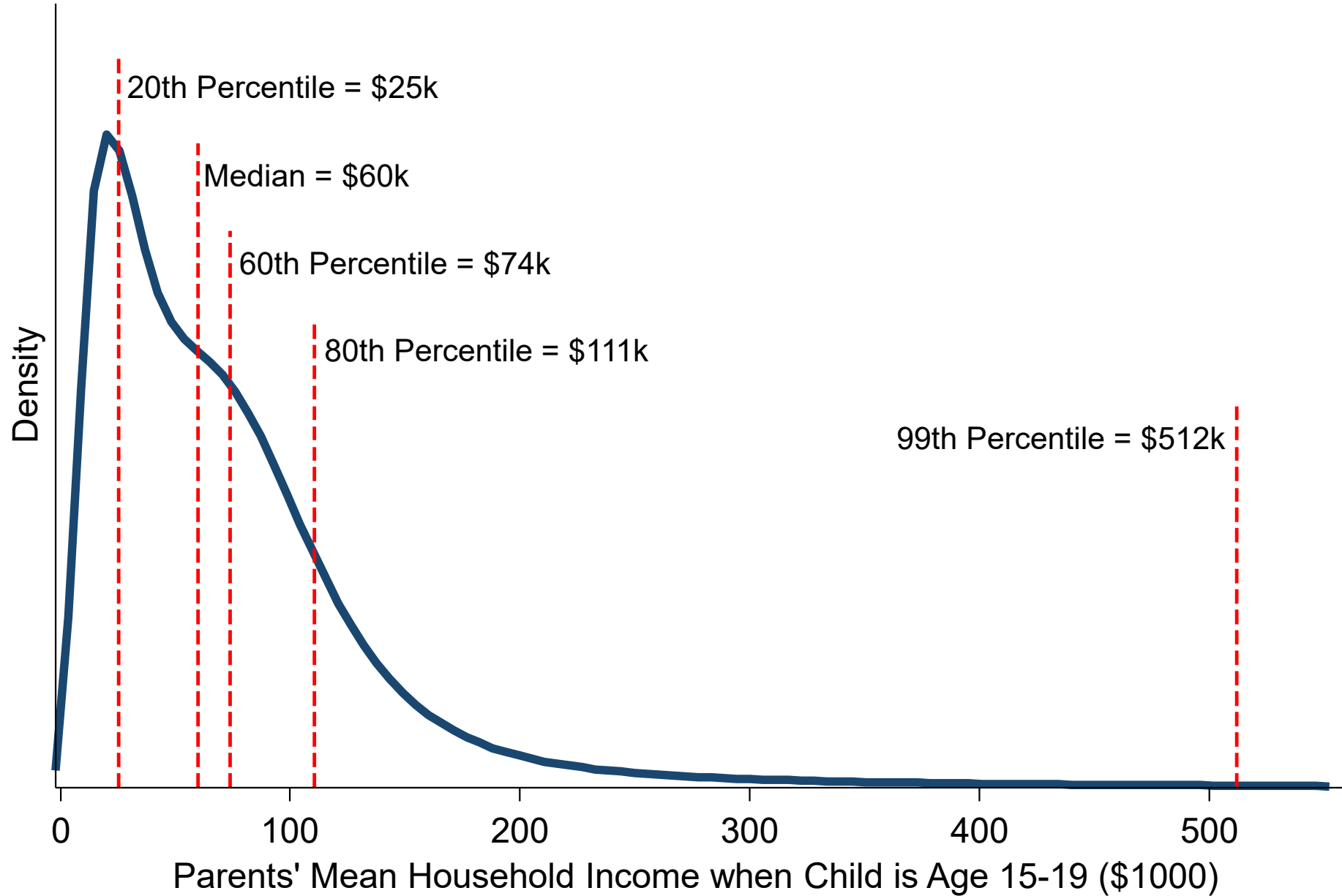
Segregation: Parents' Income Distributions by College

Measuring Parent Income

- Parent income: mean pre-tax **household** income during five year-period when child is aged 15-19
 - For filers, use Adjusted Gross Income reported on form 1040
 - For non-filers, use W-2 wage earnings + UI income
 - All incomes measured in 2015 dollars
- Focus on percentile ranks, ranking parents relative to other parents with children in same birth cohort

Parent Household Income Distribution

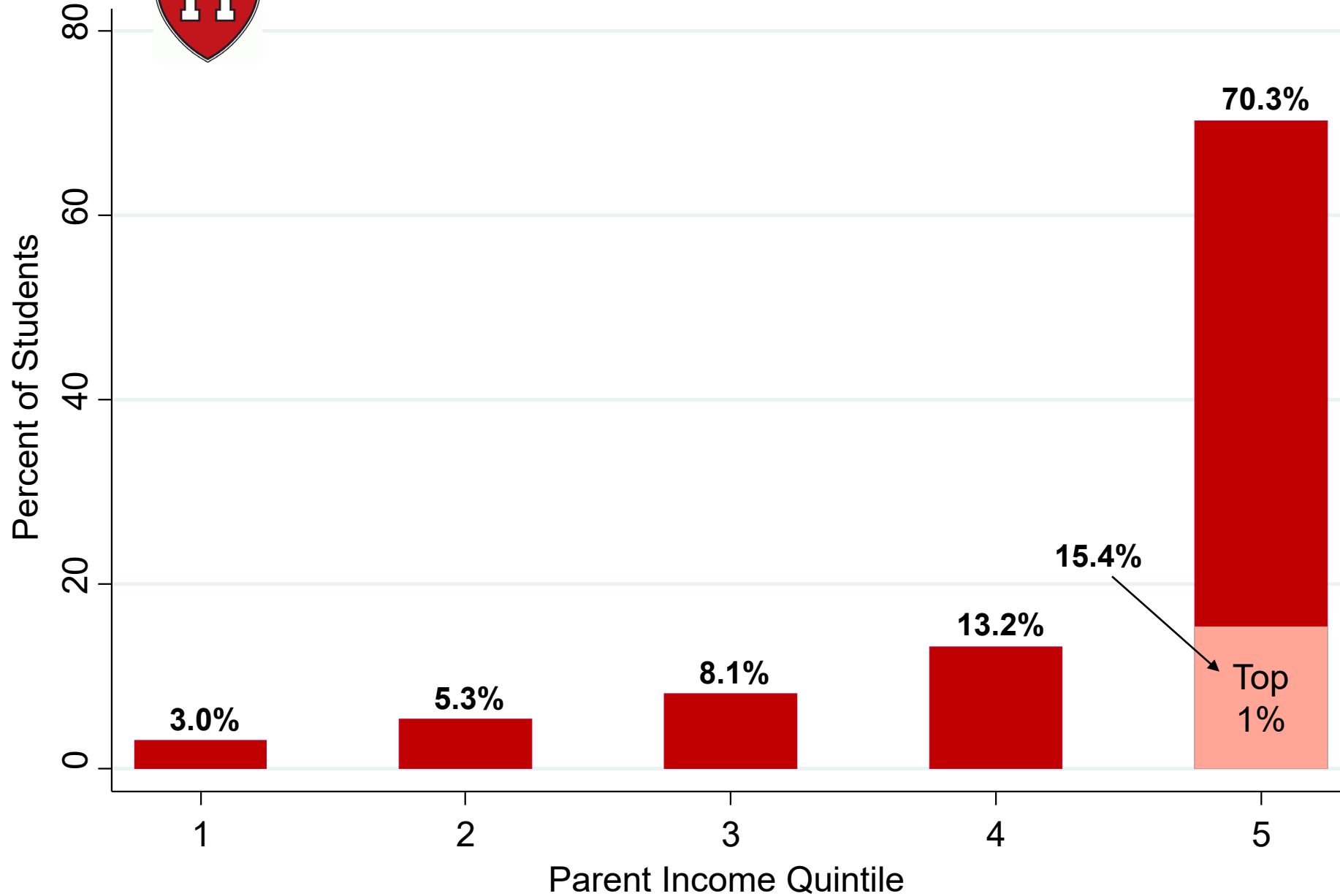
For Parents with Children in 1980 Birth Cohort





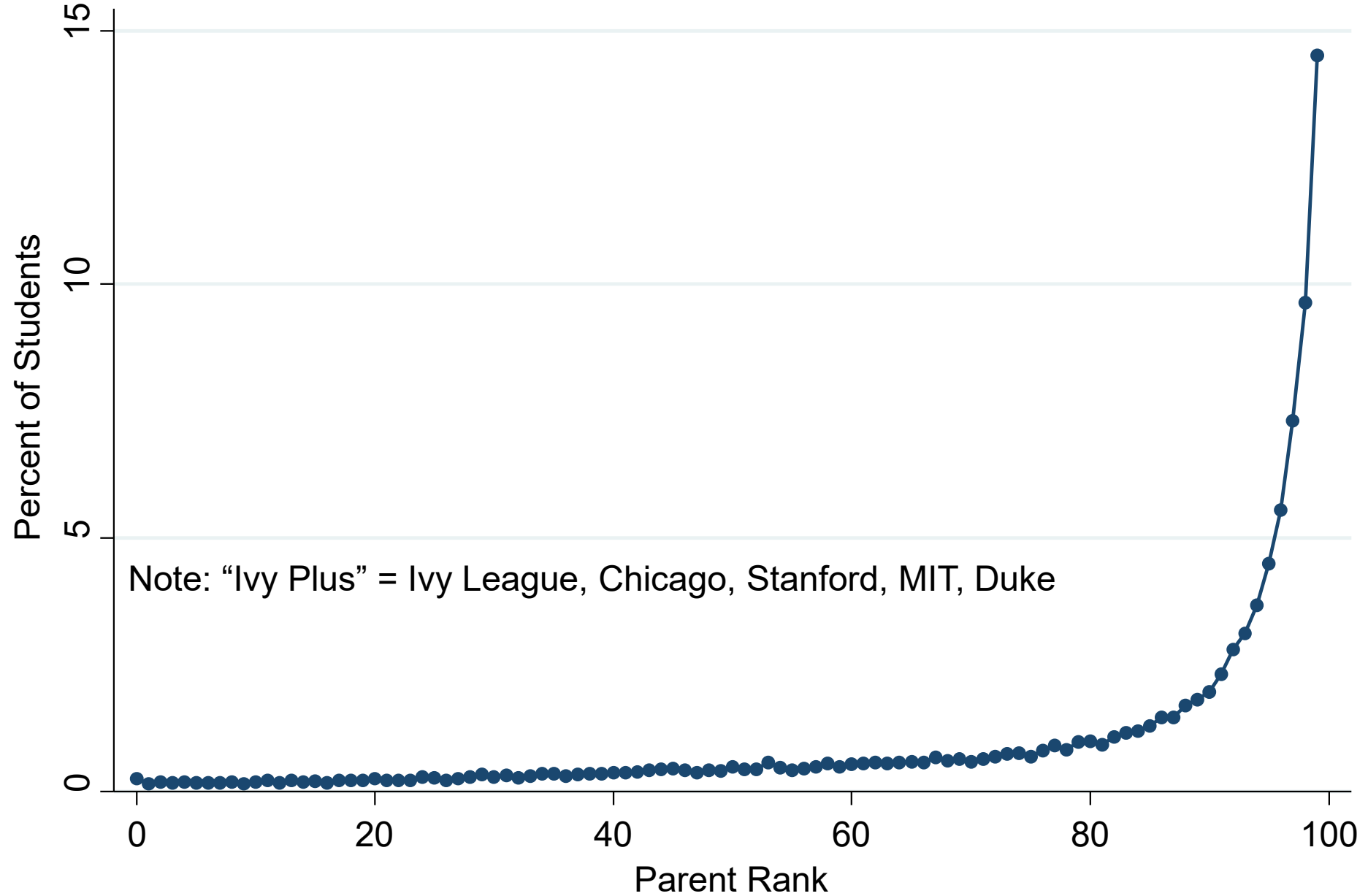
Parent Income Distribution at Harvard

1980-82 Child Birth Cohorts



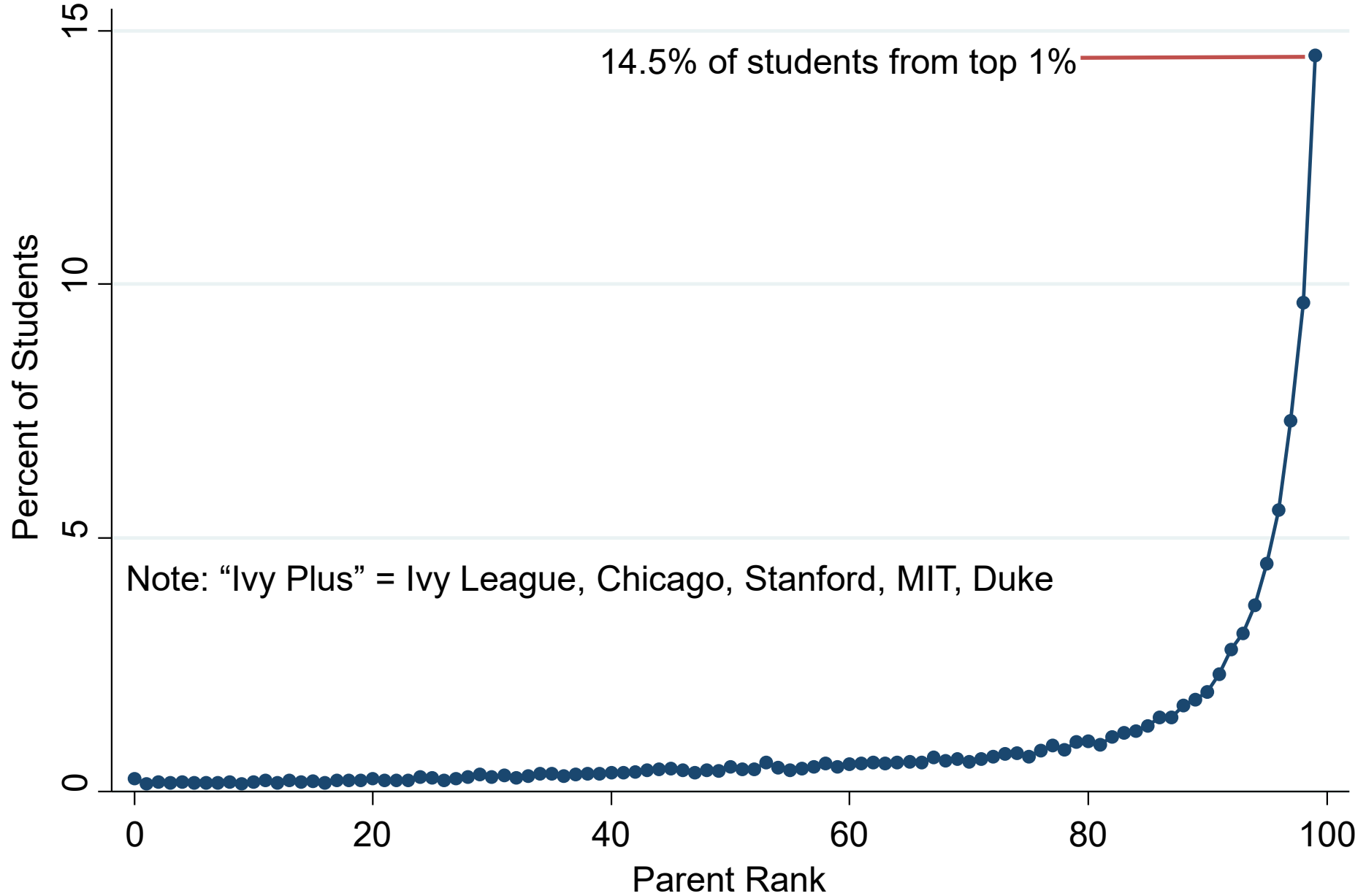
Parent Income Distribution by Percentile

Ivy Plus Colleges



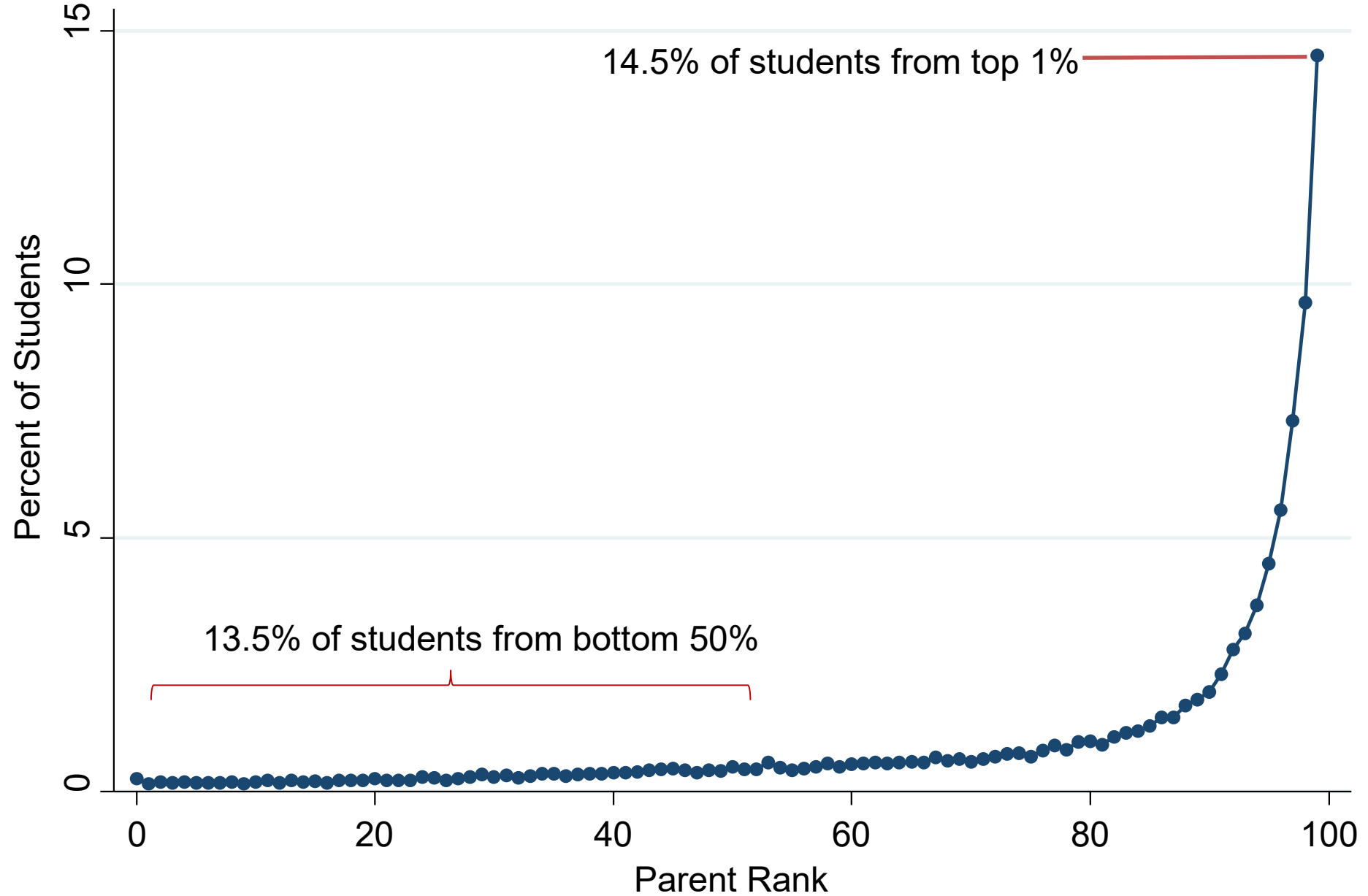
Parent Income Distribution by Percentile

Ivy Plus Colleges



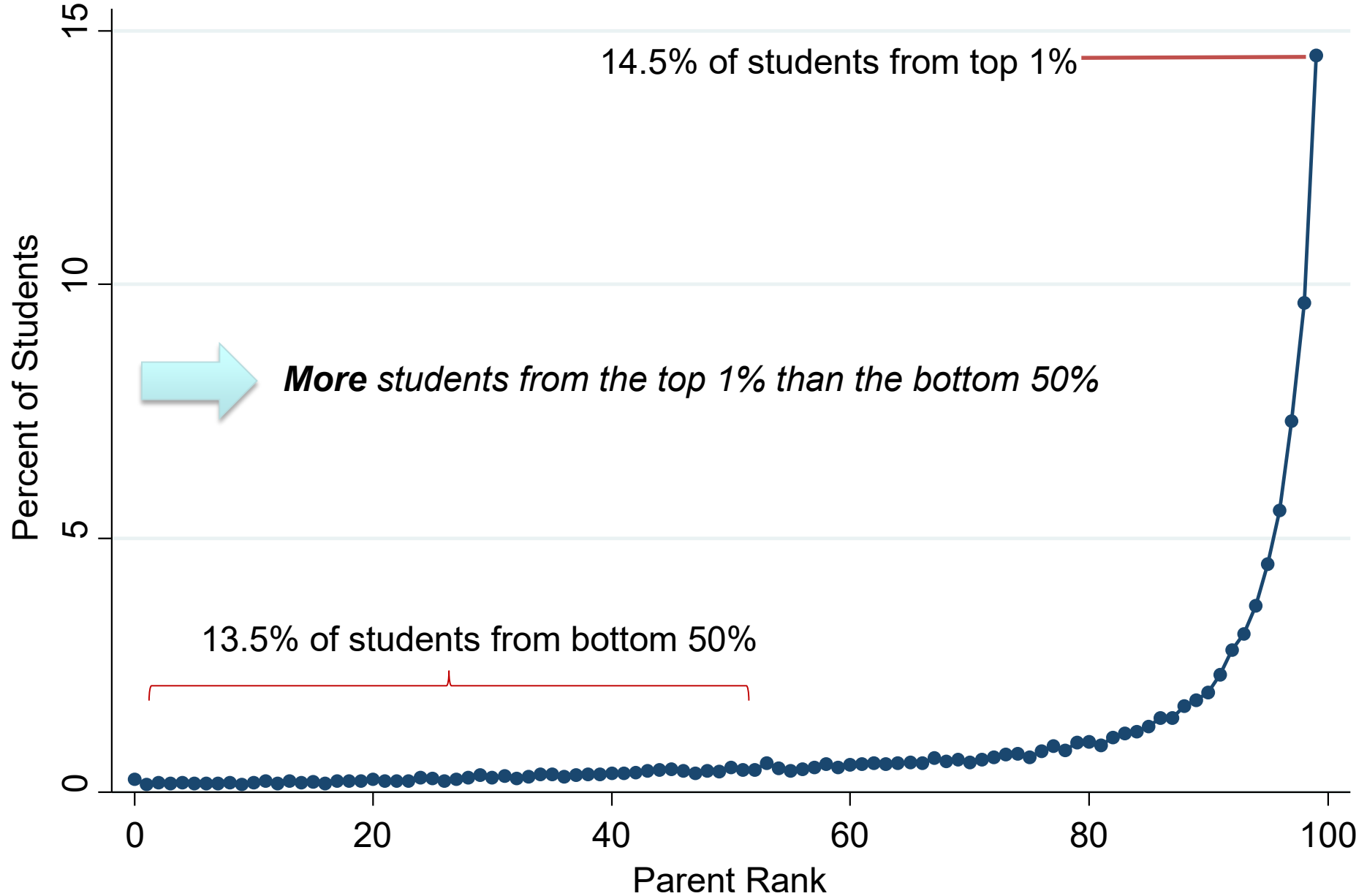
Parent Income Distribution by Percentile

Ivy Plus Colleges



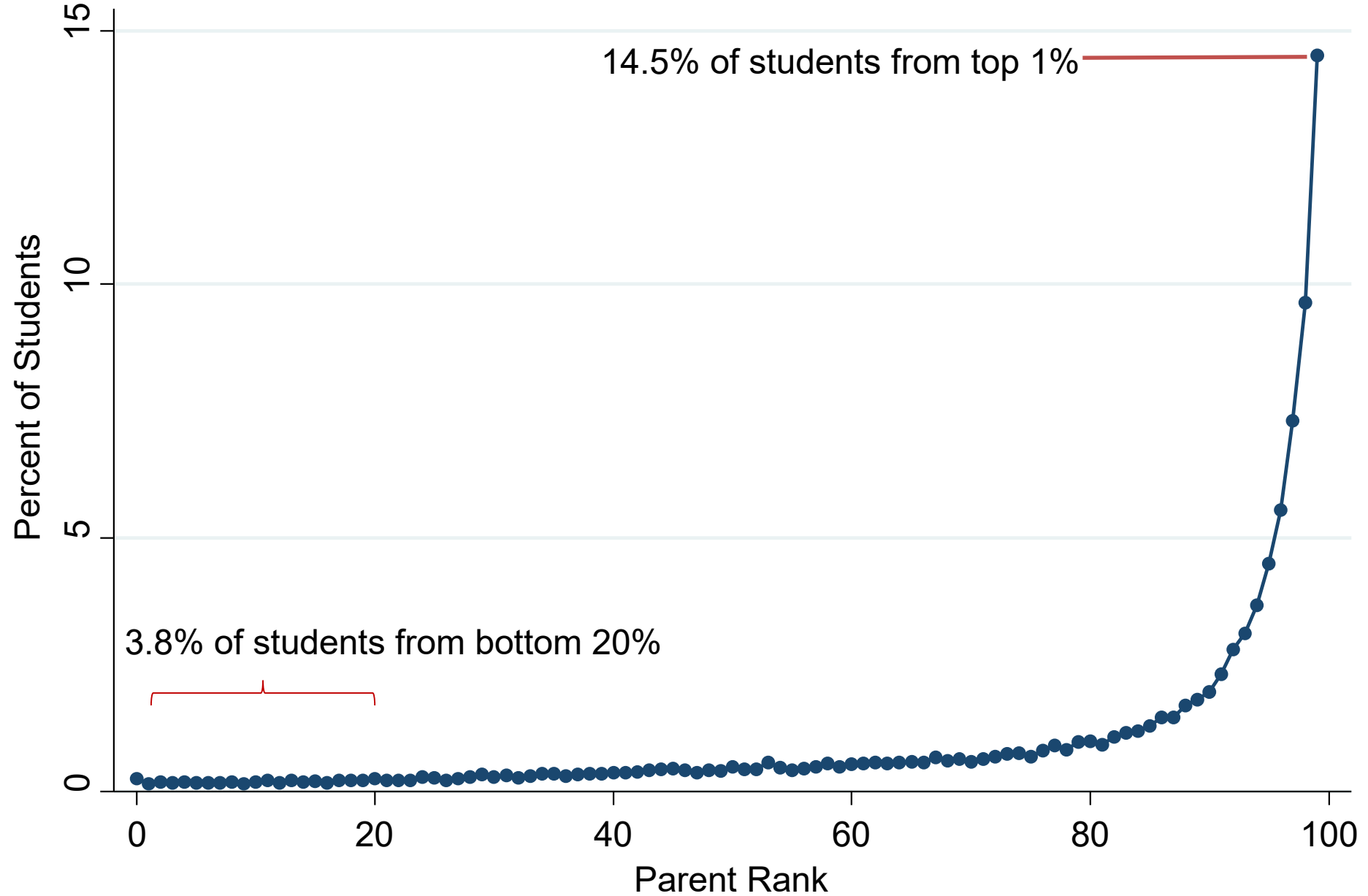
Parent Income Distribution by Percentile

Ivy Plus Colleges



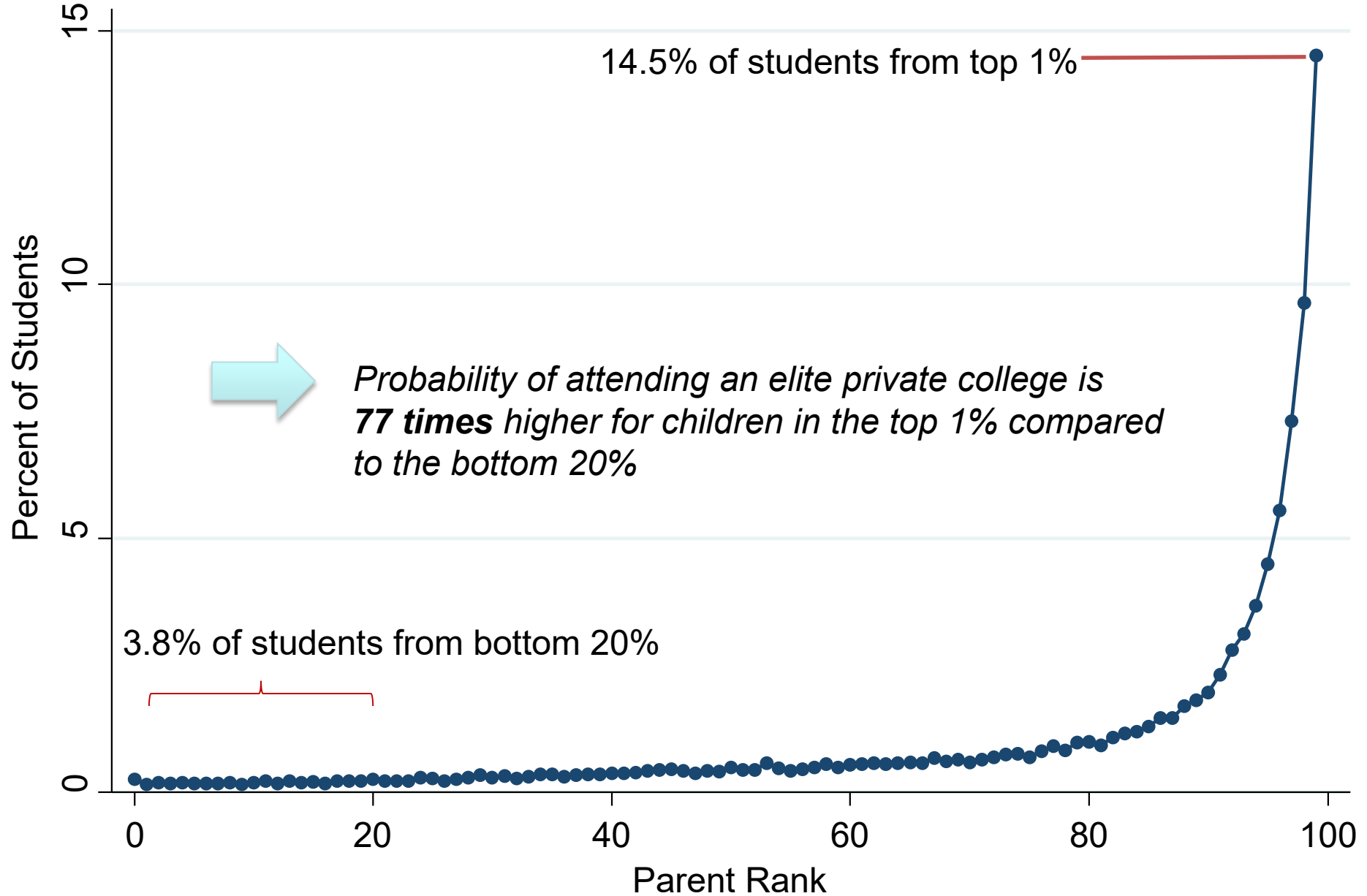
Parent Income Distribution by Percentile

Ivy Plus Colleges

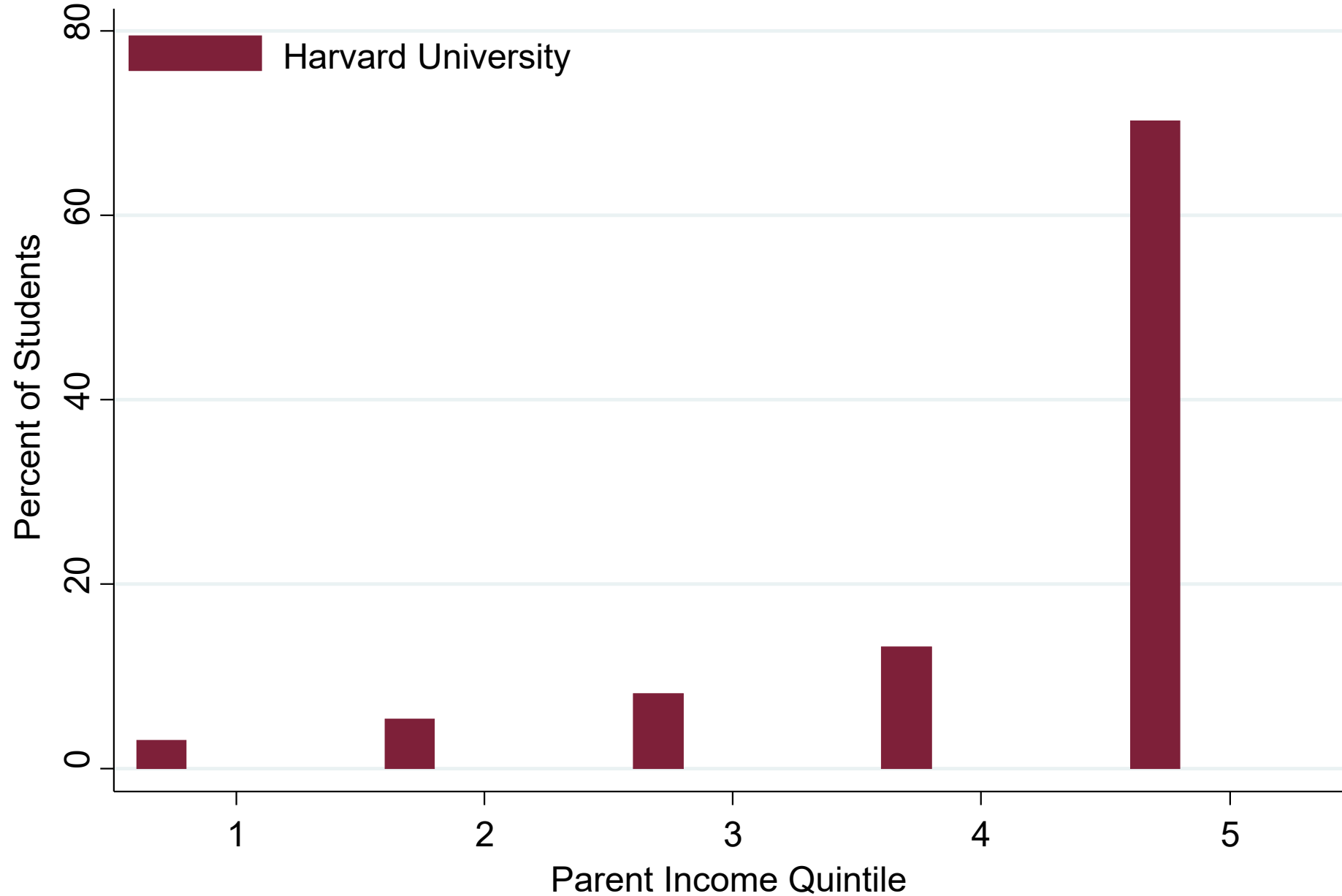


Parent Income Distribution by Percentile

Ivy Plus Colleges

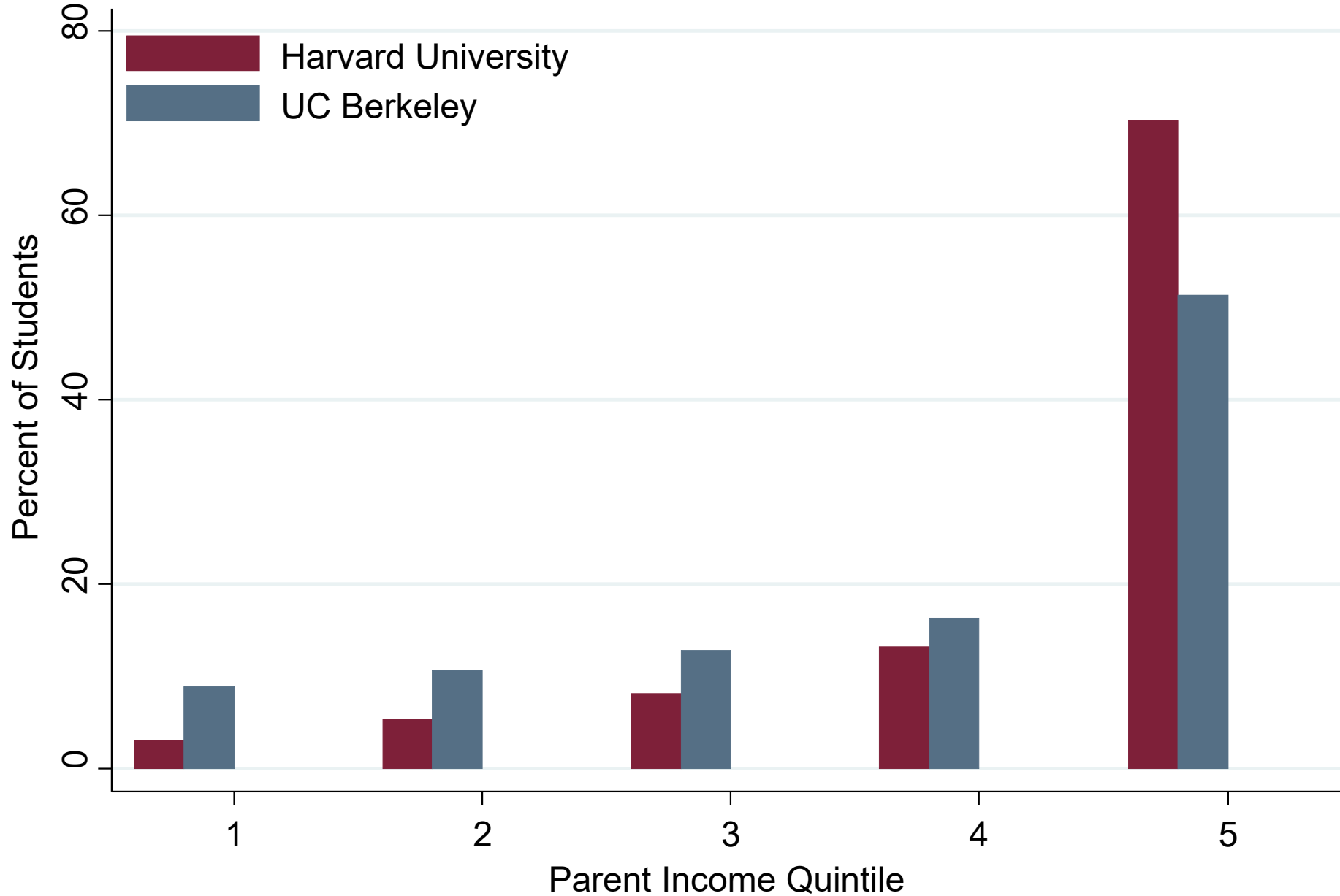


Parent Income Distributions by Quintile for 1980-82 Birth Cohorts At Selected Colleges



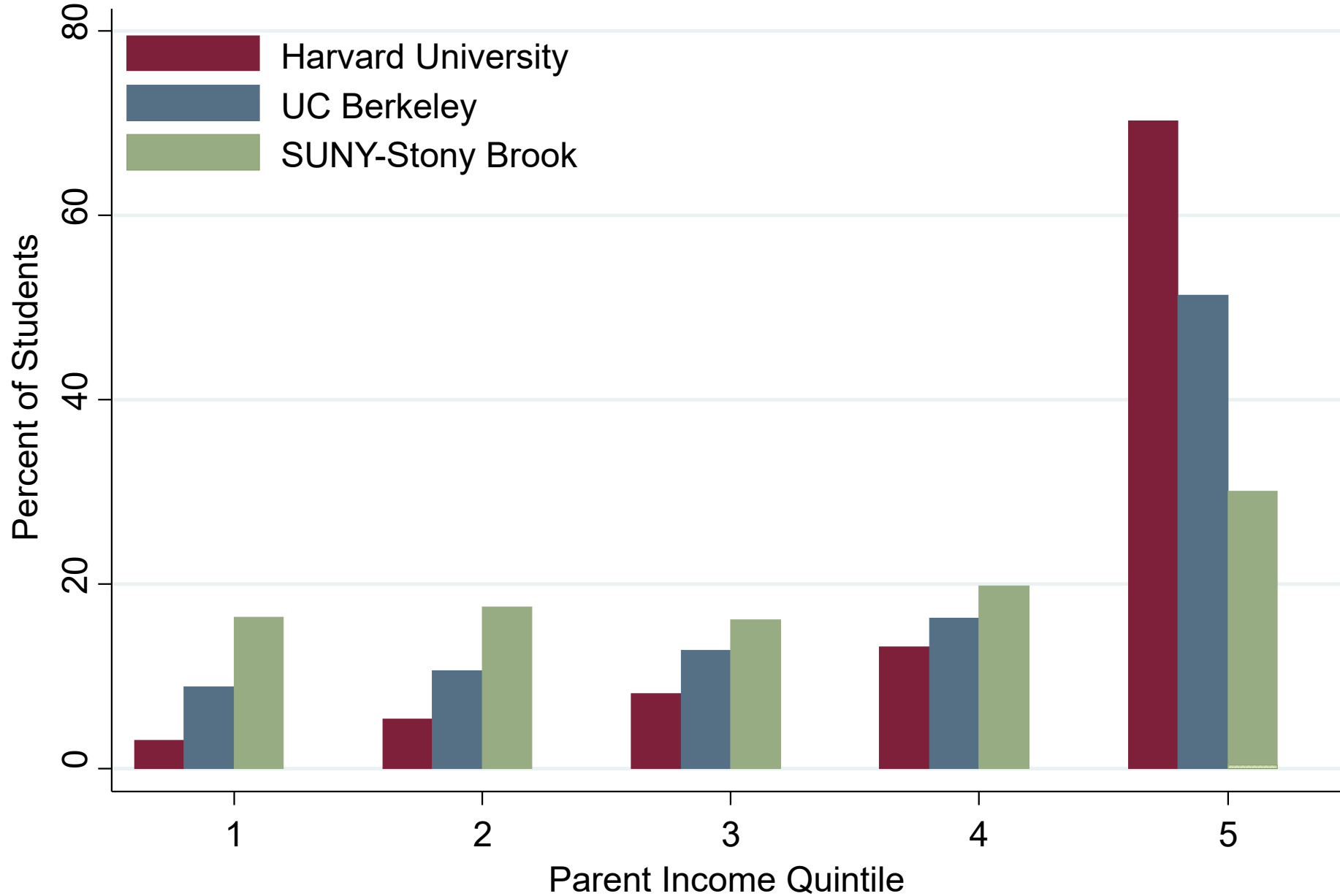
Parent Income Distributions by Quintile for 1980-82 Birth Cohorts

At Selected Colleges

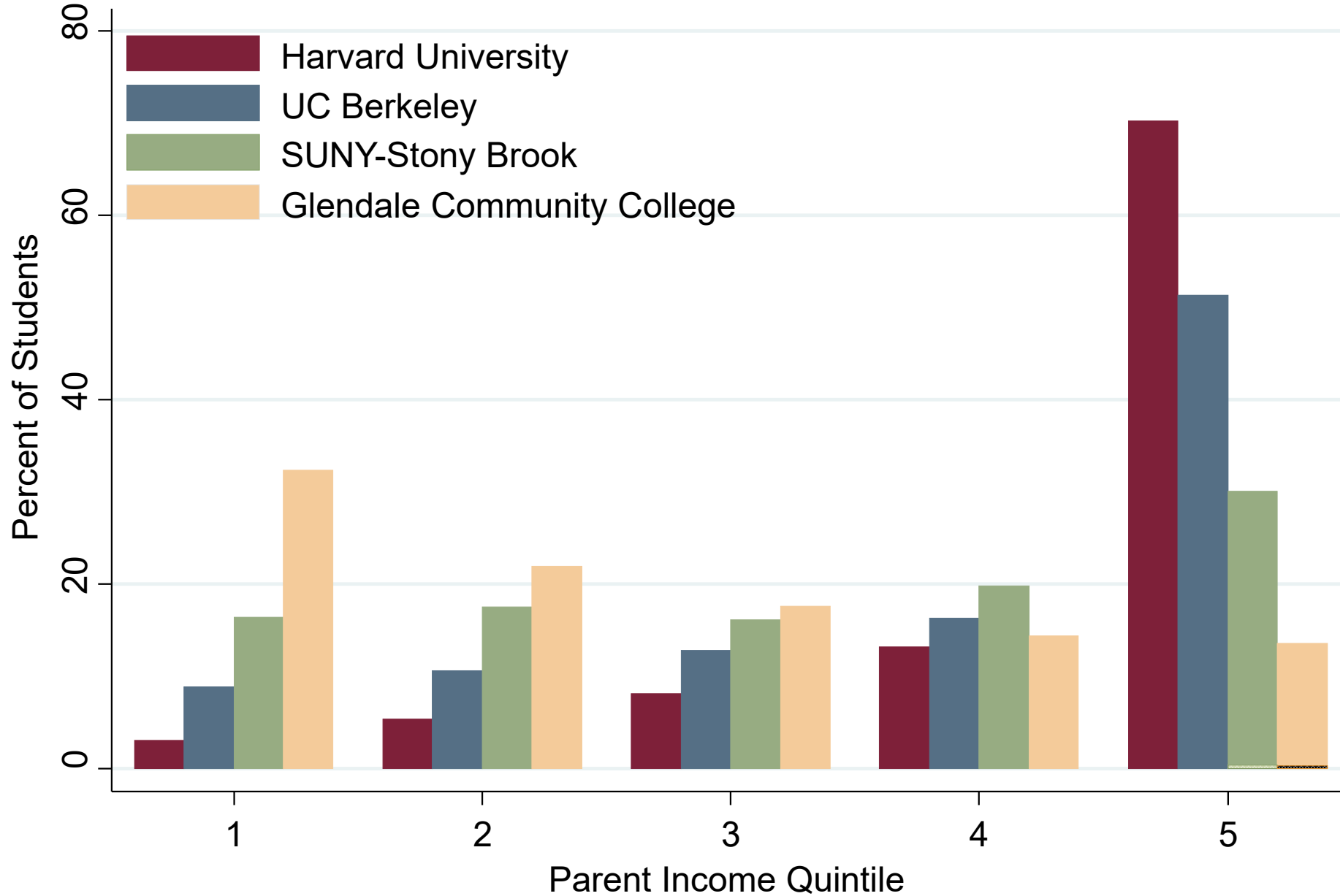


Parent Income Distributions by Quintile for 1980-82 Birth Cohorts

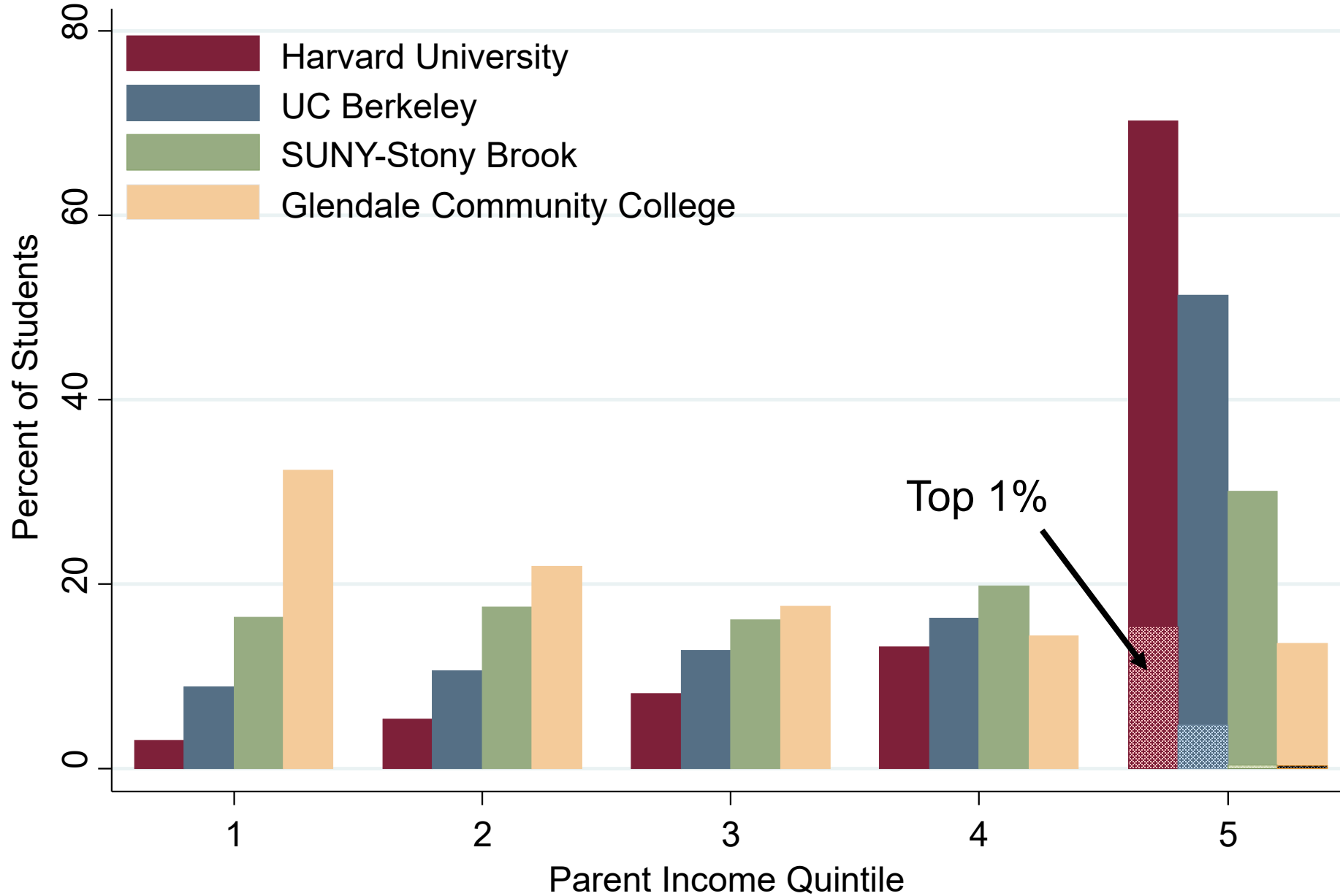
At Selected Colleges



Parent Income Distributions by Quintile for 1980-82 Birth Cohorts At Selected Colleges

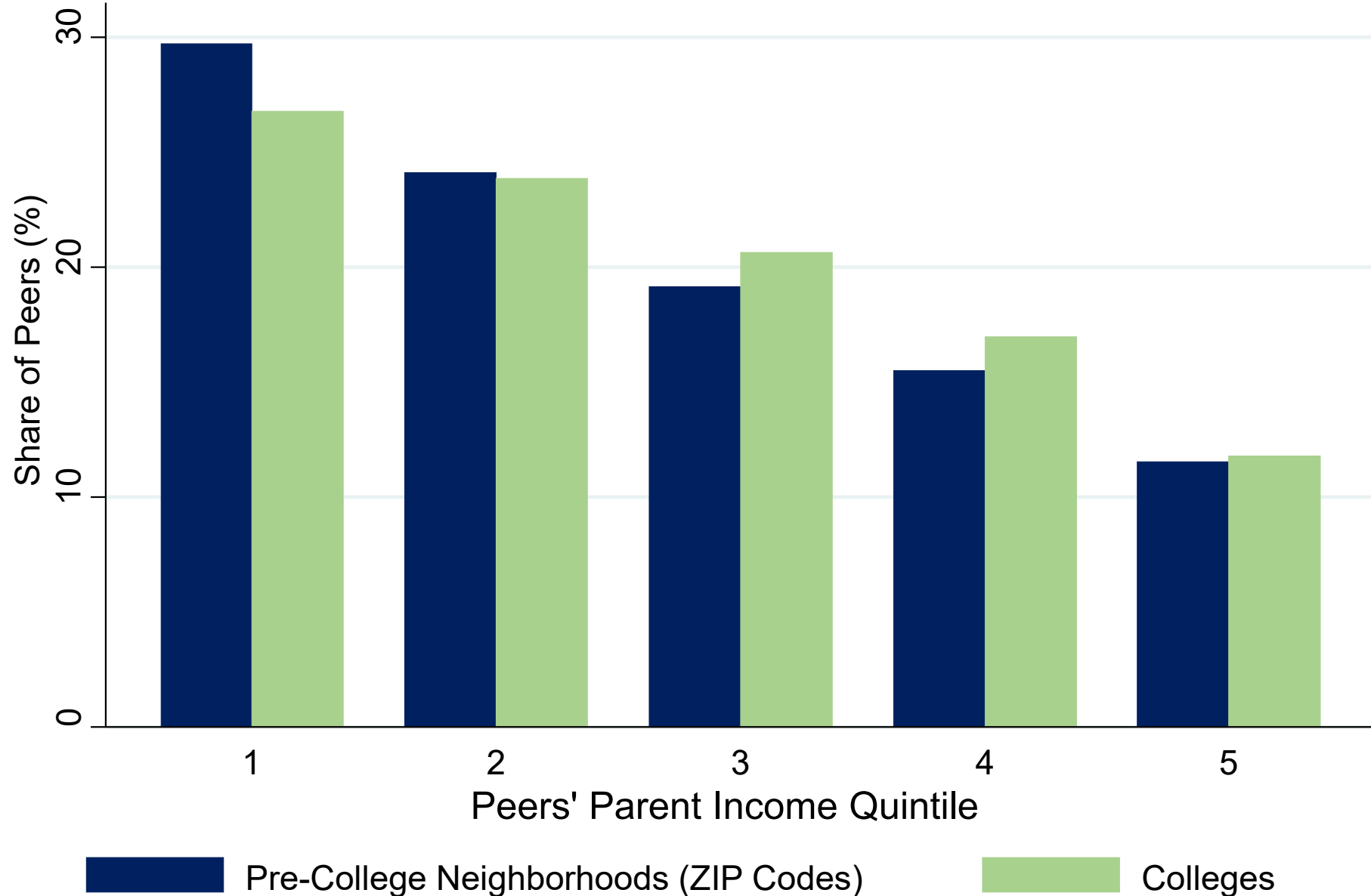


Parent Income Distributions by Quintile for 1980-82 Birth Cohorts At Selected Colleges



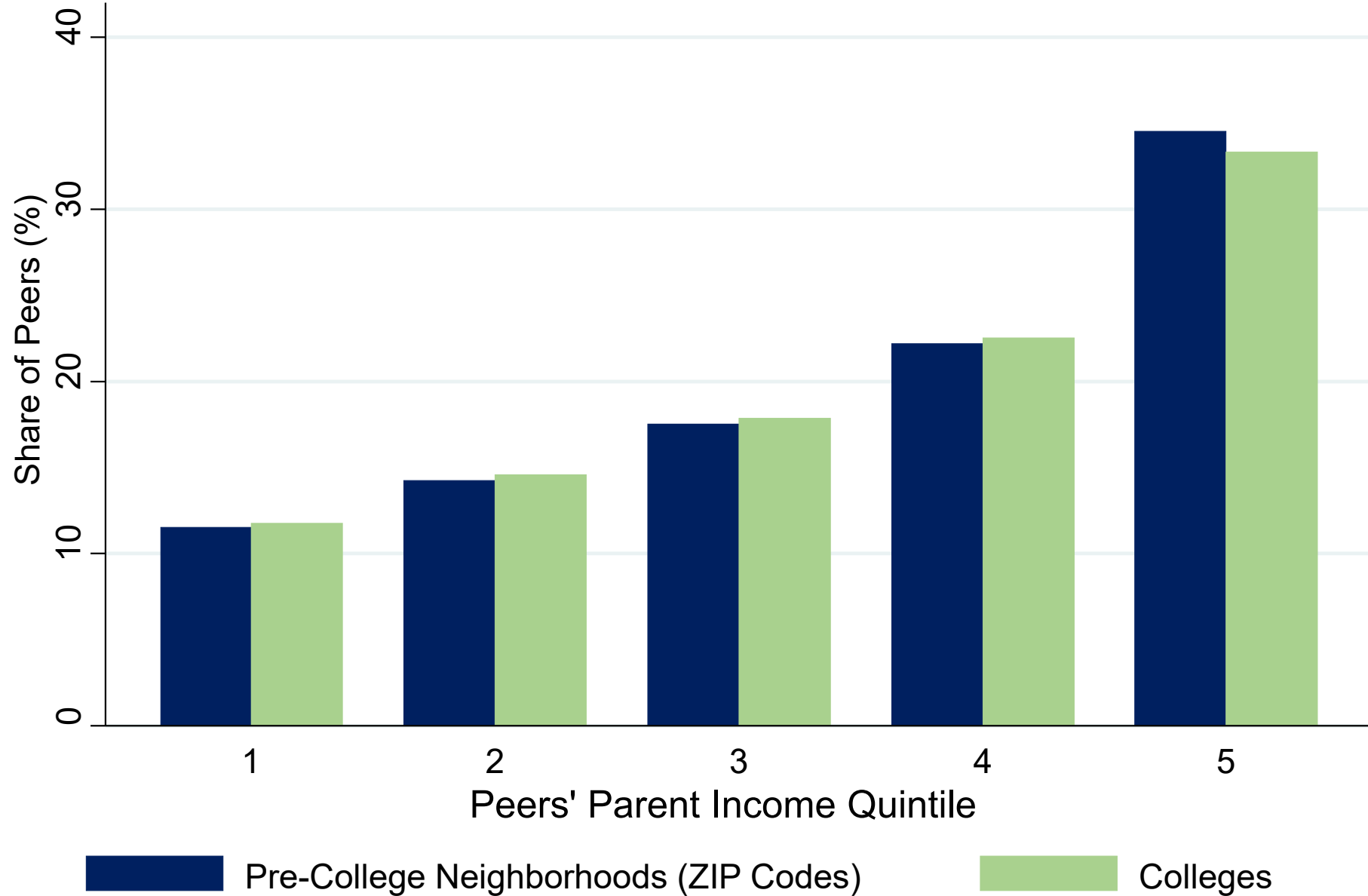
Income Segregation across Colleges vs. Pre-College Neighborhoods

Children from Bottom Quintile



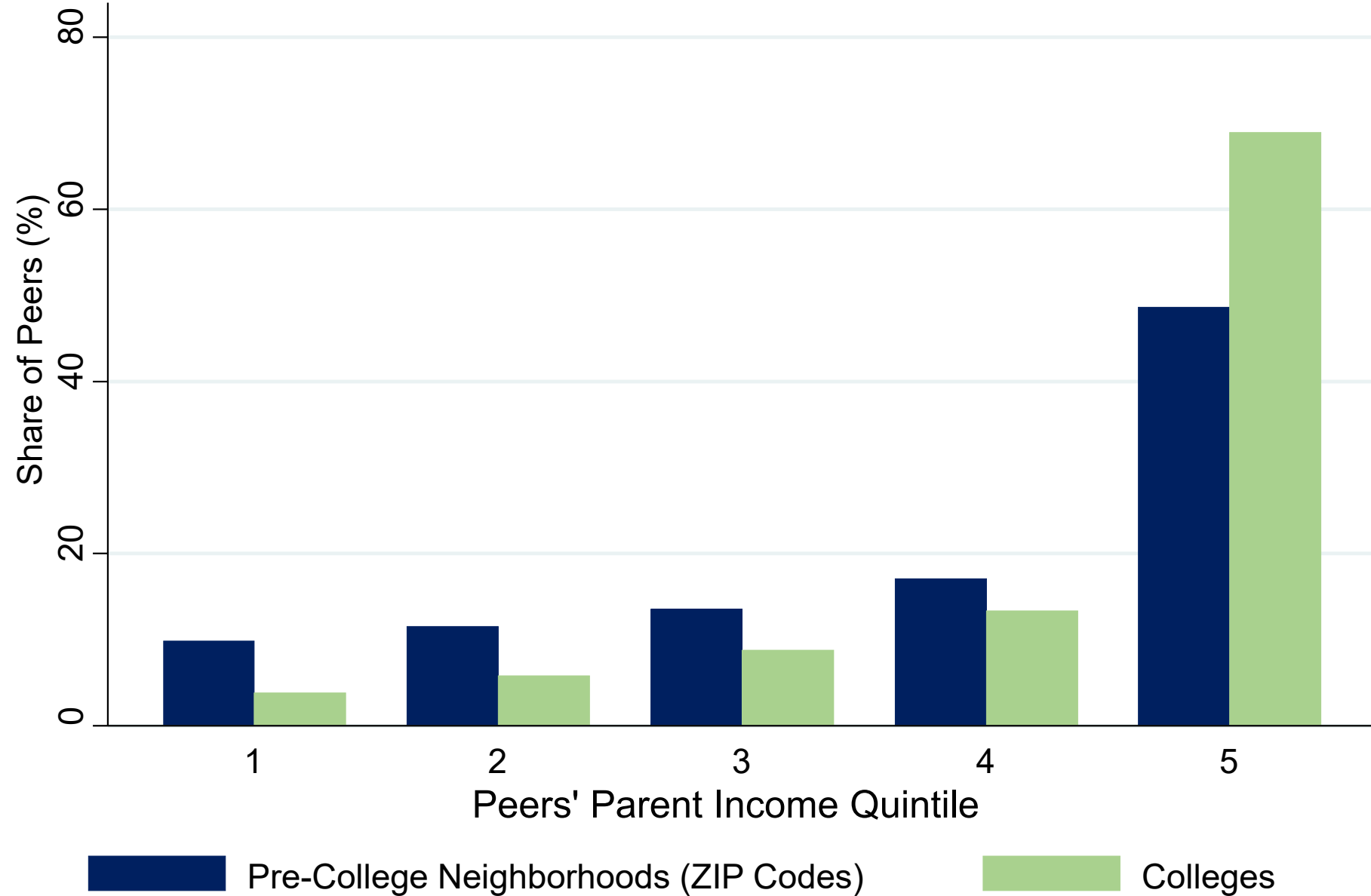
Income Segregation across Colleges vs. Pre-College Neighborhoods

Children from Top Quintile



Income Segregation across Colleges vs. Pre-College Neighborhoods

Children from Top Quintile at Ivy-Plus Colleges



Lessons on Access

- **Fact #1:** Income segregation across colleges is comparable to segregation across Census tracts in the average American city
- Even at the student level, peers in college are not substantially different than those in the home neighborhood
- Income is especially concentrated at elite private schools

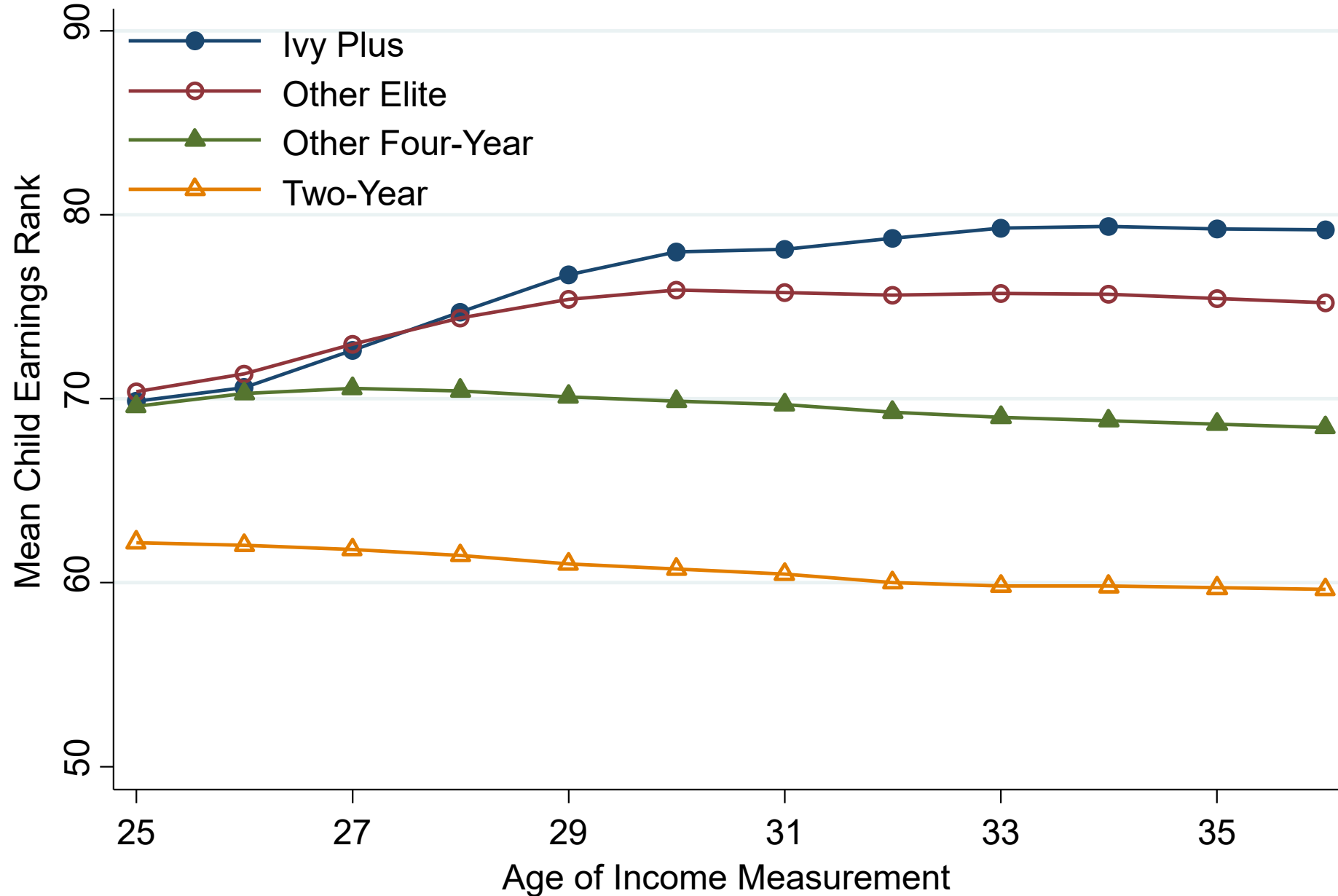
Part 2

Outcomes: Distributions of Student's Earnings by College

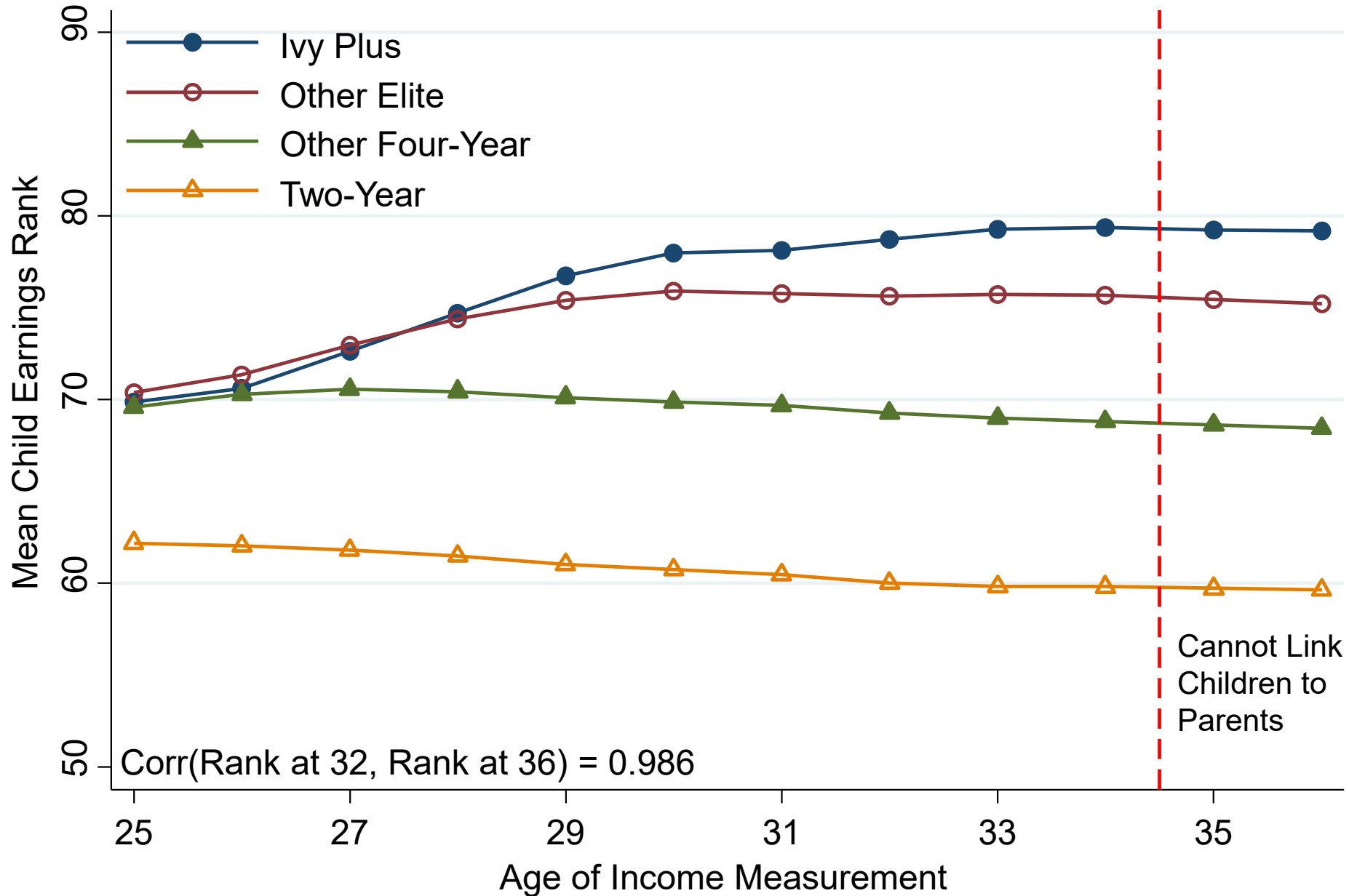
Measuring Student Earnings

- Individual labor earnings = wages + self-emp. income + foreign wages
- Compute percentile ranks by ranking children within birth cohorts
- Using data going back to 1978 cohort, we see that ranks stabilize by age 32 at all colleges

Mean Child Rank vs. Age at Income Measurement, By College Tier



Mean Child Rank vs. Age at Income Measurement, By College Tier

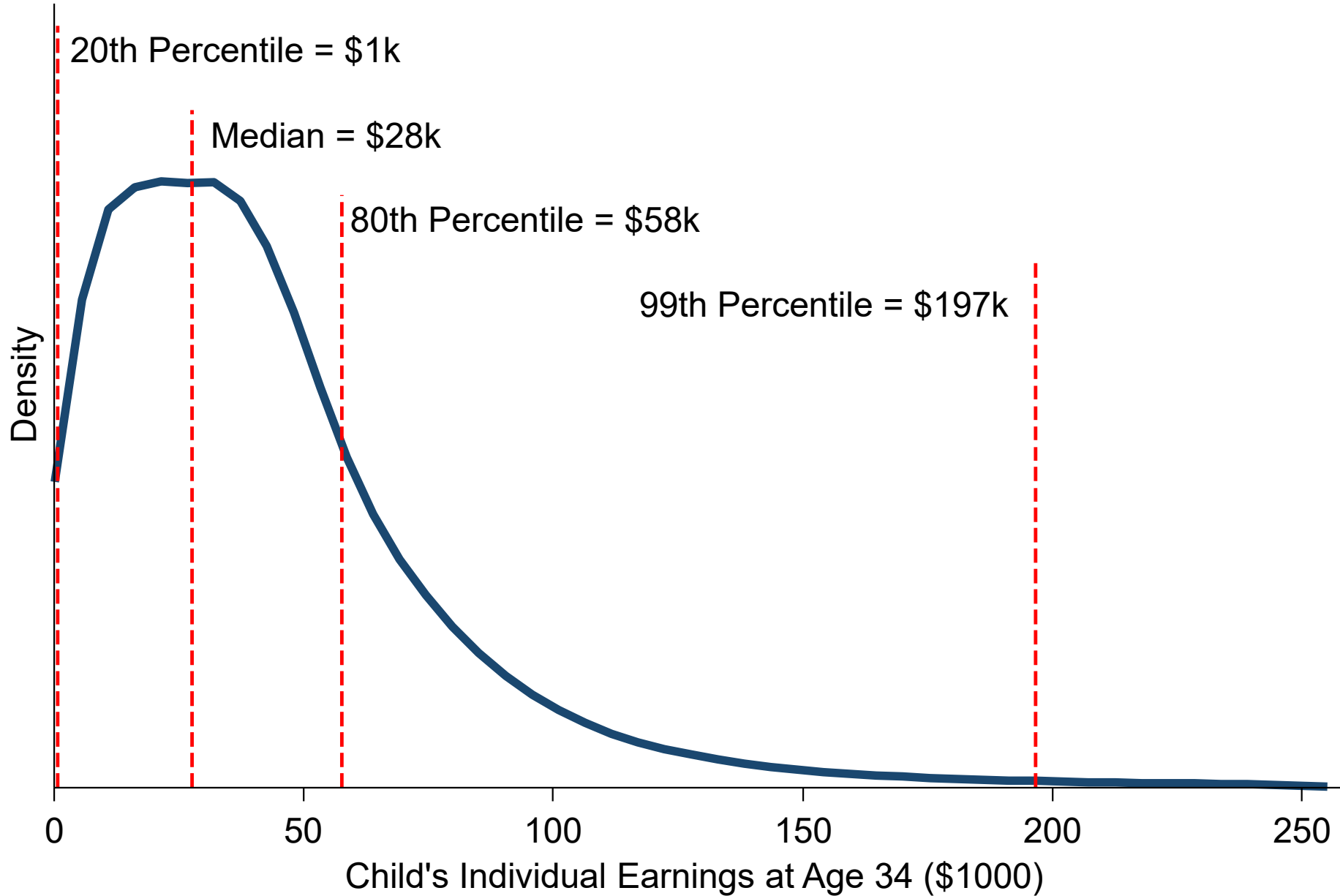


Measuring Student Earnings

- Individual labor earnings = wages + self-emp. income + foreign wages
 - Compute percentile ranks by ranking children within birth cohorts
 - Using data going back to 1978 cohort, we see that ranks stabilize by age 32 at all colleges
 - Broader income concepts (e.g., AGI) differ from individual labor earnings primarily because of marriage
- Baseline definition: individual earnings in 2014, measured at ages 32-34 for 1980-82 birth cohorts

Distribution of Children's Individual Labor Earnings at Age 34

1980 Birth Cohort

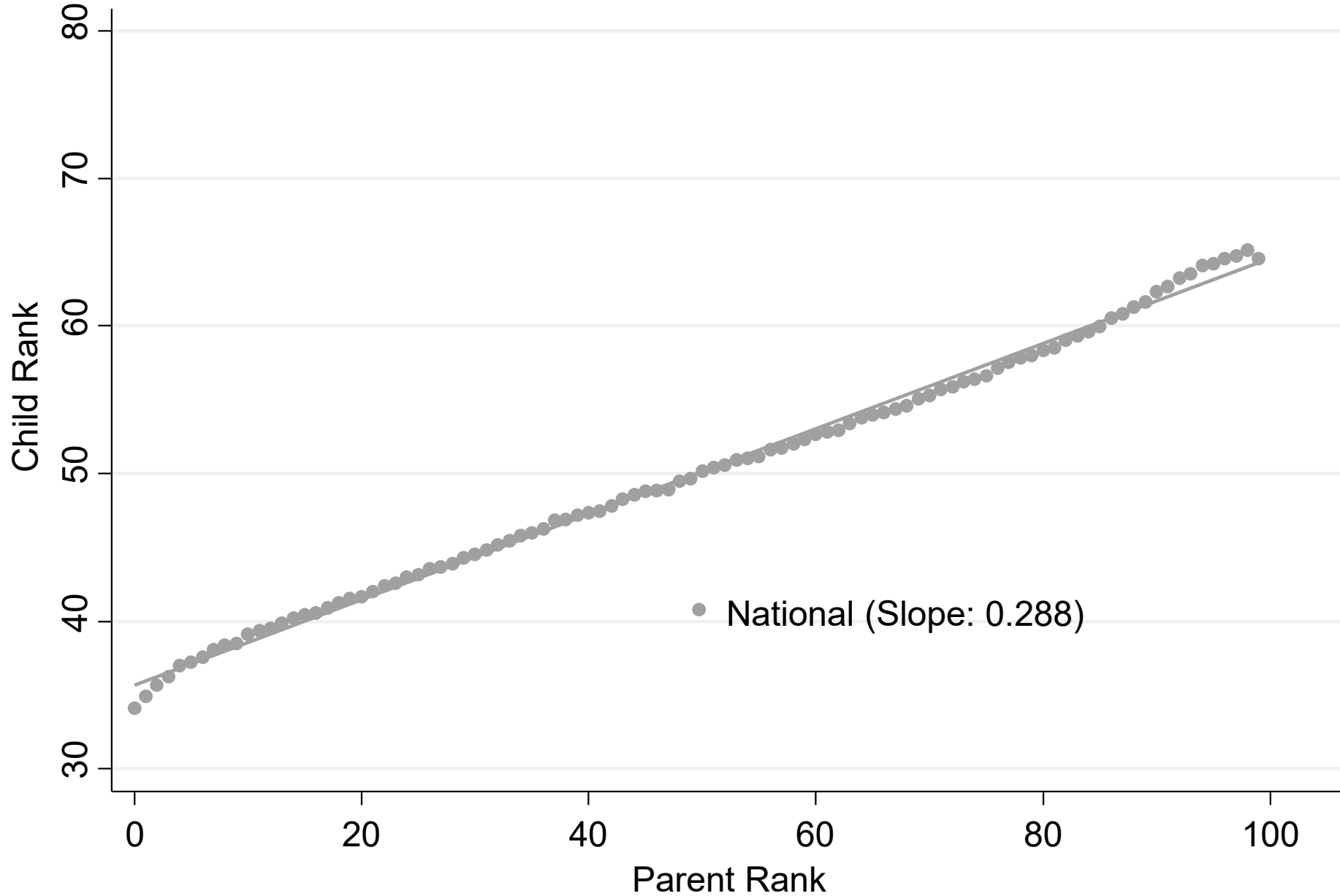


Measuring Student Earnings

- Characterize children's earnings ranks conditional on their parents' rank by college

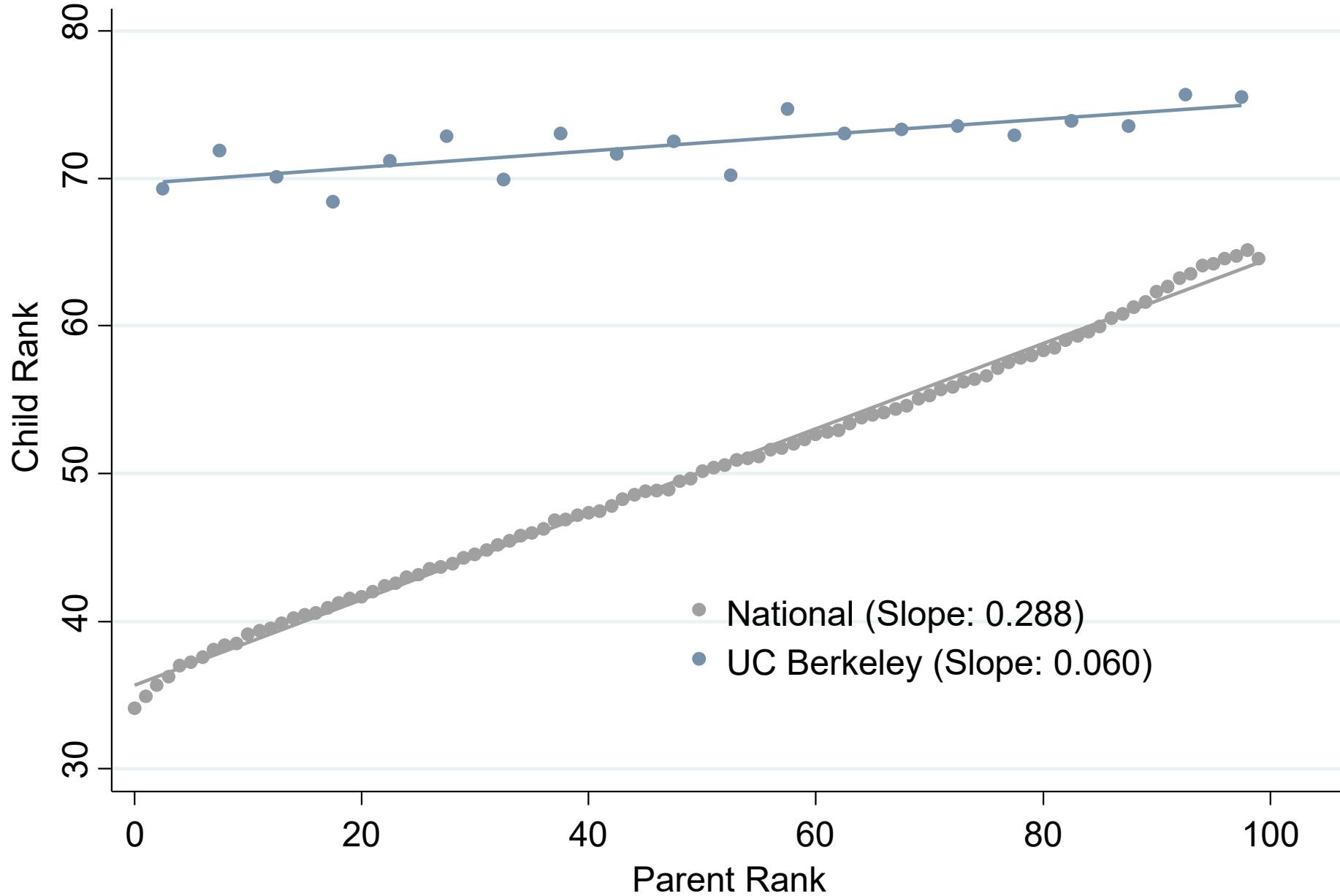
Mean Child Rank at Age 34 vs. Parent Income Rank

Full Population



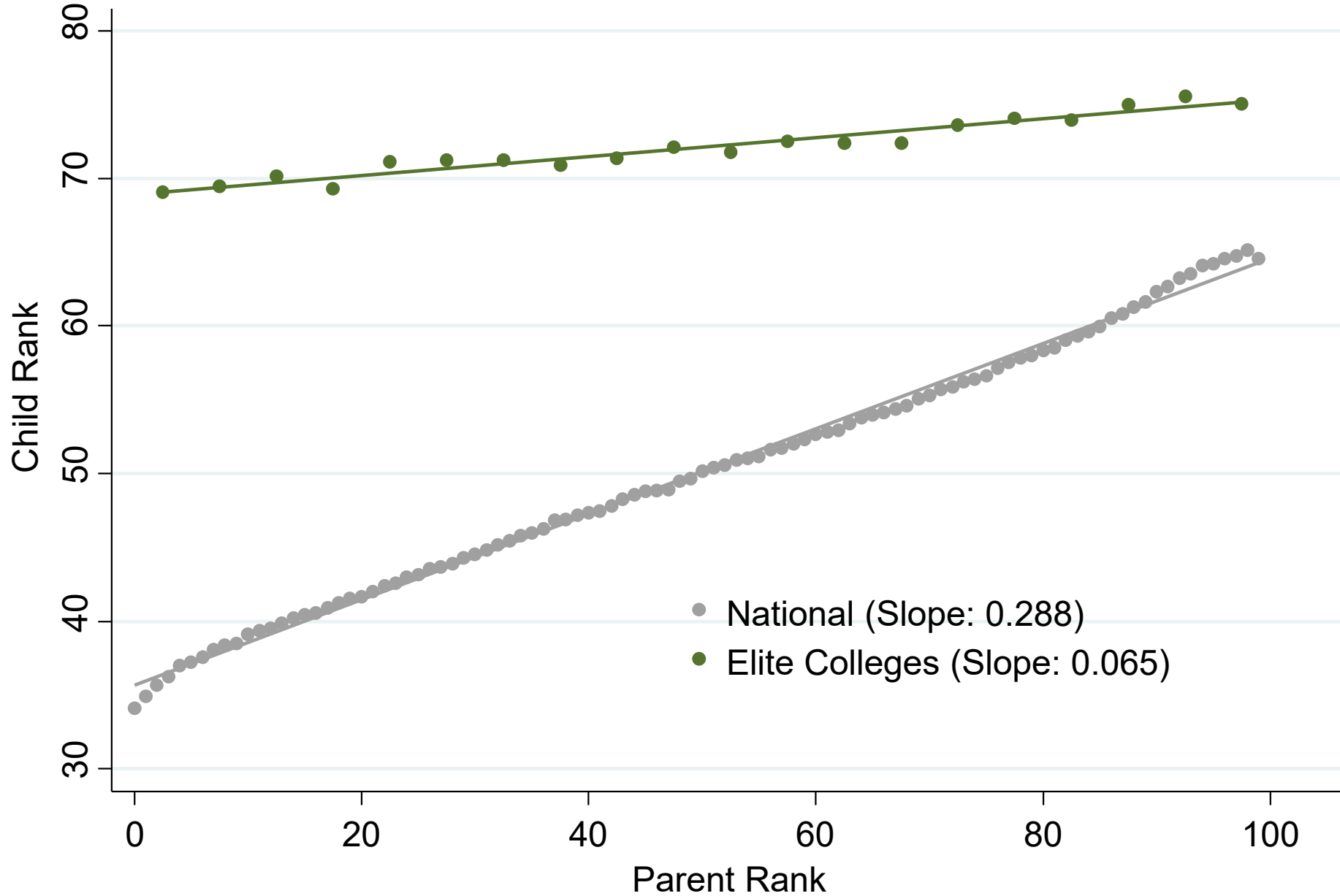
Mean Child Rank at Age 34 vs. Parent Income Rank

UC-Berkeley



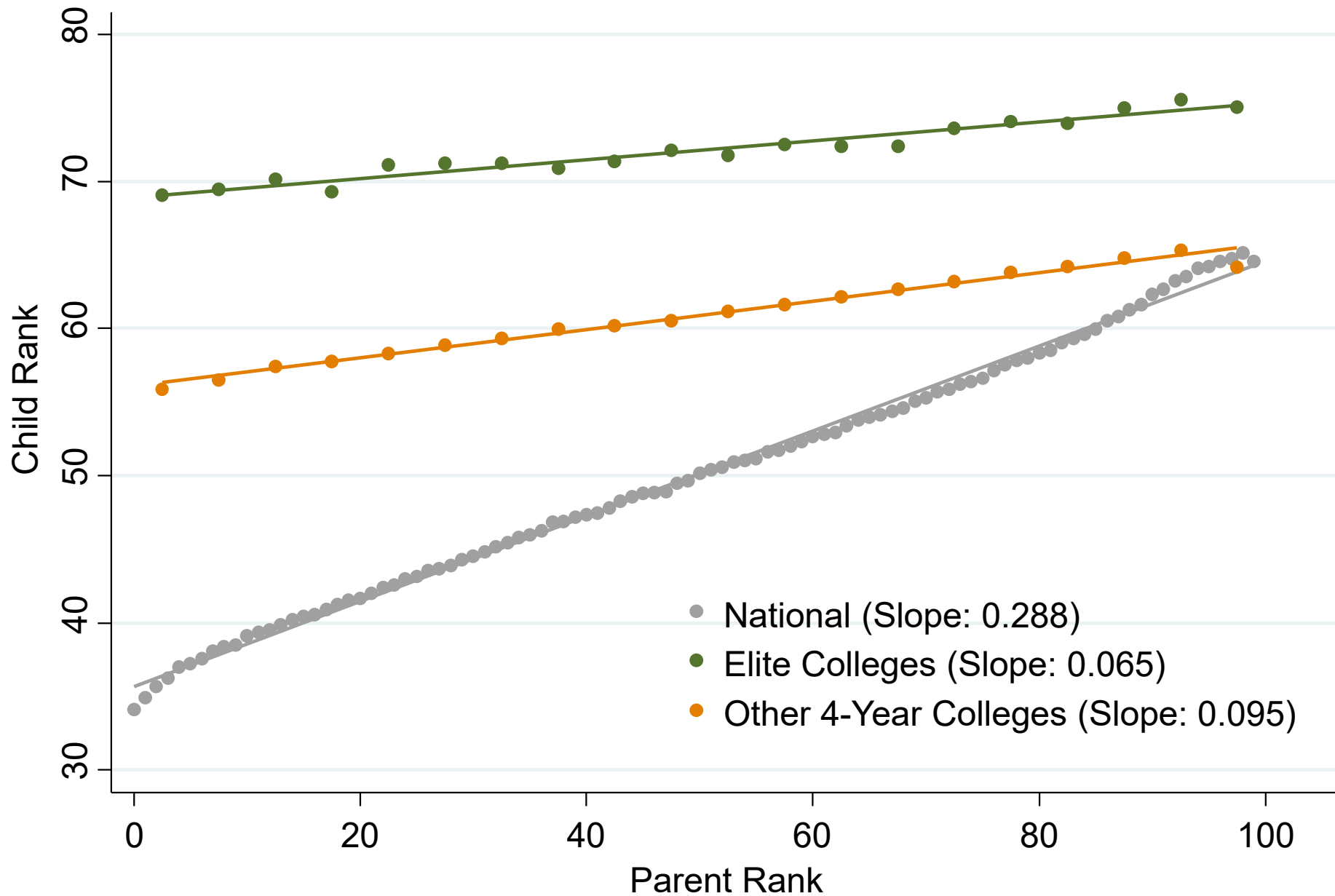
Mean Child Rank at Age 34 vs. Parent Income Rank

Elite Colleges



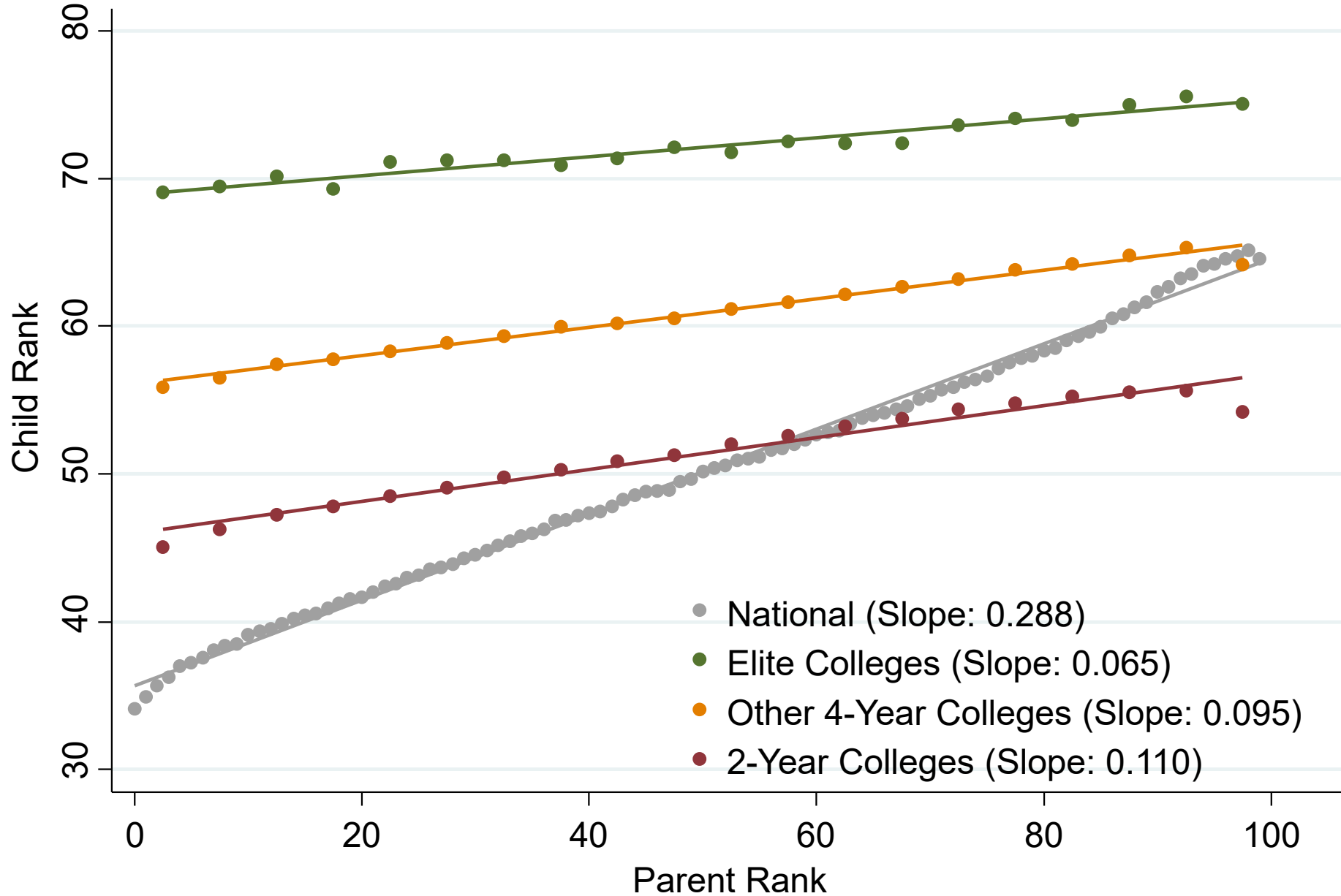
Mean Child Rank at Age 34 vs. Parent Income Rank

All 4-Year Colleges



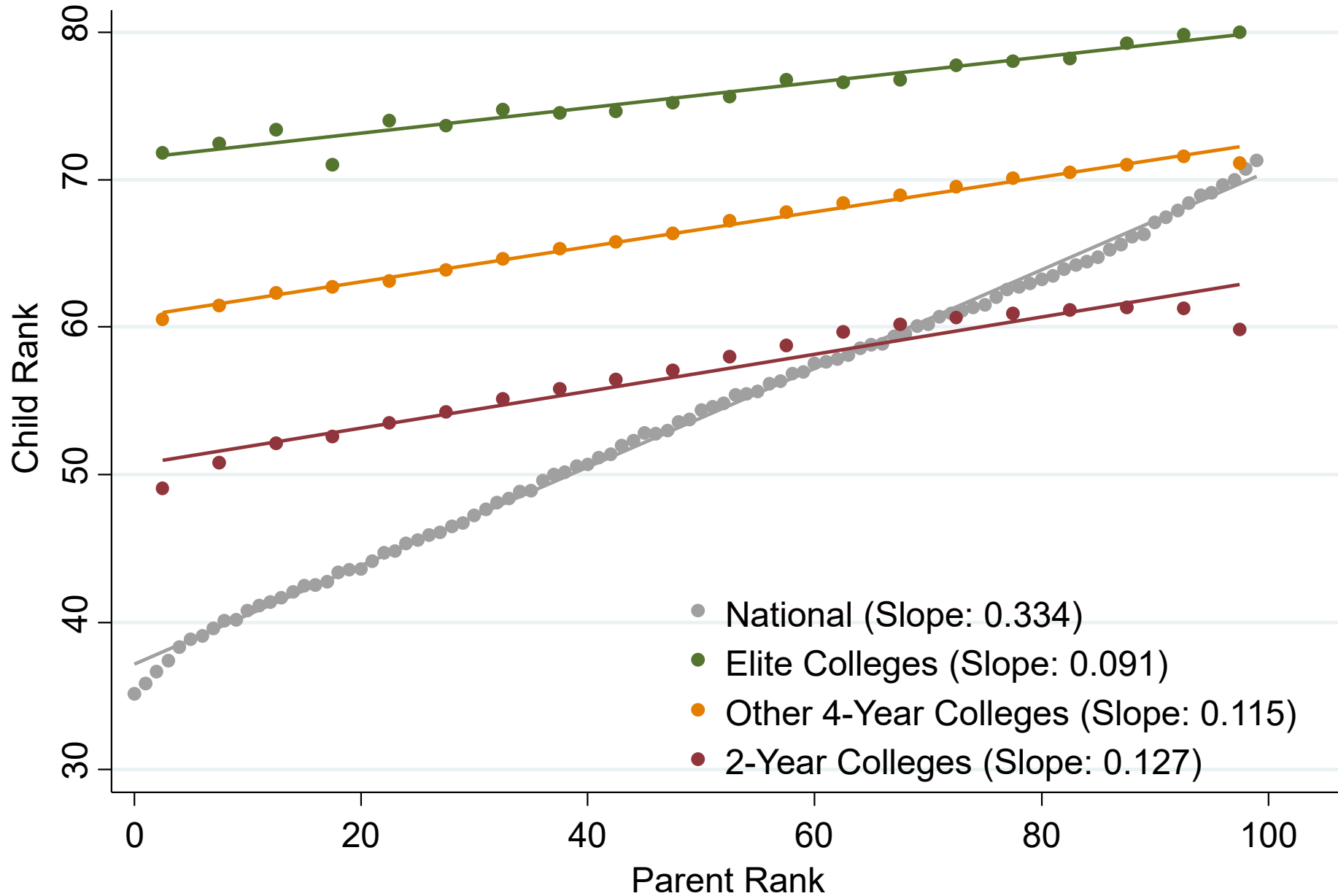
Mean Child Rank at Age 34 vs. Parent Income Rank

All Colleges



Mean Child Rank at Age 34 vs. Parent Income Rank

All Colleges – Male Children Only



Measuring Student Earnings

- **Fact #2:** Students from low- and high- income families at a given college have fairly similar earnings outcomes, especially at highly selective colleges
 - Elite colleges almost “level the playing field” across students with different socioeconomic backgrounds whom they admit

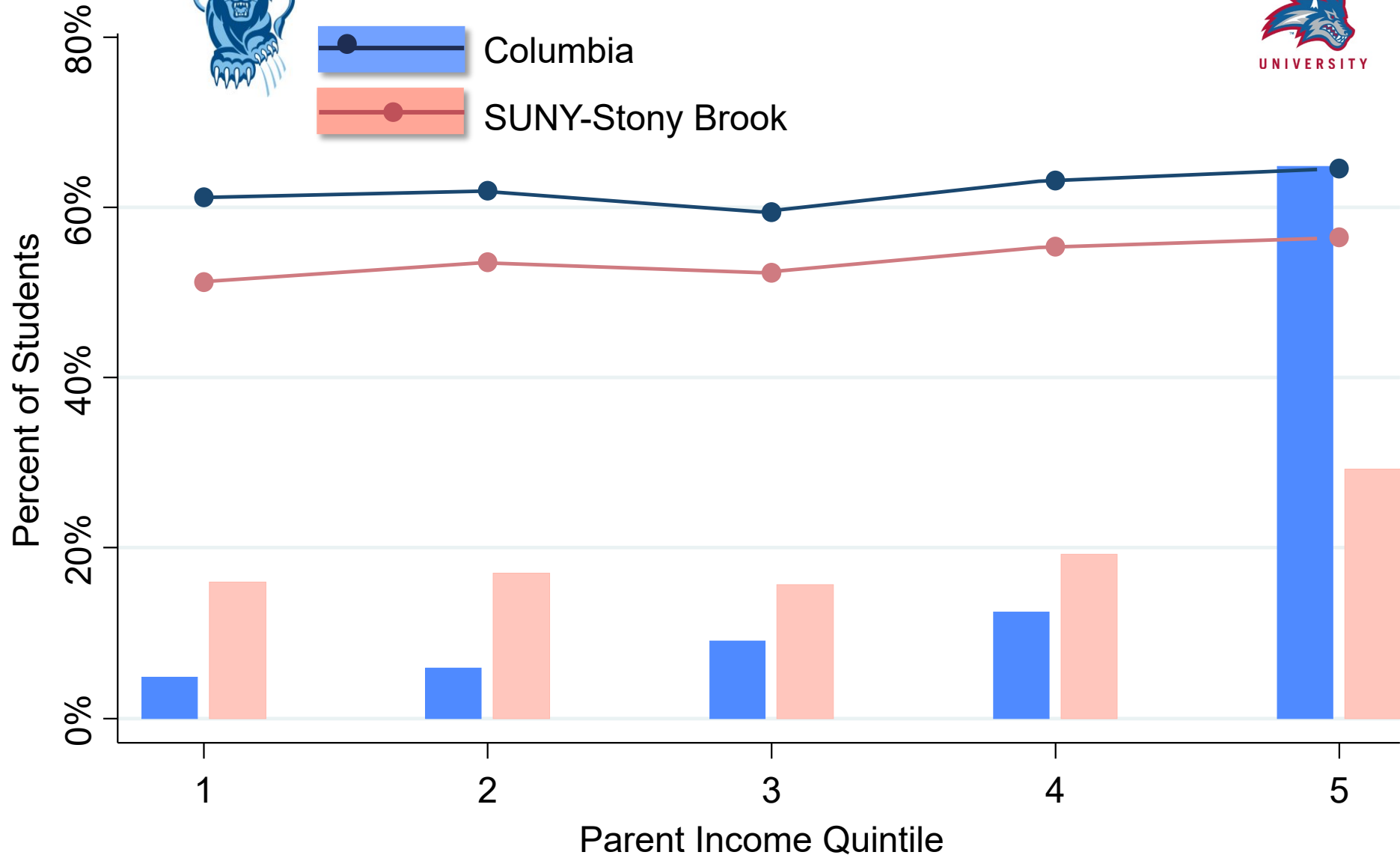
Mobility Report Cards

- Combine data on parents' incomes and students' outcomes to characterize colleges' mobility rates
- Begin by measuring upward mobility as reaching top quintile
 - Turn to upper-tail success (reaching top 1%) later



Mobility Report Cards

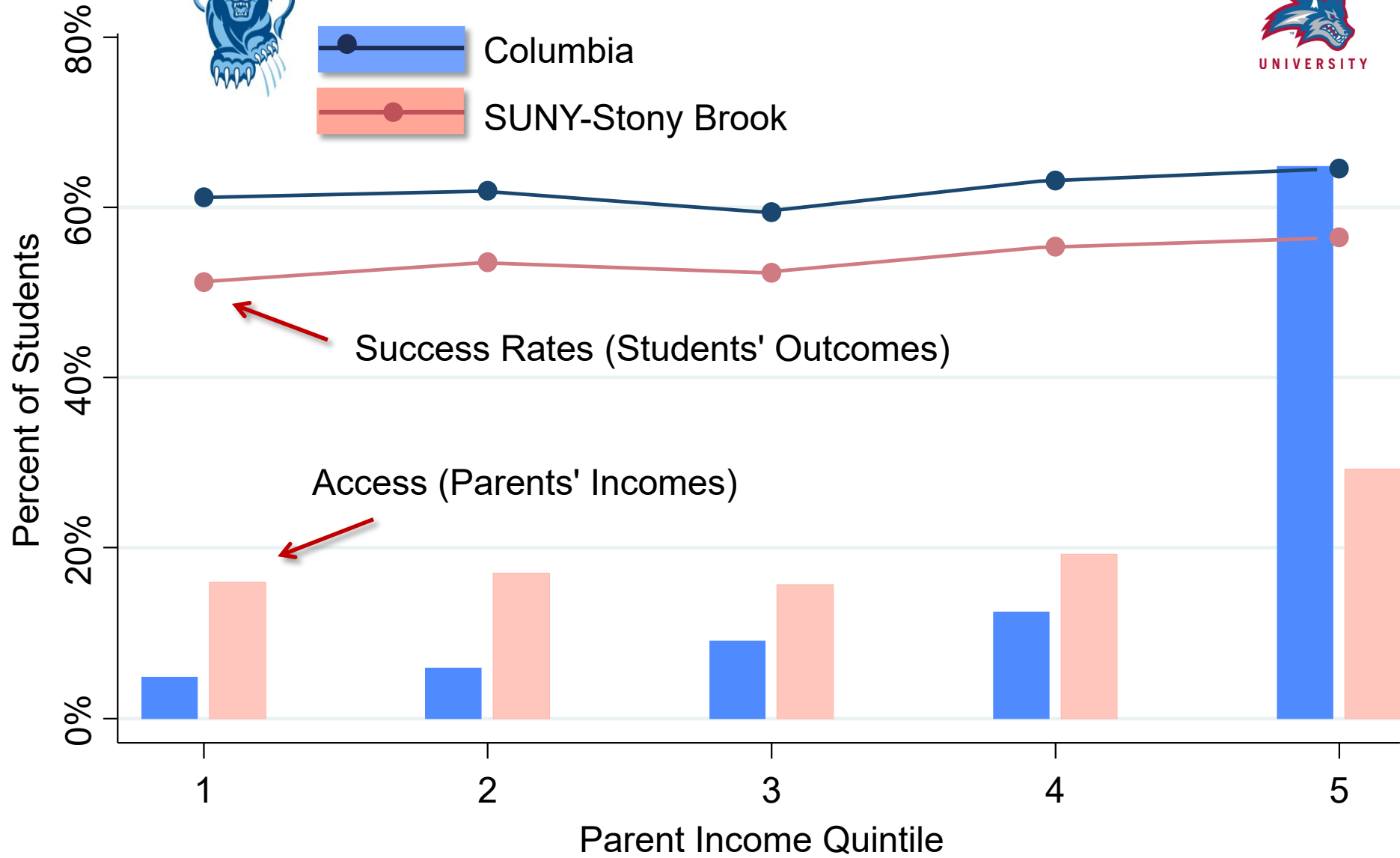
Columbia vs. SUNY-Stony Brook





Mobility Report Cards

Columbia vs. SUNY-Stony Brook



Rates of Mobility

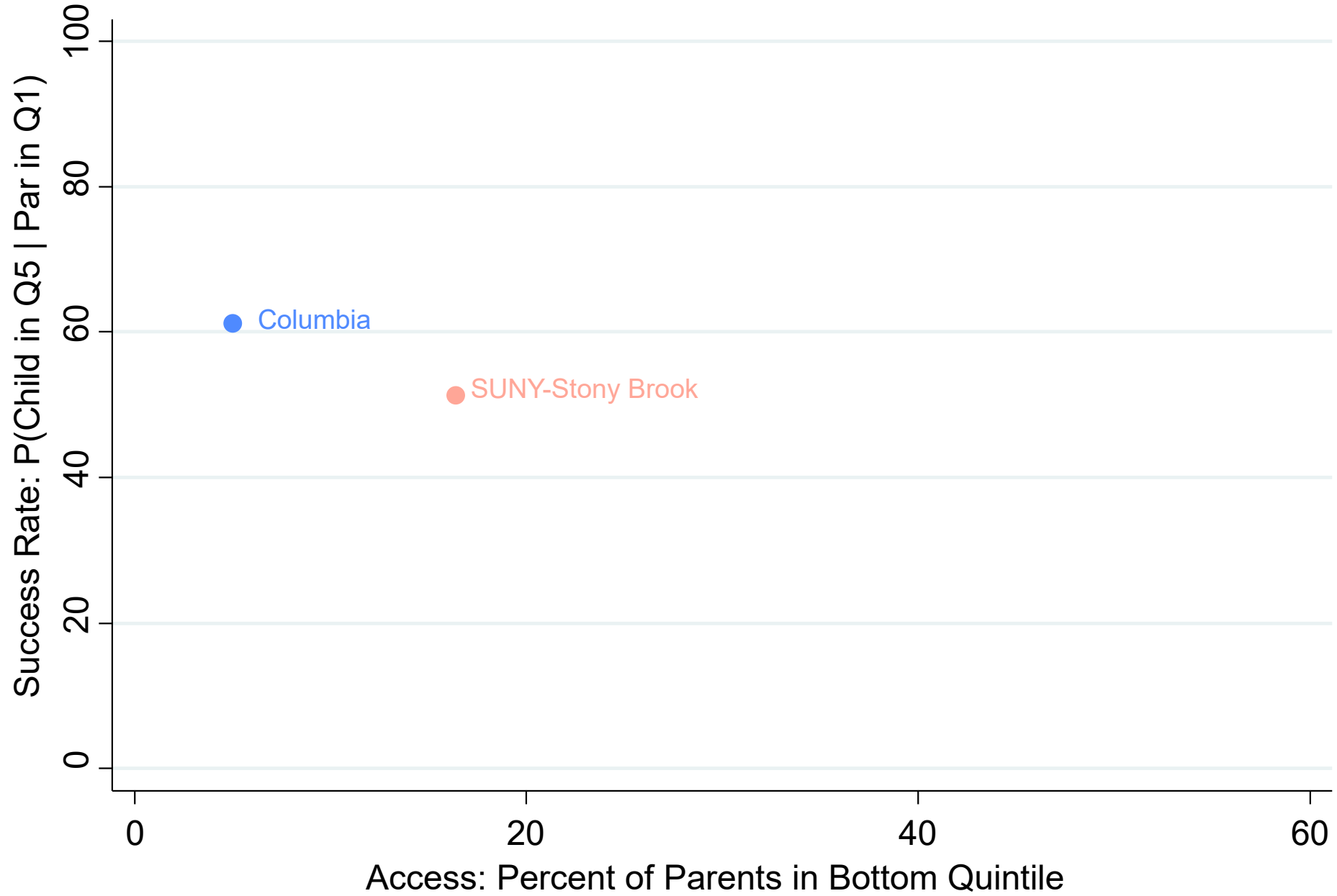
- Define a college's *mobility rate* (MR) as the fraction of its students who come from bottom quintile and end up in top quintile

$$\begin{array}{ccccc} \text{Mobility Rate} & = & \text{Success Rate} & \times & \text{Access} \\ \uparrow & & \uparrow & & \uparrow \\ P(\text{Child in Q5 \& Parent in Q1}) & & P(\text{Child in Q5} | \text{Parent in Q1}) & & P(\text{Parent in Q1}) \end{array}$$

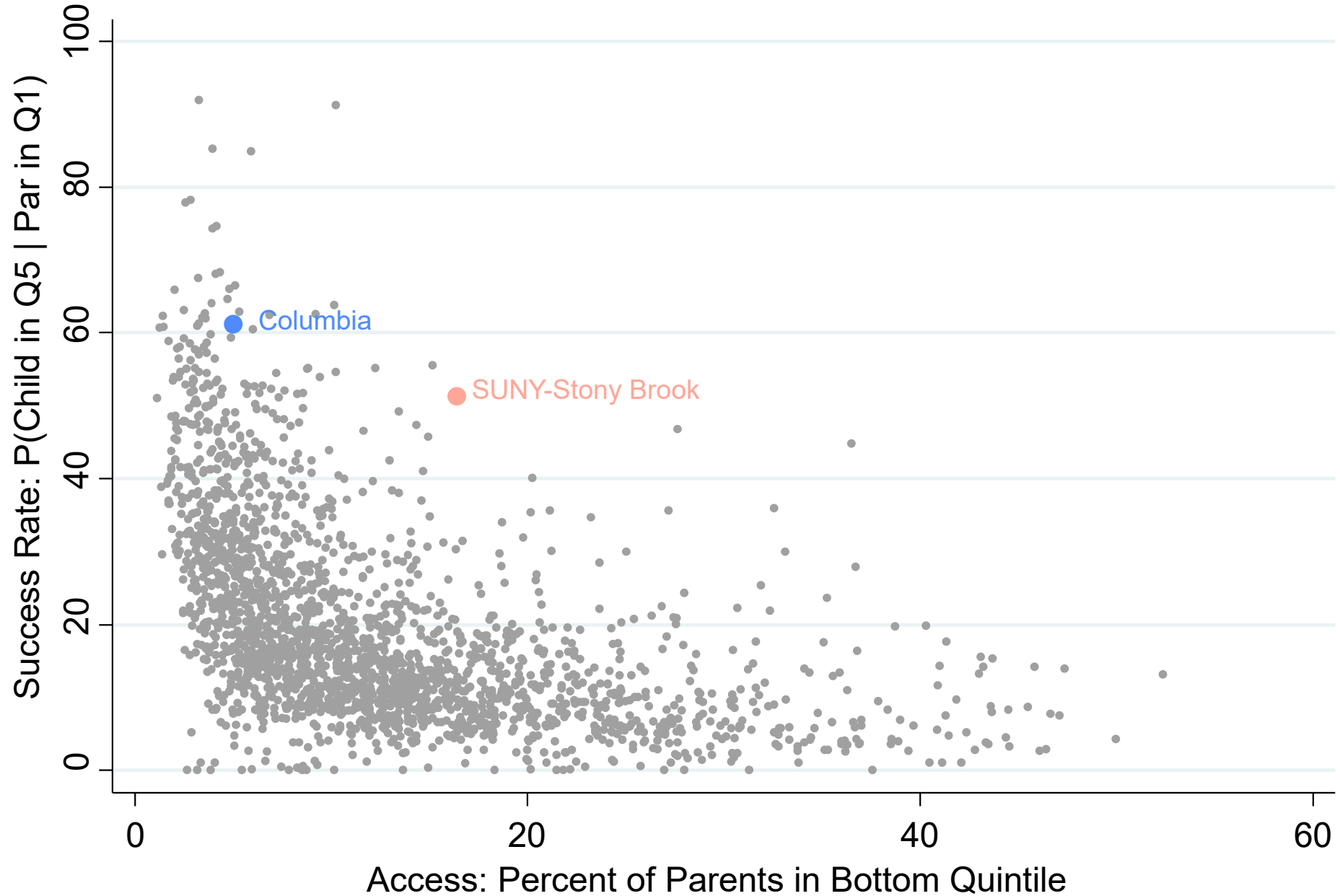
- E.g., SUNY-Stony Brook: 8.4% = 51.2% x 16.4%

- The mobility rate should be interpreted as an accounting measure rather than a causal effect

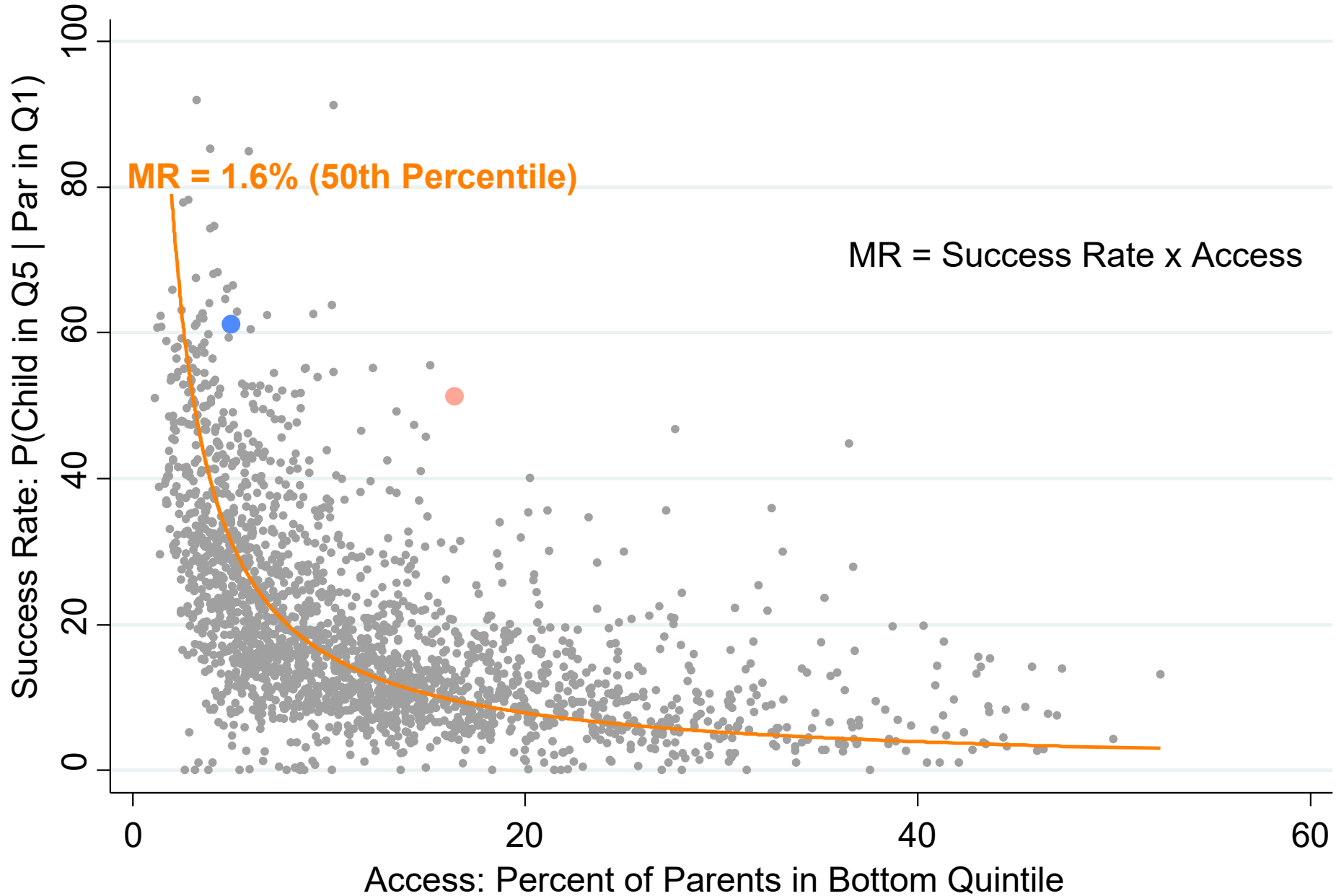
Mobility Rates: Success Rate vs. Access by College



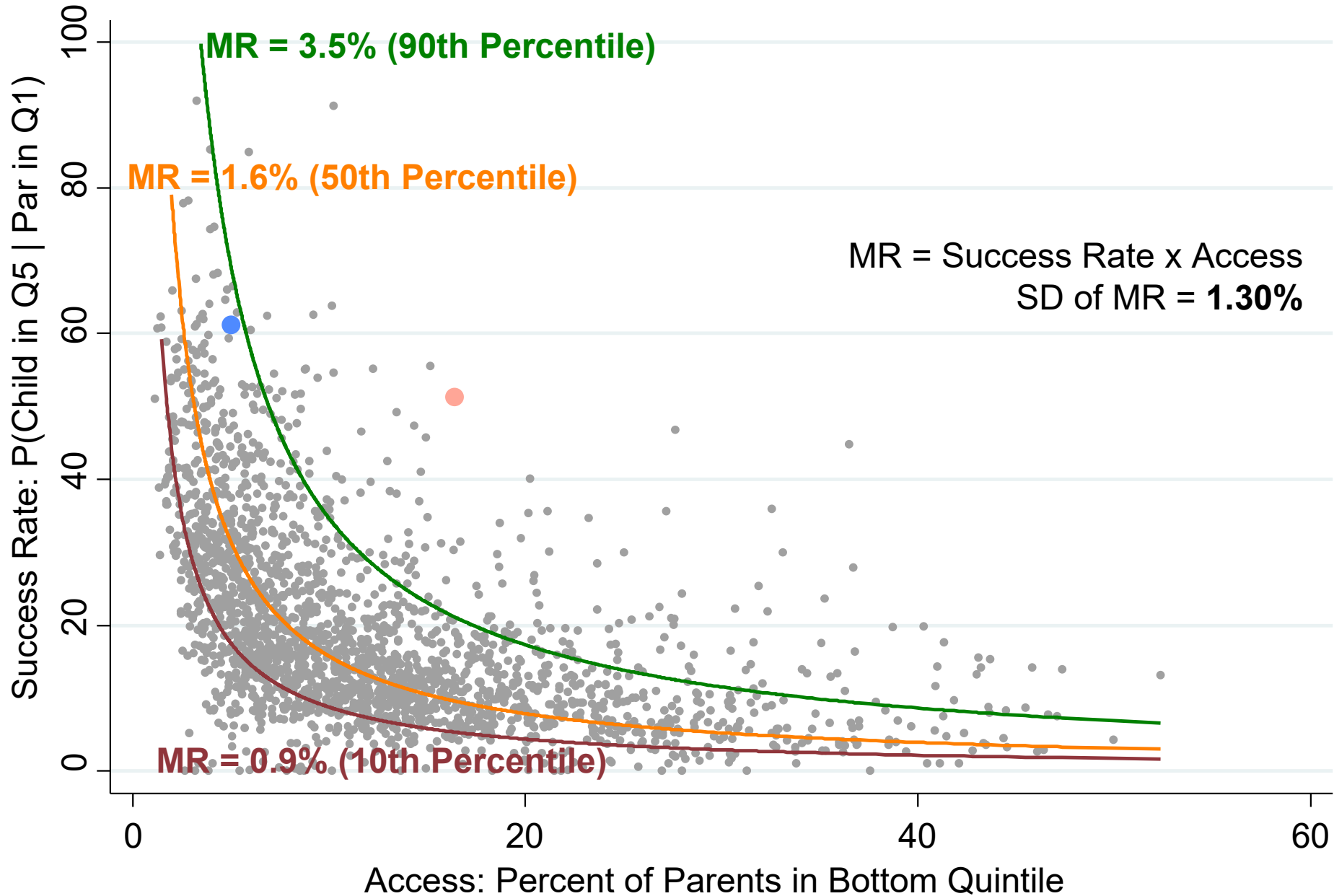
Mobility Rates: Success Rate vs. Access by College



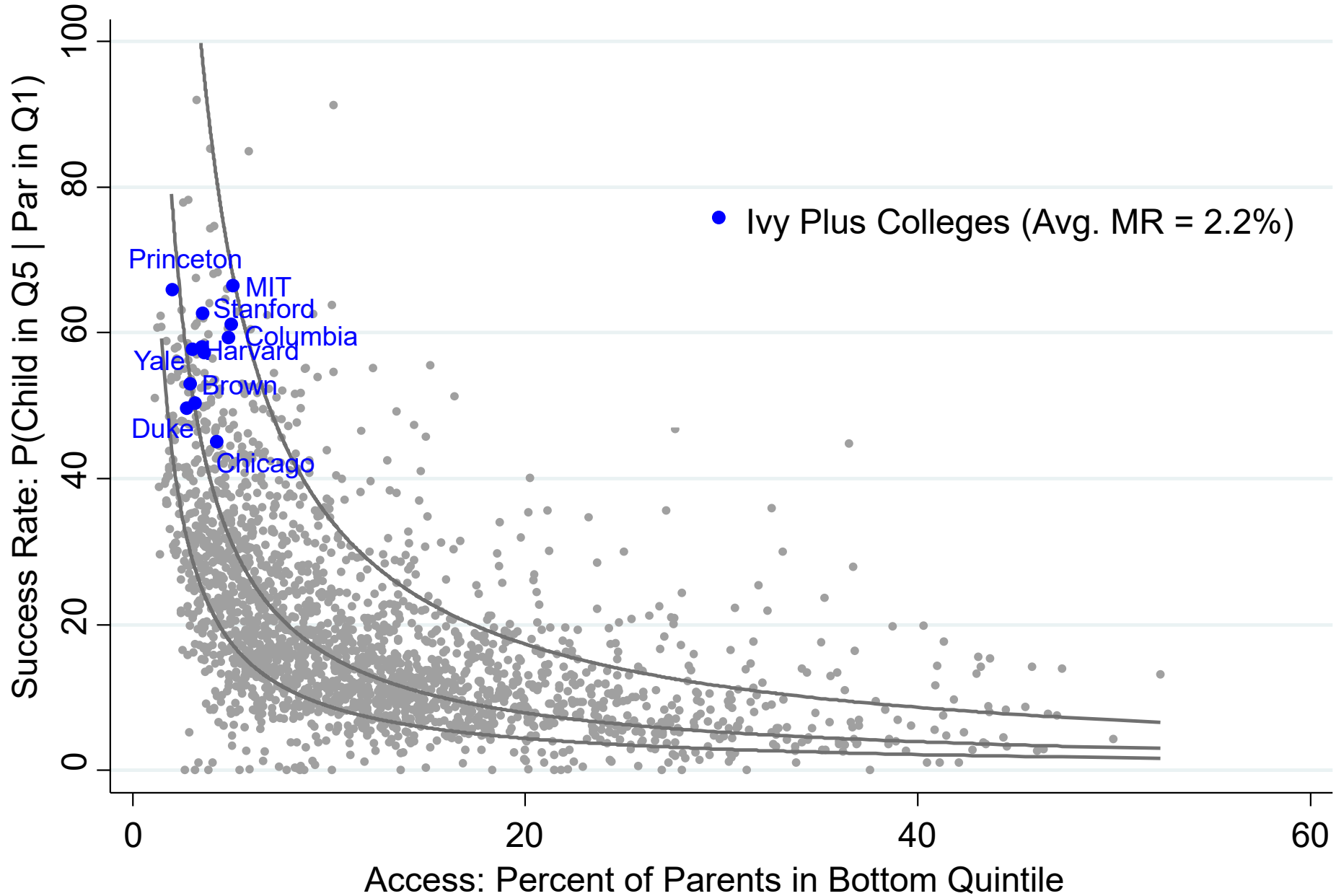
Mobility Rates: Success Rate vs. Access by College



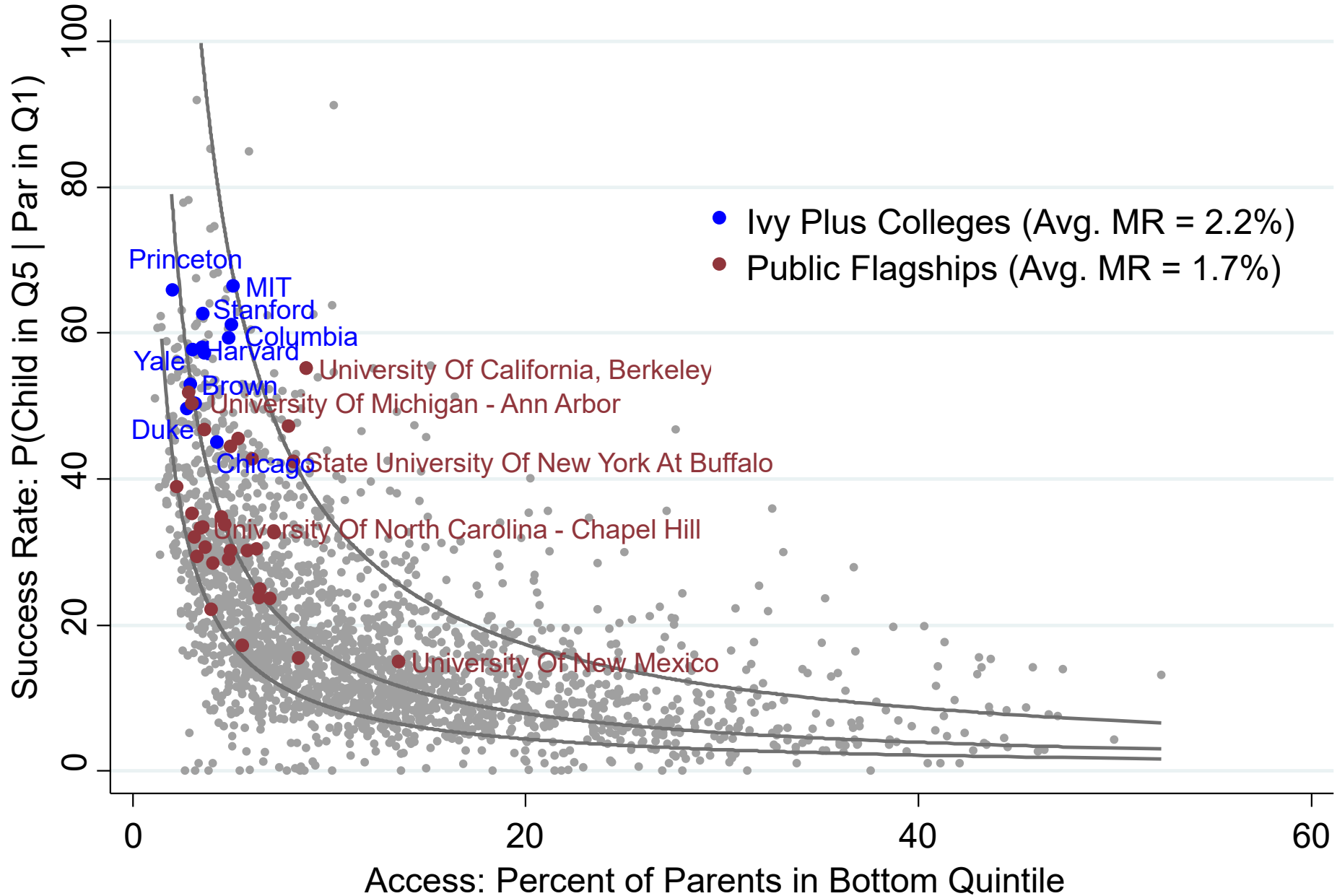
Mobility Rates: Success Rate vs. Access by College



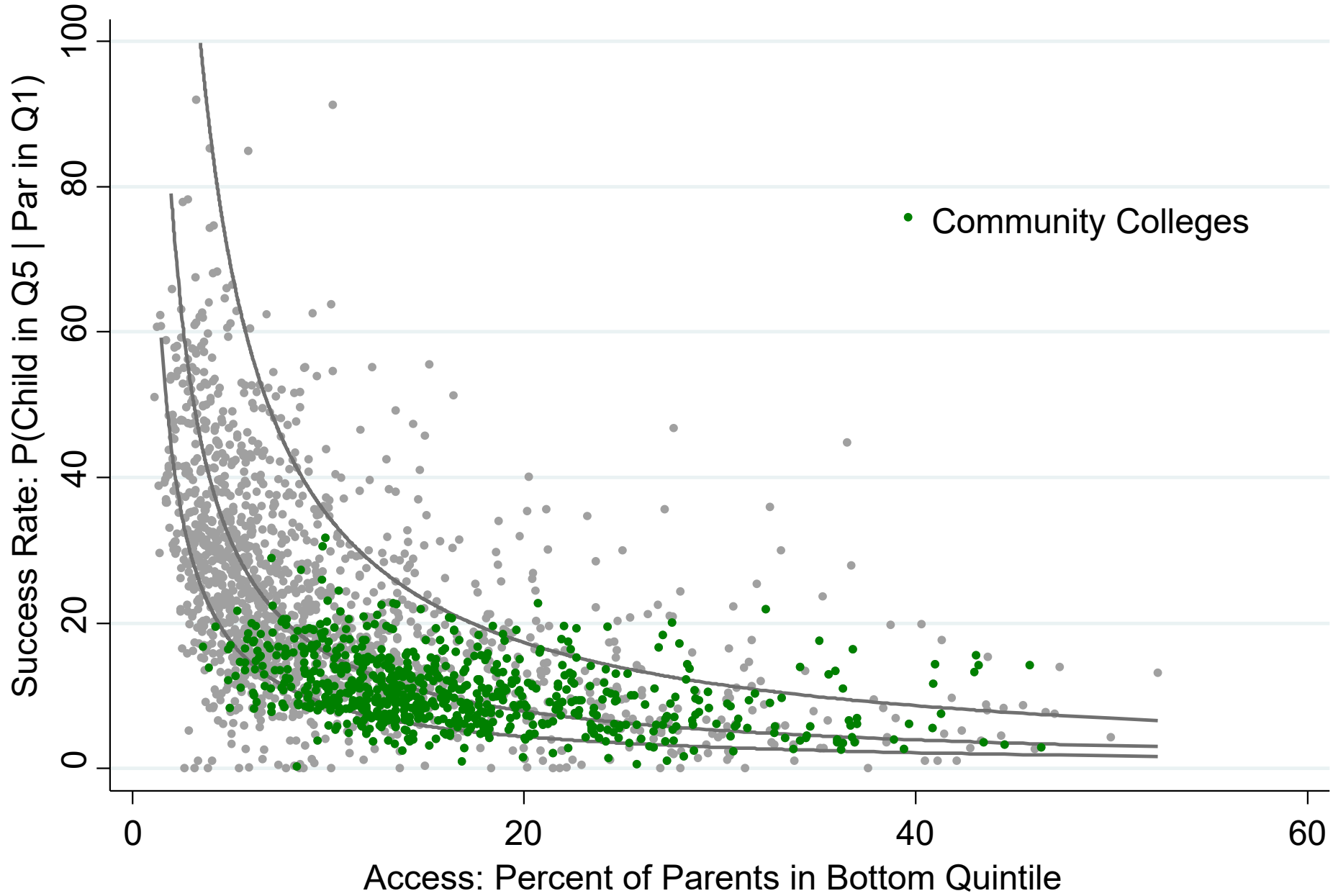
Mobility Rates: Success Rate vs. Access by College



Mobility Rates: Success Rate vs. Access by College



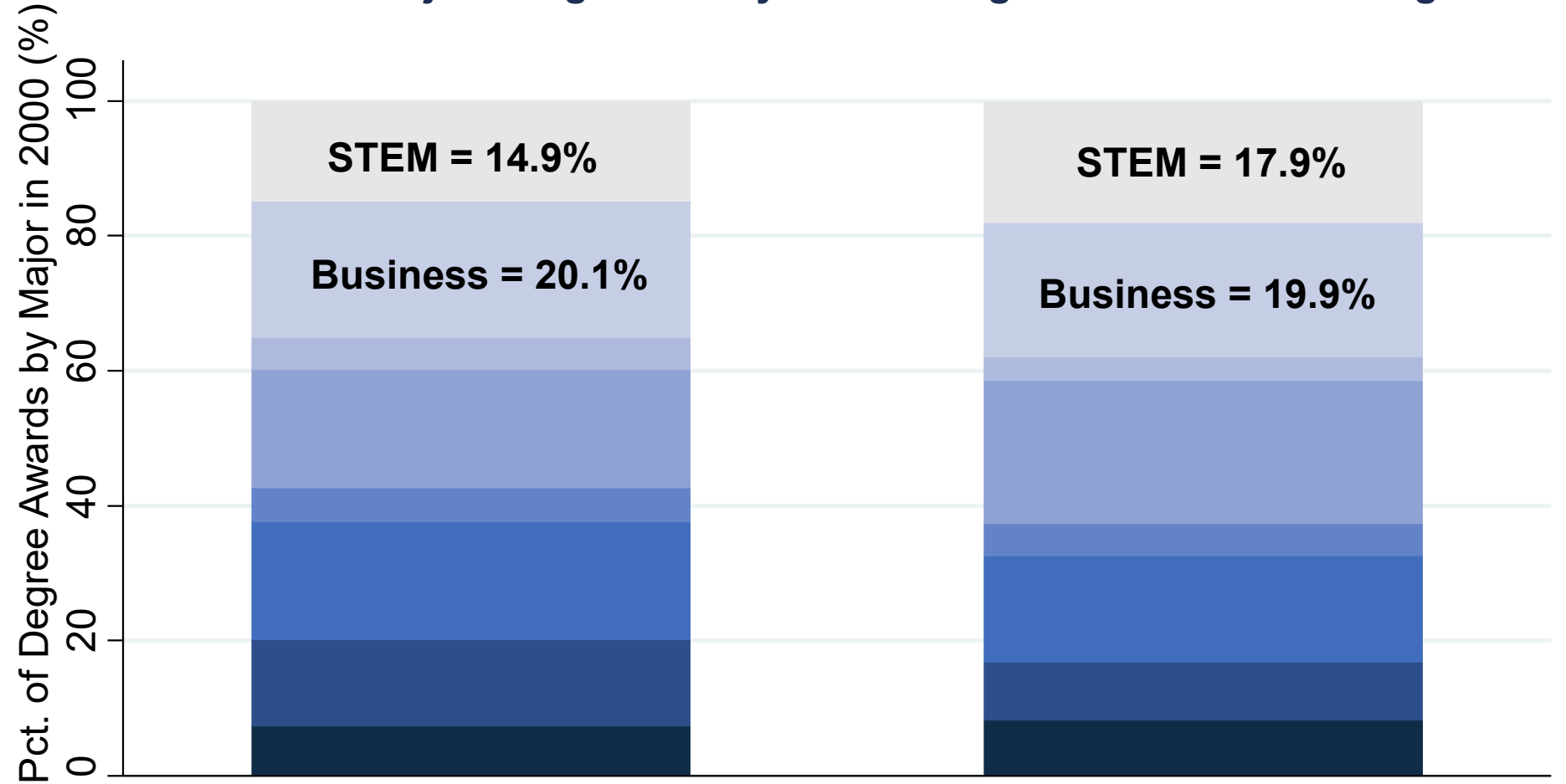
Mobility Rates: Success Rate vs. Access by College



Top 10 Colleges by Mobility Rate (Bottom to Top 20%)

Rank	Name	Fraction Low- Income	x	Top-Quintile Outcome Rate	=	Mobility Rate
1	Cal State University – LA	33.1%		29.9%		9.9%
2	Pace University – New York	15.2%		55.6%		8.4%
3	SUNY – Stony Brook	16.4%		51.2%		8.4%
4	University of Texas – Pan American	38.7%		19.8%		7.6%
5	CUNY System	28.7%		25.2%		7.2%
6	Glendale Community College	32.4%		21.9%		7.1%
7	South Texas College	52.4%		13.2%		6.9%
8	Cal State Polytechnic – Pomona	14.9%		45.8%		6.8%
9	University of Texas – El Paso	28.0%		24.4%		6.8%
10	St John’s University – Queens, NY	14.3%		47.4%		6.8%

Distribution of Majors: High-Mobility-Rate Colleges vs. All Other Colleges

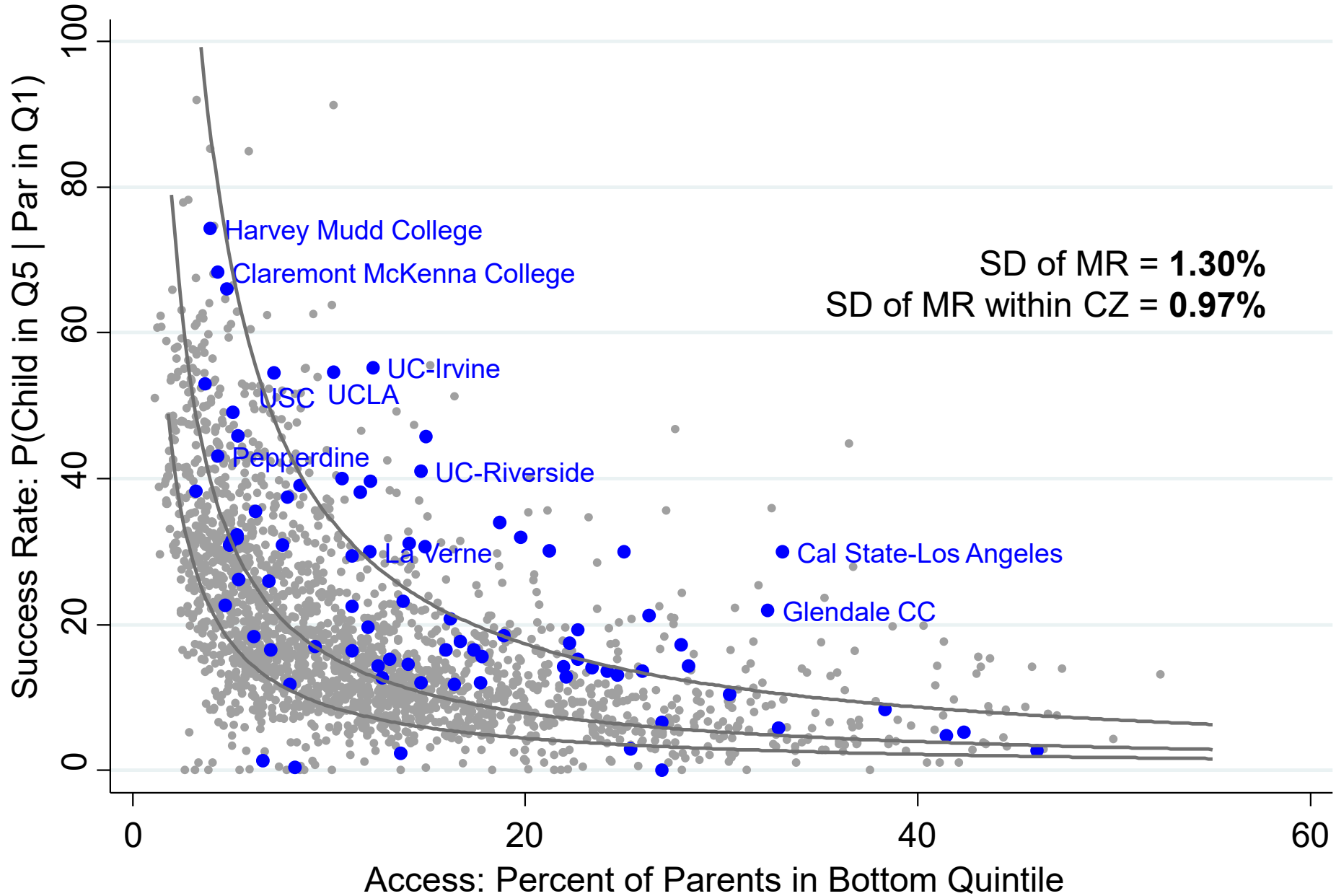


All Other Colleges

High Mobility Rate Colleges

- STEM
- Business
- Trades and Personal Services
- Social Sciences
- Public and Social Services
- Multi/Interdisciplinary Studies
- Health and Medicine
- Arts and Humanities

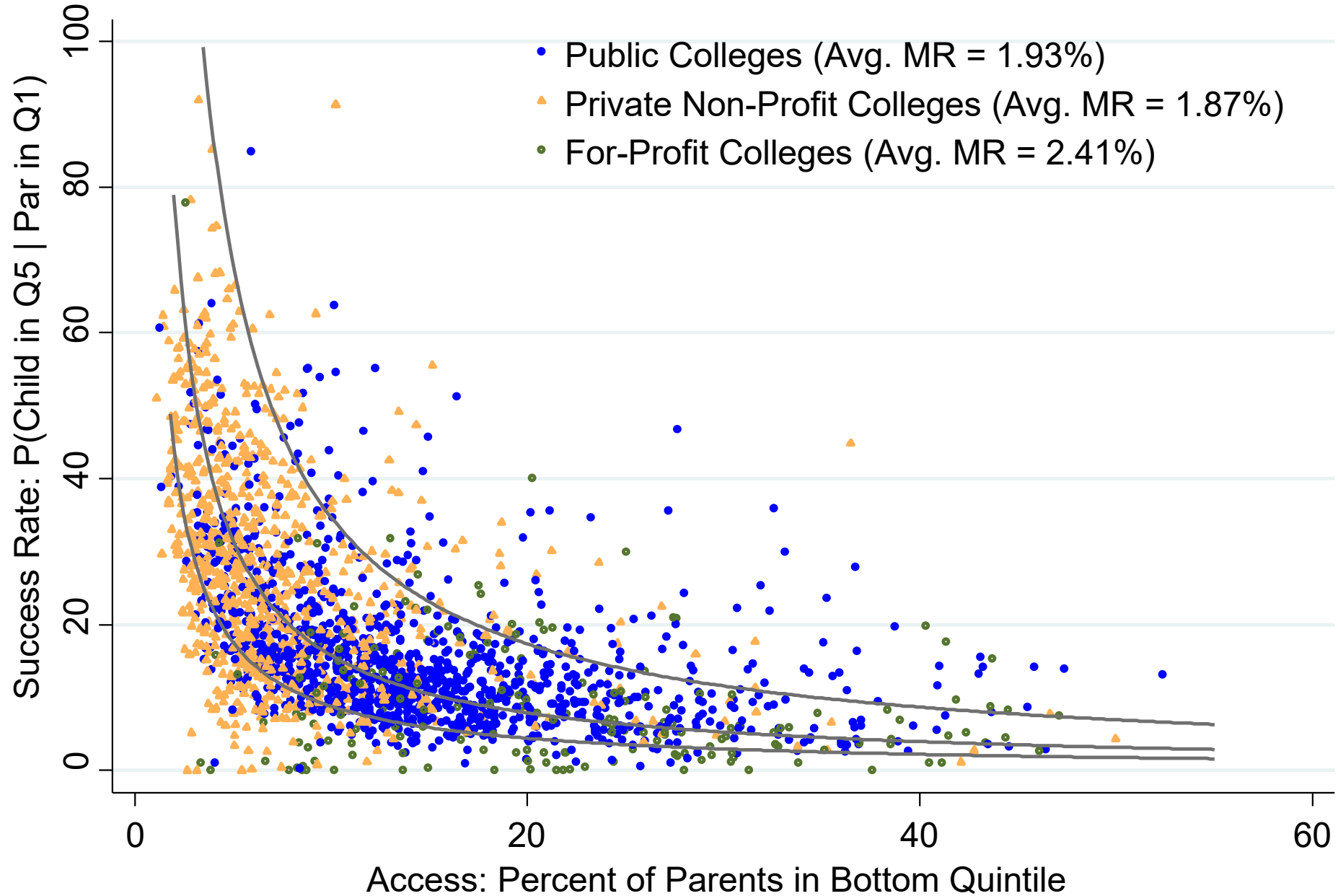
Mobility Rates: Colleges in Los Angeles



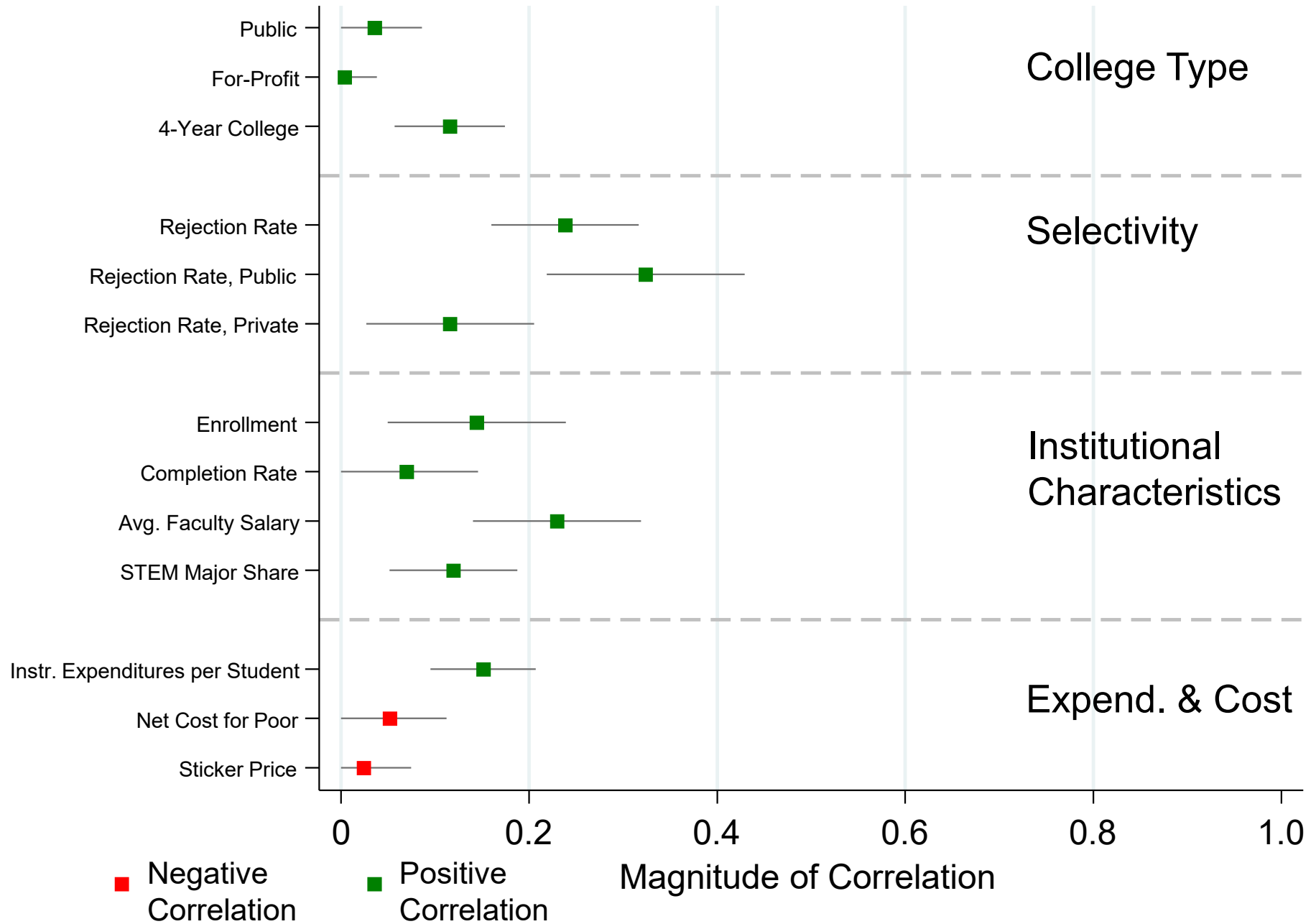
Which Colleges Have the Highest Mobility Rates?

- Characterize the types of colleges with high vs. low rates of mobility
 - Correlate Mobility Rate, $P(\text{Child in Q5 and Parent in Q1})$, with various college characteristics
- Analysis in this section is purely descriptive; explore causal effects of colleges toward the end of the talk

Mobility Rates: Success Rate vs. Access by College



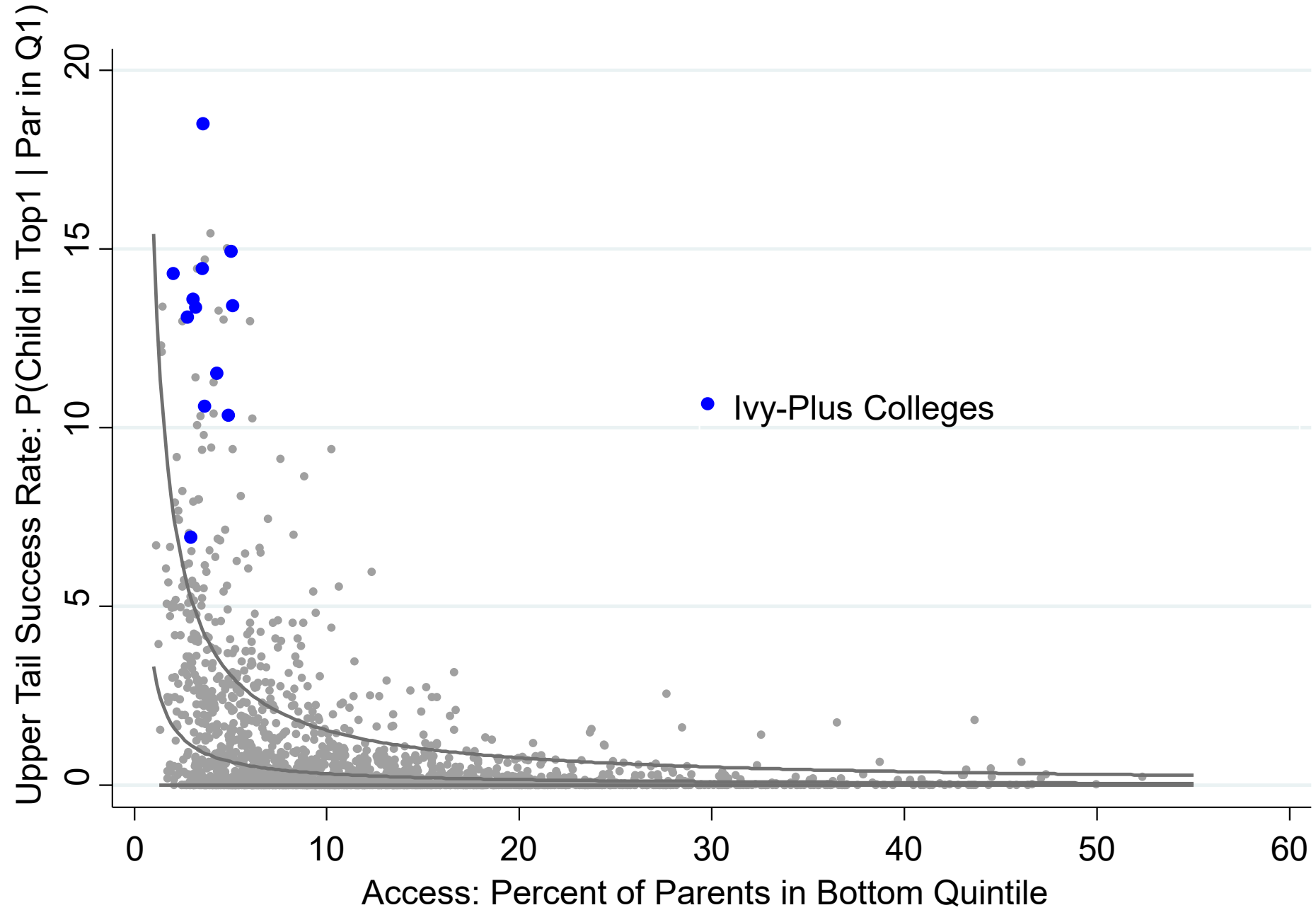
Correlates of Top 20% Mobility Rate



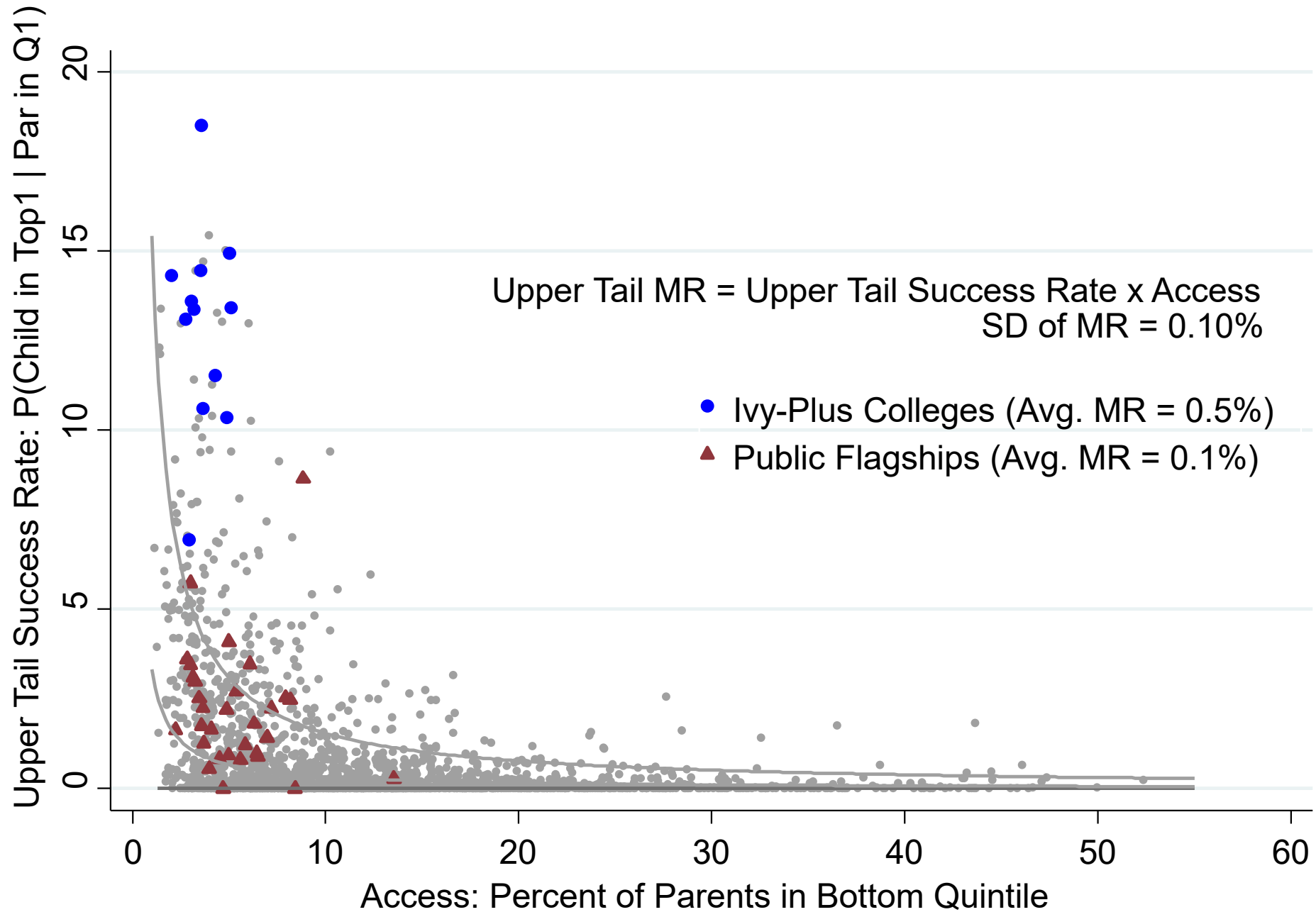
Mobility Rates for Upper-Tail Success

- Now examine mobility rates for upper tail success: fraction of students who come from bottom quintile and reach top 1%

Upper-Tail Success Rate (Top 1%) vs. Access by College



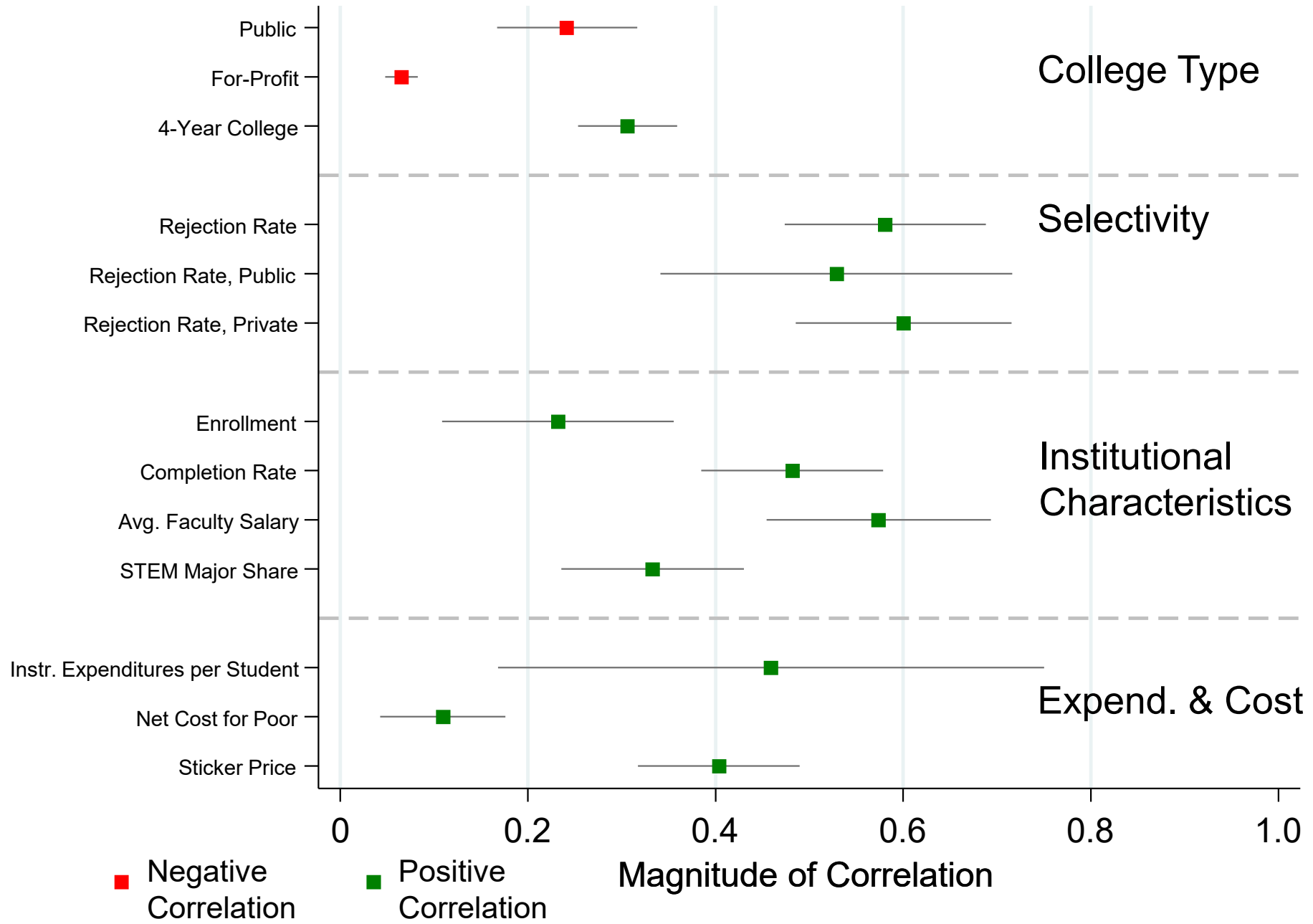
Upper-Tail Success Rate (Top 1%) vs. Access by College



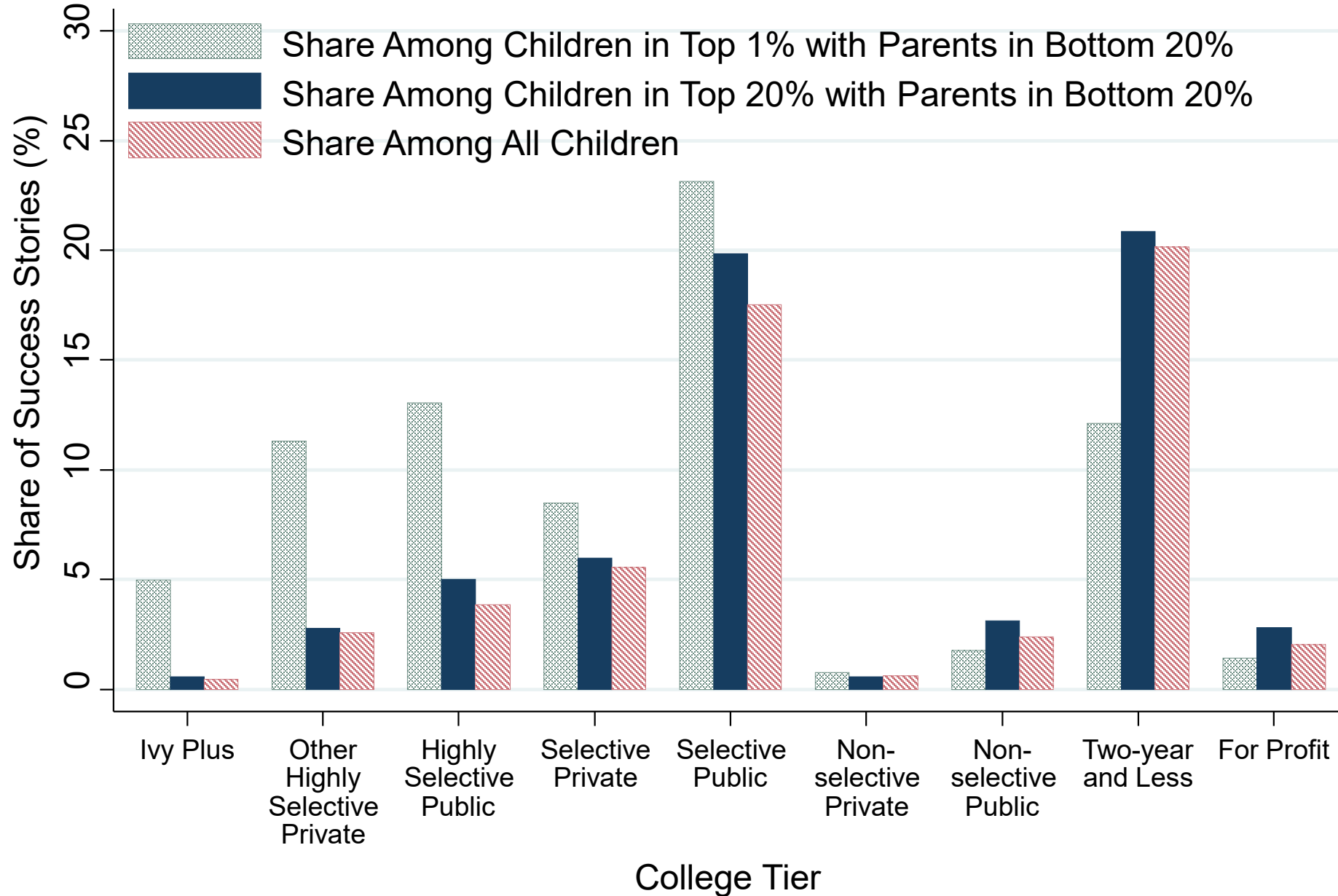
Top 10 Colleges by Mobility Rates for Upper-Tail Success (Top 1%)

Rank	Name	Mobility Rate	=	Access	x	Upper-Tail Success
1	University of California – Berkeley	0.76%		8.8%		8.6%
2	Columbia University	0.75%		5.0%		14.9%
3	MIT	0.68%		5.1%		13.4%
3	Stanford University	0.66%		3.6%		18.5%
4	Swarthmore College	0.61%		4.7%		13.0%
6	Johns Hopkins University	0.54%		3.7%		14.7%
7	New York University	0.52%		6.9%		7.5%
8	University of Pennsylvania	0.51%		3.5%		14.5%
9	Cornell University	0.51%		4.9%		10.4%
10	University of Chicago	0.50%		4.3%		11.5%

Correlates of Top 1% Mobility Rate



Fraction of Success Stories by School Type



Lessons on Mobility Rates

- **Fact #3:** Certain mid-tier public institutions (e.g., CUNY, Cal-State) have the highest bottom-to-top quintile mobility rates
- But highly selective institutions (e.g., Berkeley) have the highest bottom-to-top-1% mobility rates

Part 3

Counterfactual Student Allocations: Undermatching and Intergenerational Mobility

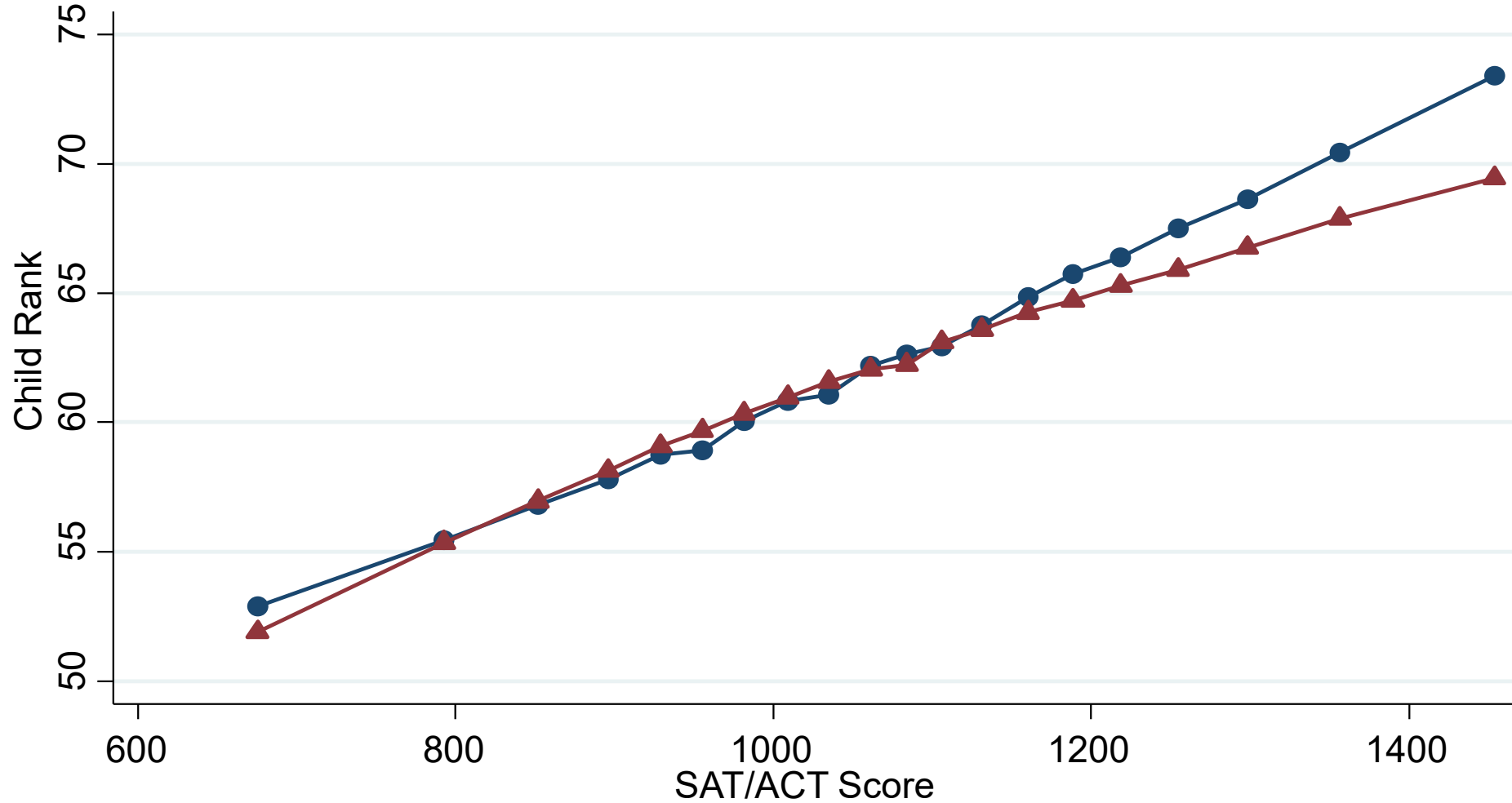
Income-Neutral Student Allocation

- Now simulate how income segregation across colleges and intergenerational mobility would change if students were allocated to colleges differently
- Begin by evaluating the extent to which differences in parental income distributions across colleges can be explained by differences in academic preparation at point of application

Test Score Data

- Use data on SAT/ACT scores for 1980-82 cohorts as a proxy for pre-college academic qualifications
 - Impute scores for college students who do not register a test by matching within parent income quintile, state, selectivity tier, and earnings level in adulthood
 - Validate using out-of-sample testing
- Test scores are a useful proxy for academic qualification as relevant for college enrollment
 - 100 SAT points is associated with a \$5,307 (2.2 percentiles) annual increase in earnings, controlling for race, gender, parent income, and HS fixed effects

Relationship Between SAT Scores and Earnings in Adulthood



—●— No Controls

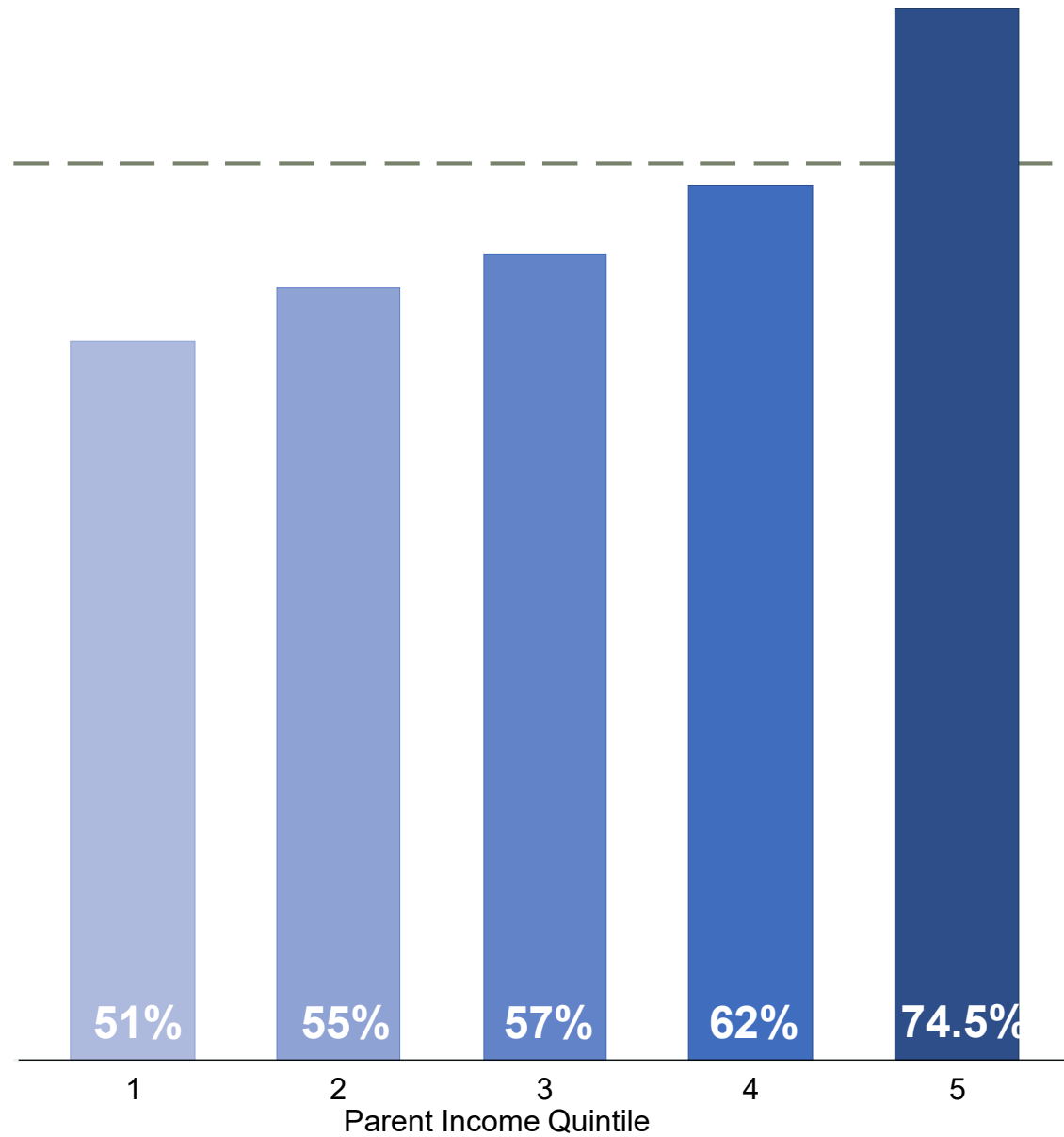
—▲— Controlling for race x sex x cubic in parent rank and HS FE's

Undermatching

- First step: observe that at any given level of SAT/ACT scores, children from higher-income families attend more selective colleges

Attendance Rates at Selective Colleges

Among All Students with Median Selective College SAT (1080)



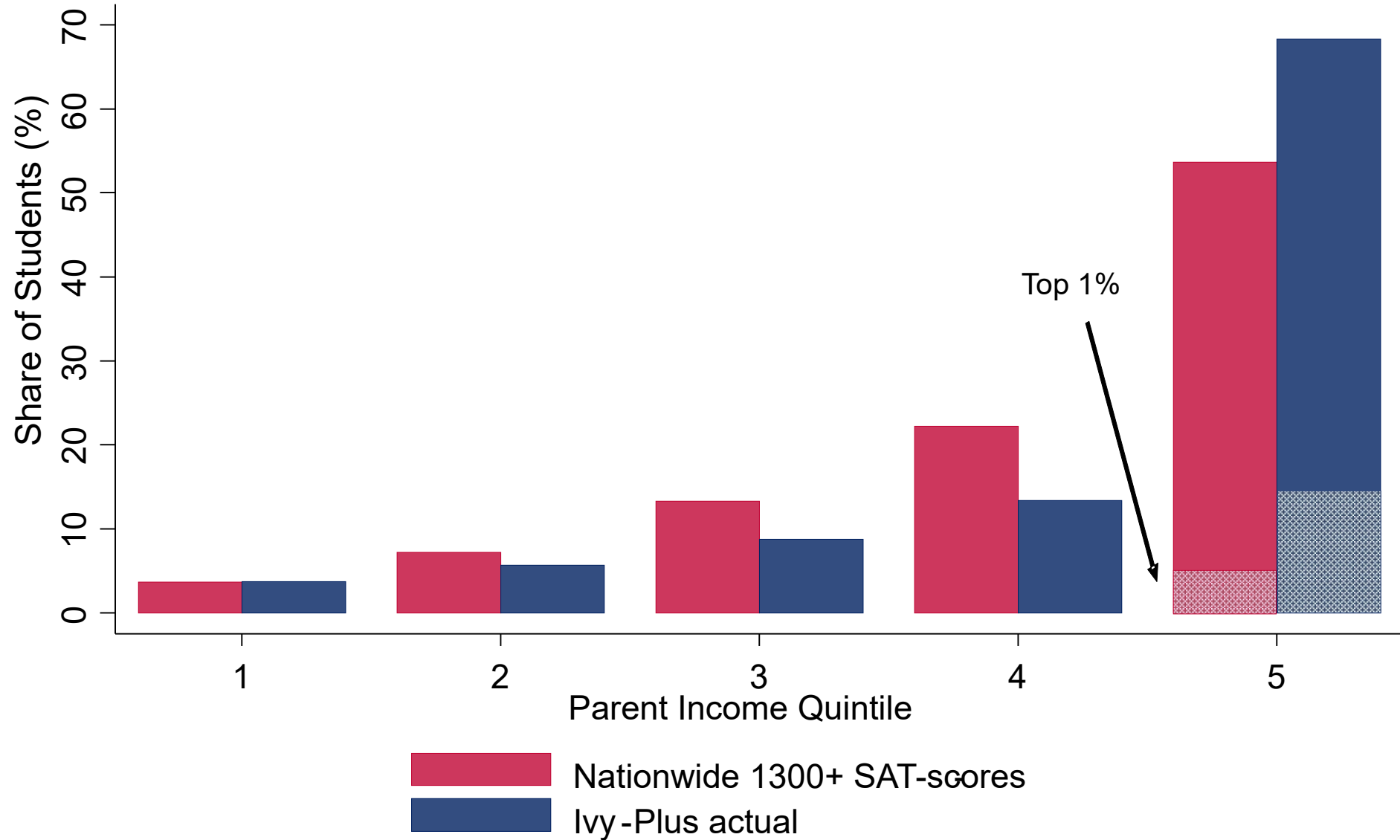
Income-Neutral Student Allocation

- To quantify degree of undermatching, consider reallocating students to equalize attendance rates between students from all parent income backgrounds, conditional on scores
 - Hold number of students, distribution of scores, home state mix, and racial diversity constant at each school
- Such a reallocation would substantially reduce parental income segregation across colleges
 - Closes 38% of gap in bottom-quintile share relative to benchmark in which all colleges have the same bottom-quintile share

Undermatching at Ivy-Plus Colleges

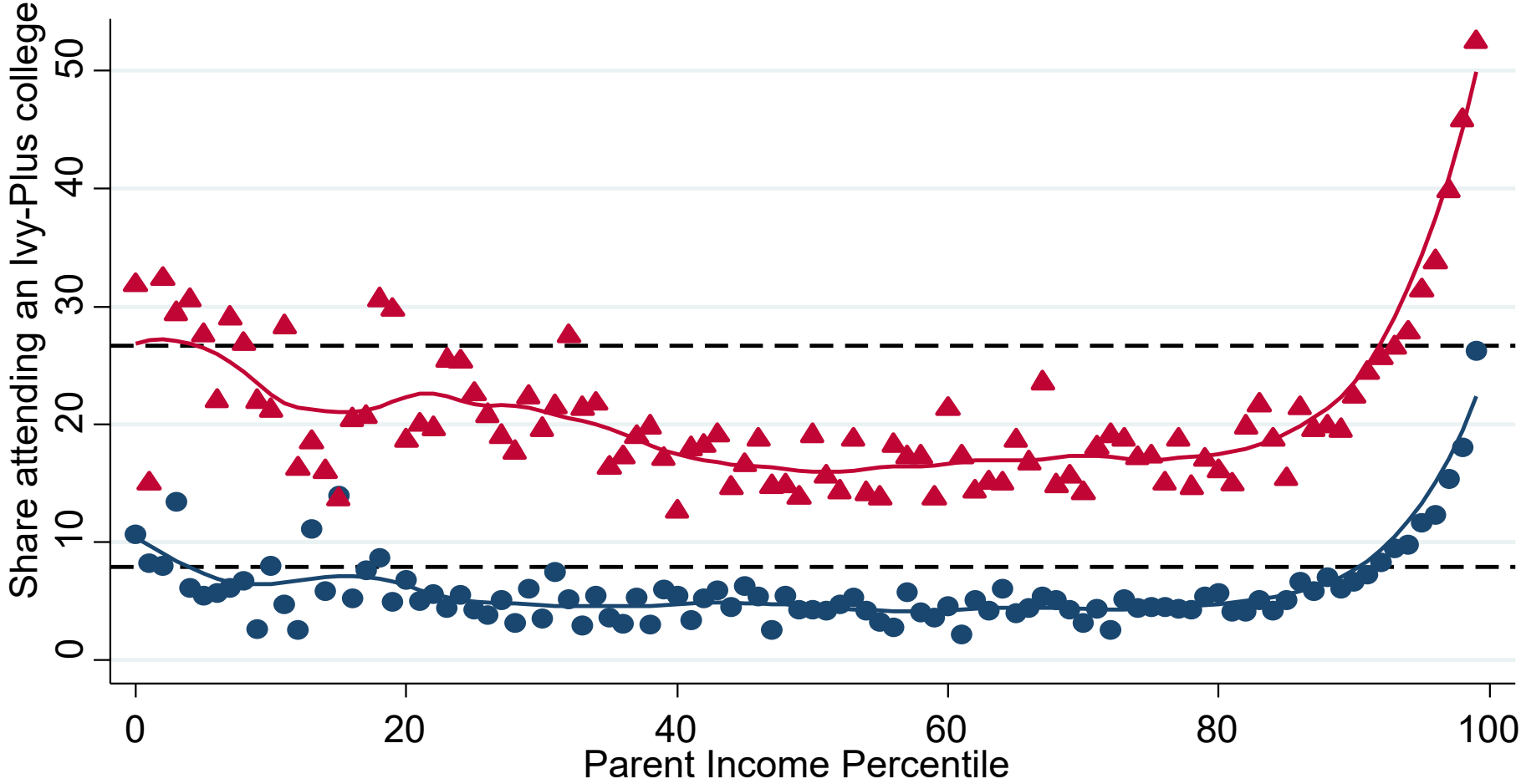
- At Ivy-plus schools, we find substantial undermatching of middle-income students, but little undermatching of low-income students
 - There are very few students from low-income families with sufficiently high test scores to get into such colleges
 - Among students scoring 1300+ on the SAT/ACT, only 3.7% come from the lowest quintile of parent income

Parental Income among High-SAT Scorers vs. Ivy-Plus Colleges



Ivy-Plus Attendance Rates by Parental Income Conditional on SAT/ACT Scores

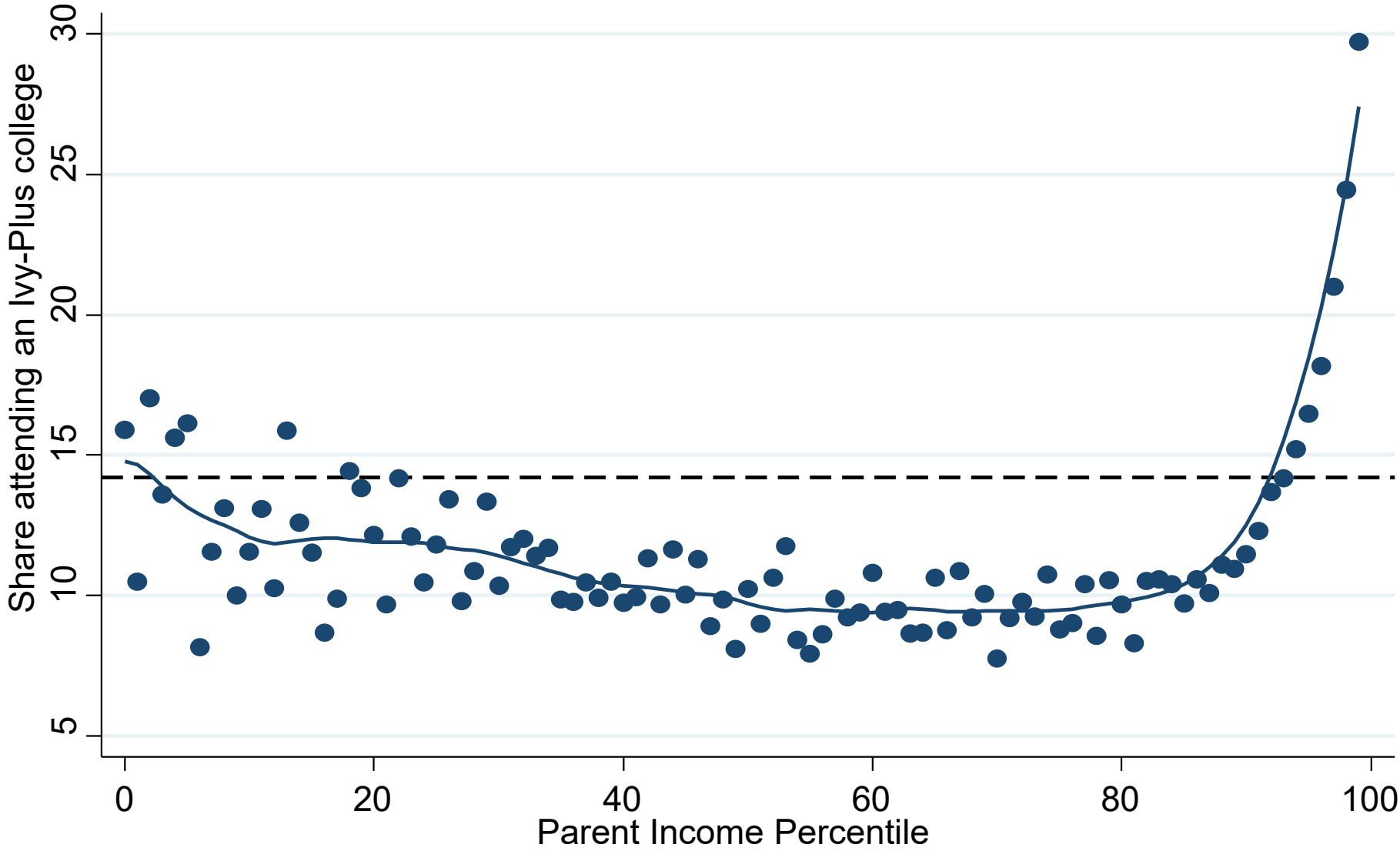
Students with Scores of 1400 or 1500 on SAT/ACT



- ▲ Children with a 1500 SAT/ACT
- Children with a 1400 SAT/ACT

Ivy-Plus Attendance Rates by Parental Income Conditional on SAT/ACT Scores

All SAT/ACT Scores, Reweighted to Match Ivy-Plus SAT/ACT Score Distribution



Undermatching at Ivy-Plus Colleges

- At Ivy-plus schools, we find substantial undermatching of middle-income students, but little undermatching of low-income students
 - There are very few students from low-income families with sufficiently high test scores to get into such colleges
 - Among students scoring 1300+ on the SAT/ACT, only 3.7% come from the lowest quintile of parent income
- Why do we find fewer high-achieving, low-income children relative to Hoxby and Avery (2013)?
 - Individual measurement of parent income (rather than tract-based imputation)
 - Measure income and percentile cutoffs in the same dataset

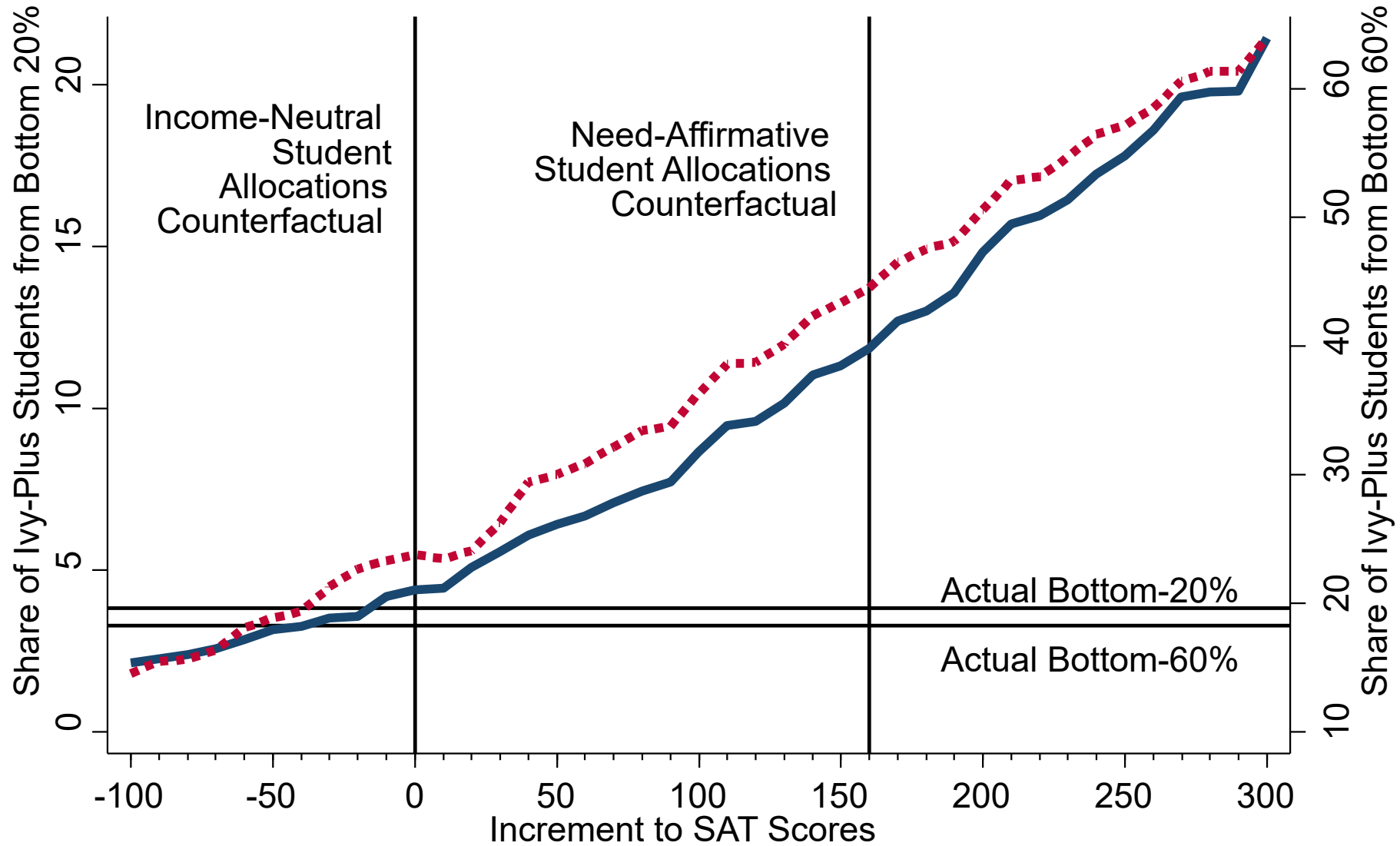
Need-Affirmative Counterfactual Allocation

- To further increase representation of lowest-income students at elite private colleges, need to give them some preference in application/admissions process
- We model this by giving students from low-income families an increment to their SAT score
 - Sliding-scale increment by income: increase SAT scores by 80%, 60%, and 40% as much in 2nd-4th parent income quintiles
 - Hold number of students, home state mix, and racial diversity constant at each school

Need-Affirmative Counterfactual Allocation

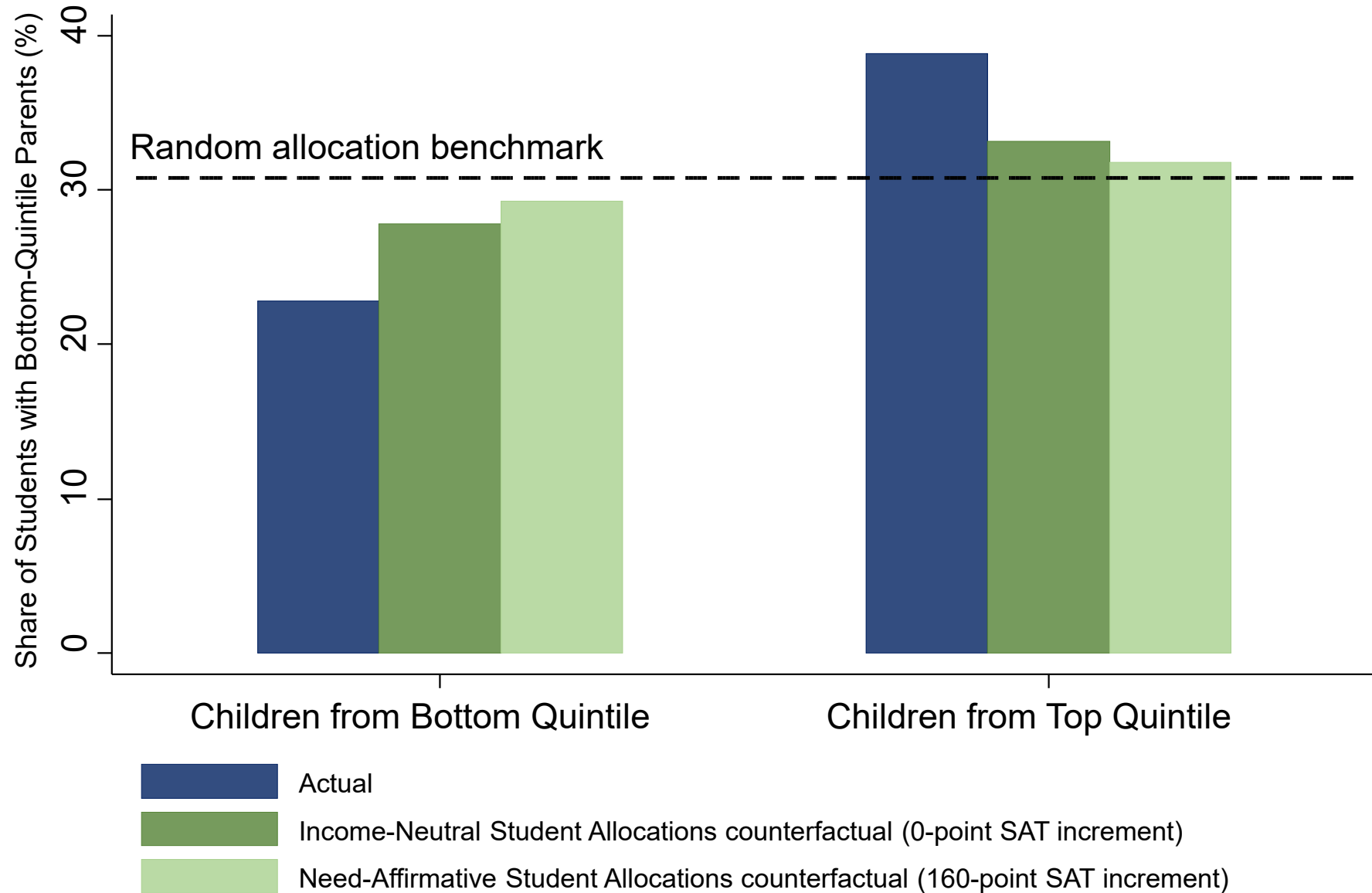
- Finding: Adding 160 points to the SAT scores of students from the bottom quintile would desegregate higher education system
 - Roughly the same distribution of parent income across all college selectivity tiers
- At Ivy-Plus colleges, this would require increasing attendance rate for bottom-quintile students from 7.3% to 35.8%
 - Similar to the preference given to legacies, athletes, minority students at elite private colleges [Espenshade et al. 2004, Arcidiacono et al. 2019]

Counterfactual Low-Income Shares at Ivy-Plus Colleges



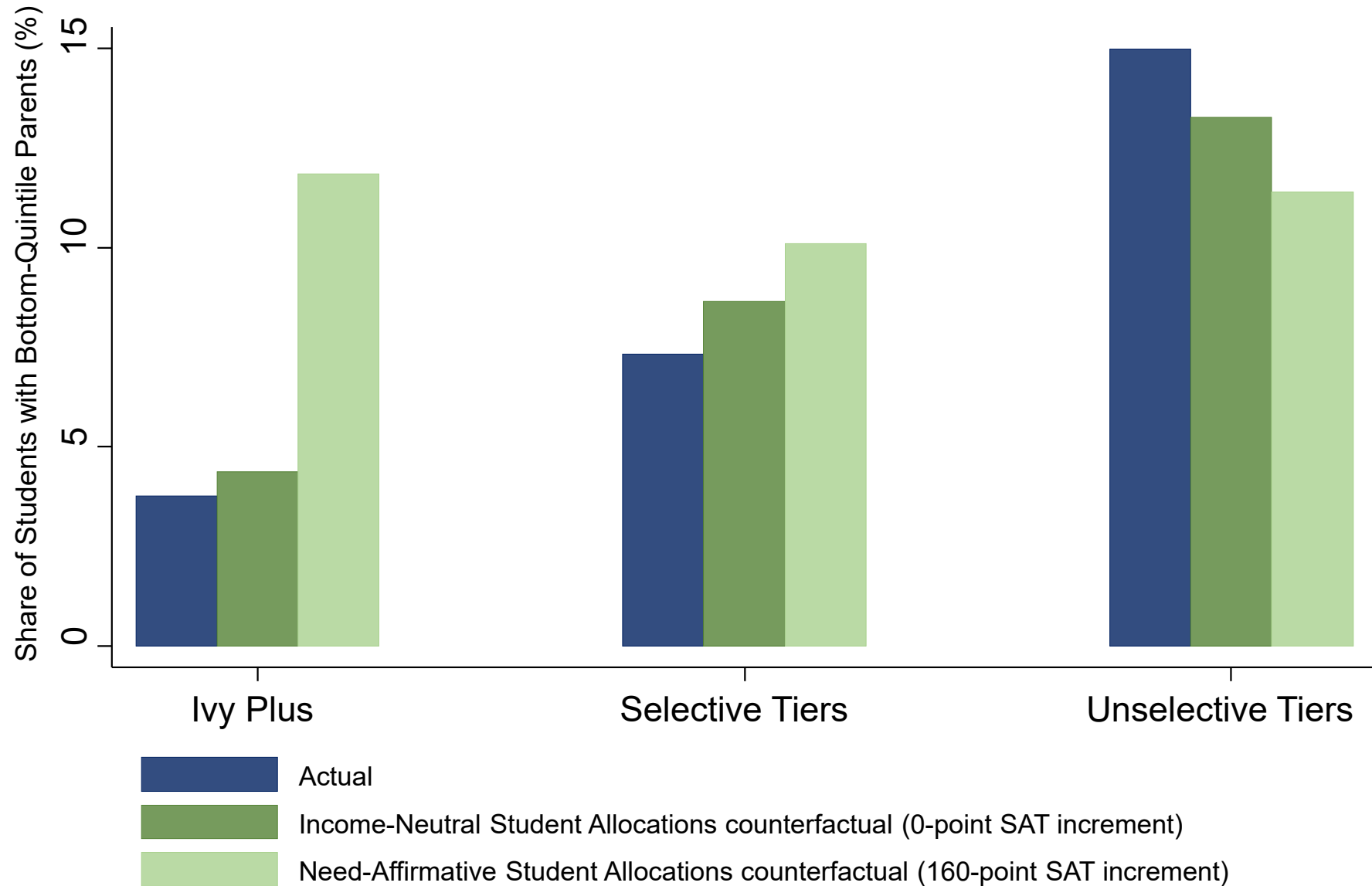
Impacts of Counterfactuals on Income Segregation and Intergenerational Mobility

Fraction of Peers from Top-Quintile for Students from Bottom vs. Top Quintiles



Impacts of Counterfactuals on Income Segregation and Intergenerational Mobility

Fraction of Students with Low-Income Parents in Selected Tiers



Potential Effect on Intergenerational Mobility

- Estimate causal effect of each school following Dale and Krueger (2002):
 - Begin with differences in earnings across colleges that control flexibly only for SAT and parent income
 - Adjusting for race, gender, location, and application set only reduces differences by 10-15%
 - We conservatively assume 80% of SAT-controlled differences are causal

Fraction of Differences in Earnings Across Colleges Due to Causal Effects

Dependent Variable: college fixed effect, conditional on parental income, race and SAT/ACT

	Race, Gender * SAT/ACT (1)	HS FE's (2)	HS FE's * Race (3)	Application Set and HS FEs (4)	Application Set, HS FEs * Race (5)	Bottom Quintile Only (6)
College fixed effect (conditional on parent inc., race, SAT/ACT)	1.003 (0.006)	0.907 (0.010)	0.903 (0.010)	0.857 (0.012)	0.850 (0.012)	0.850 (0.015)
Adj. R-squared	0.968	0.886	0.883	0.889	0.886	0.750

Controls

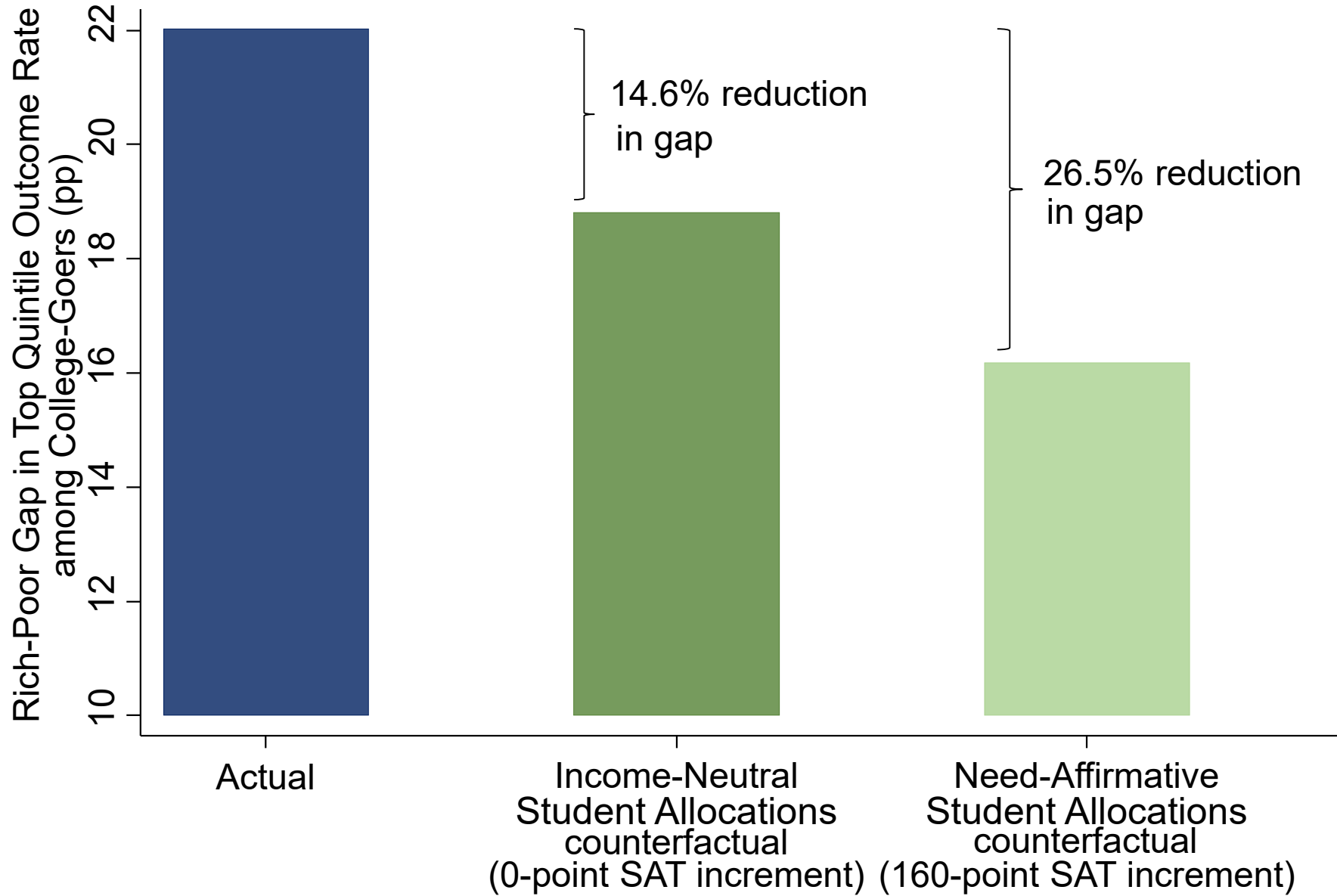
Interactions of race, gender w/SAT/ACT	X	X	X	X	X	X
High school FE's		X	X	X	X	X
High school FE's interacted with race			X		X	X
Mean SAT of schools to which scores were sent				X	X	X

Potential Effect on Intergenerational Mobility

- Estimate causal effect of each school following Dale and Krueger (2002):
 - Begin with differences in earnings across colleges that control flexibly only for SAT and parent income
 - Adjusting for race, gender, location, and application set only reduces differences by 10-15%
 - We conservatively assume 80% of SAT-controlled differences are causal
- Income-neutral counterfactual would reduce top-quintile student earnings gap by 14.6%
 - Need-affirmative counterfactual would reduce top-quintile student earnings gap by 26.5%

Impacts of Counterfactuals on Income Segregation and Intergenerational Mobility

Gaps in Chance of Reaching Top Earnings Quintile

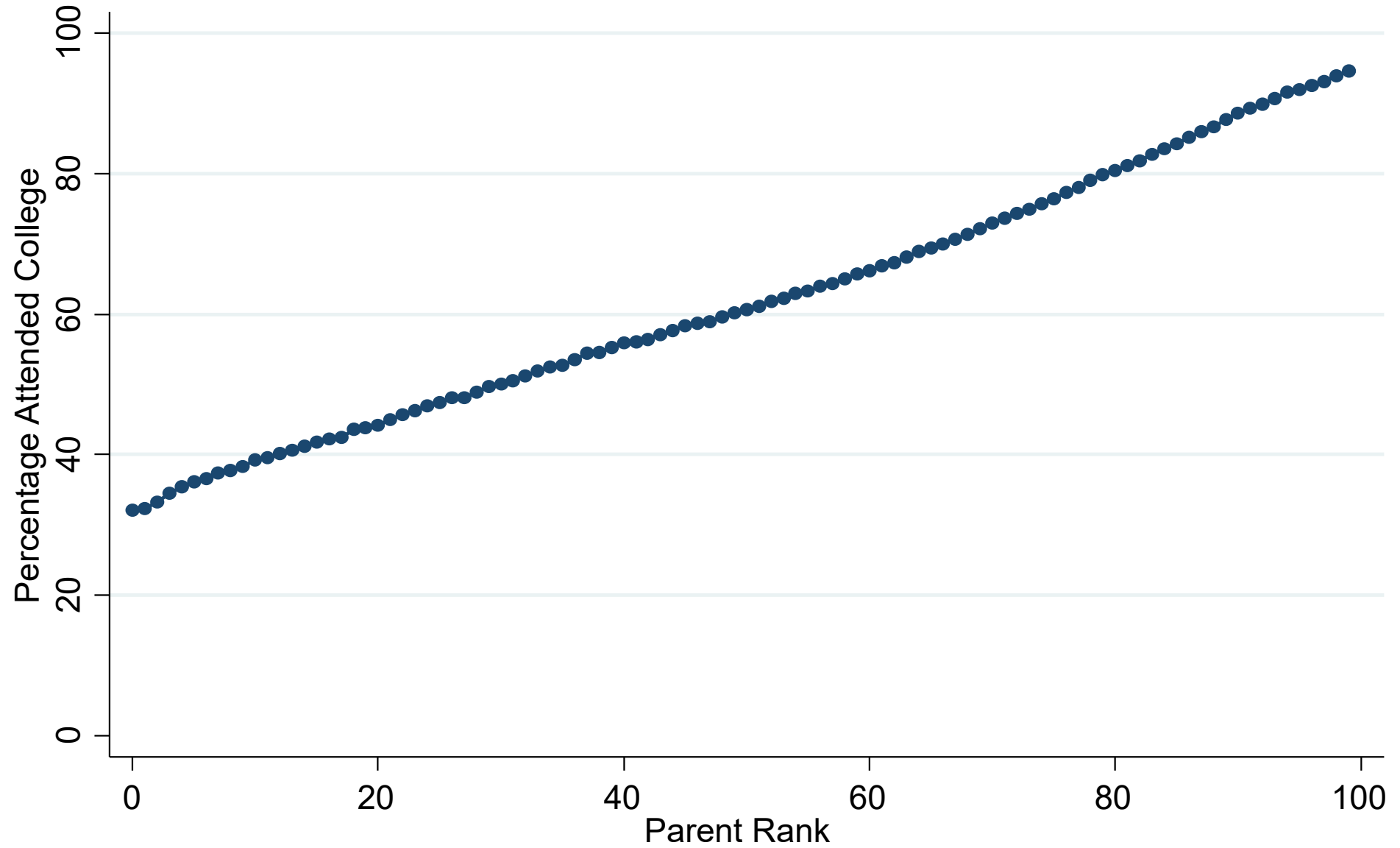


Conclusion

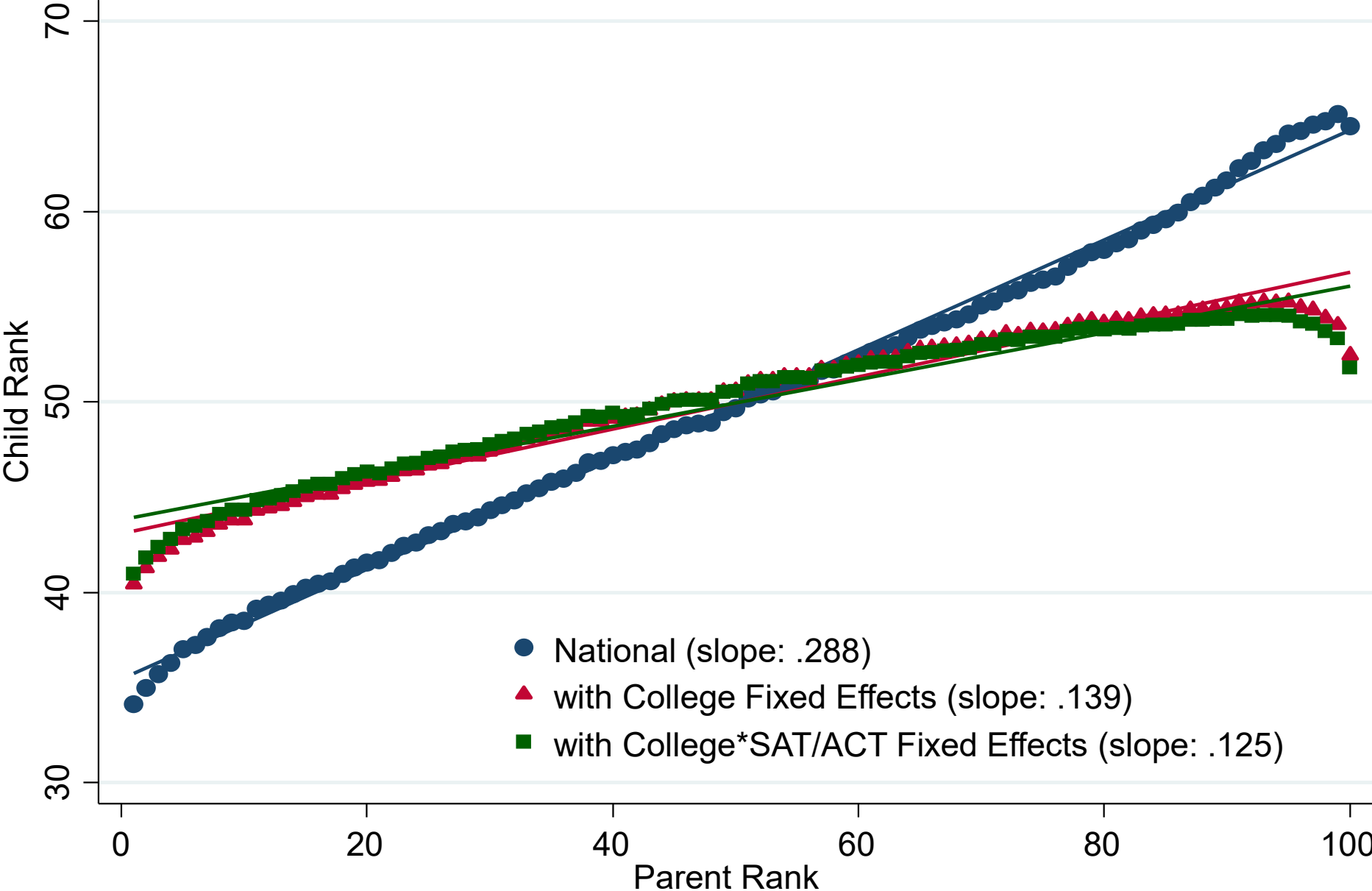
- Feasible changes in the allocation of students to colleges could change segregation across colleges substantially and increase intergenerational mobility significantly
 - Even without any changes in “value-added” of colleges (or national expenditure on higher education) or in pre-college environmental disparities
- Studying precisely how to increase the representation of students from lower-income families at more selective colleges is of great value
 - Do disparities arise due to differences in applications, admissions, or matriculation rates?
 - How can these disparities be tackled most effectively?

Supplementary Figures

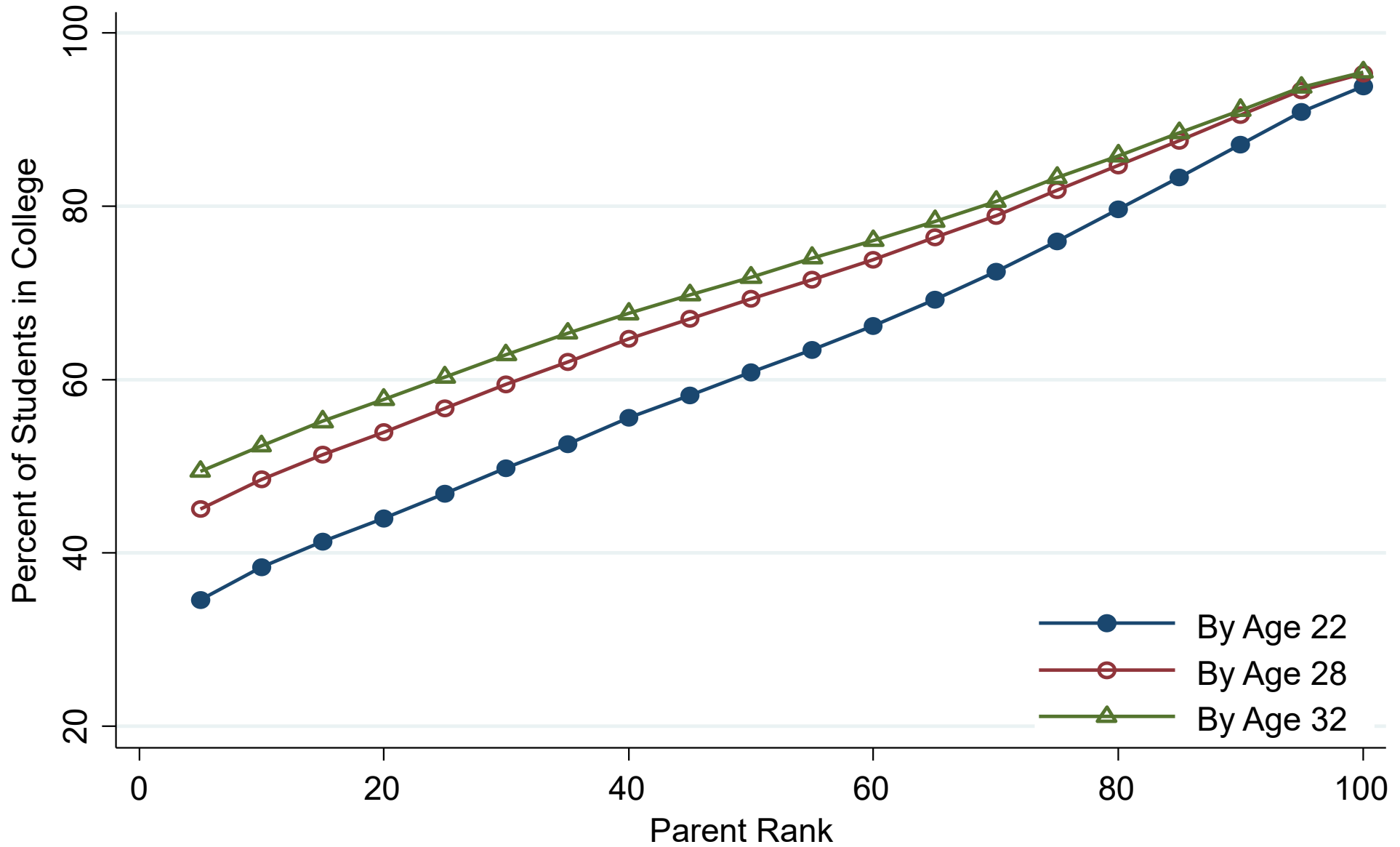
College Attendance Rates by Parent Income Percentile



Relationship Between Children's and Parents' Ranks within Colleges

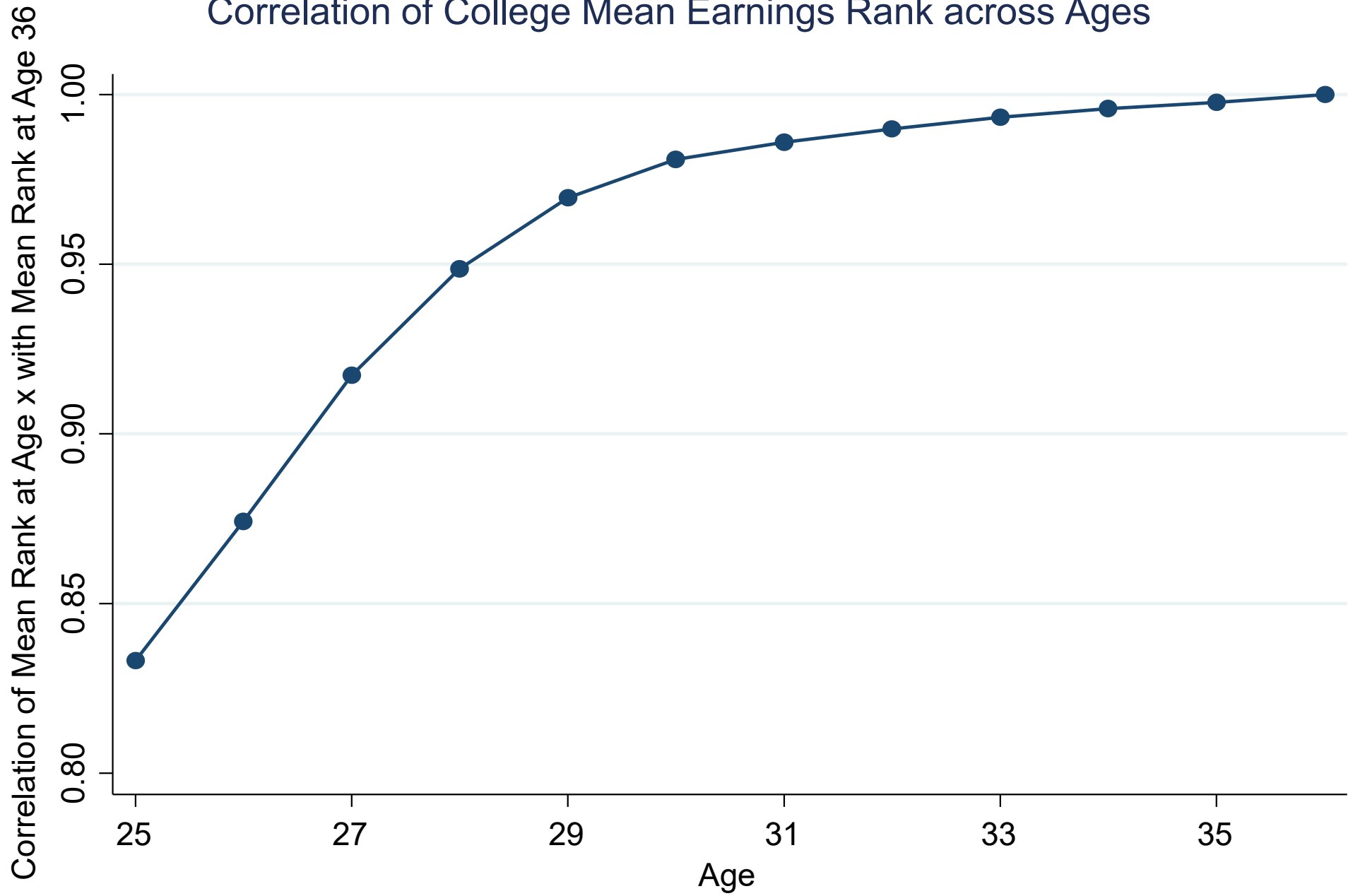


College Attendance Rates by Parental Income and Age



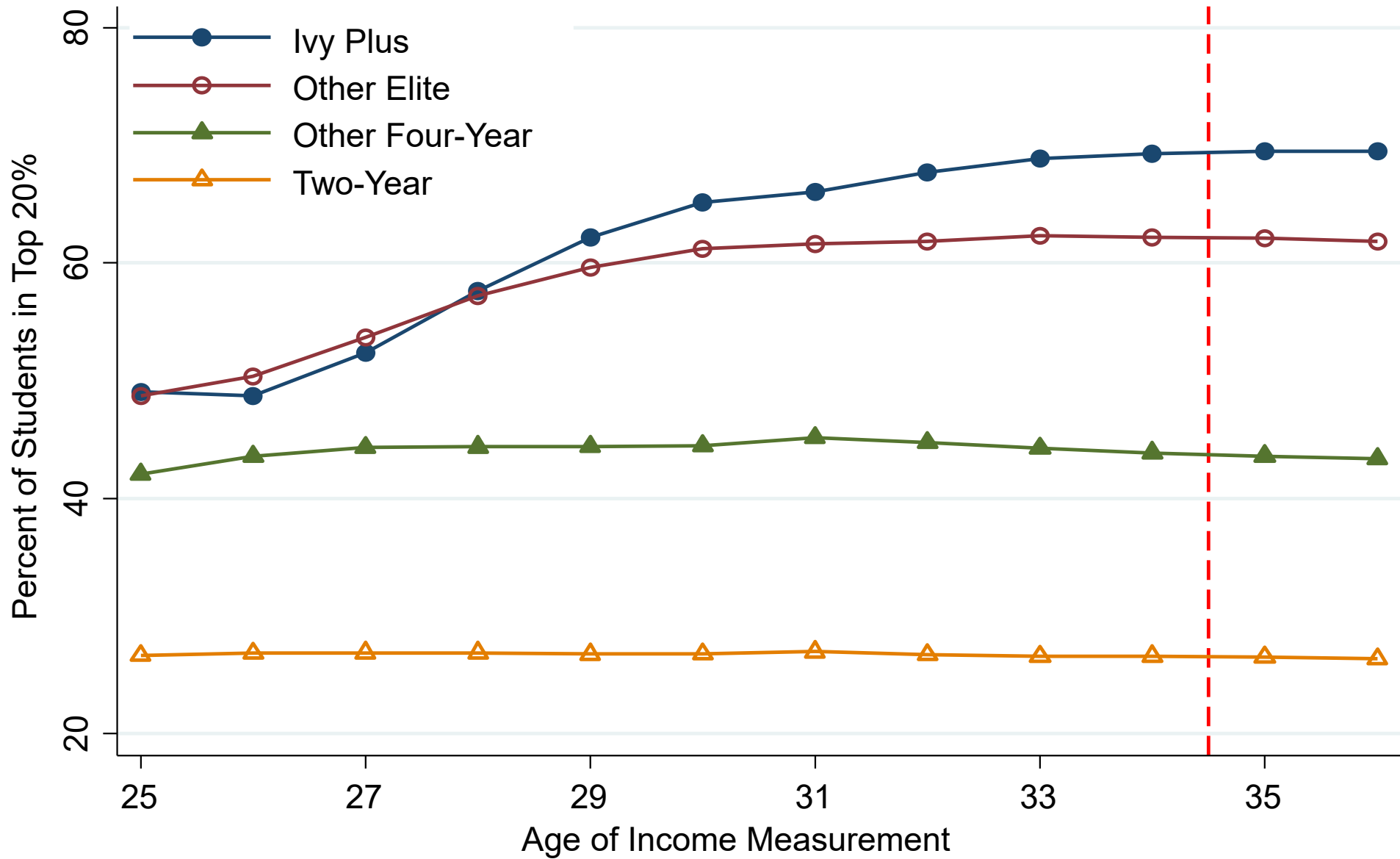
Children's Earnings Ranks by Age of Earnings Measurement

Correlation of College Mean Earnings Rank across Ages



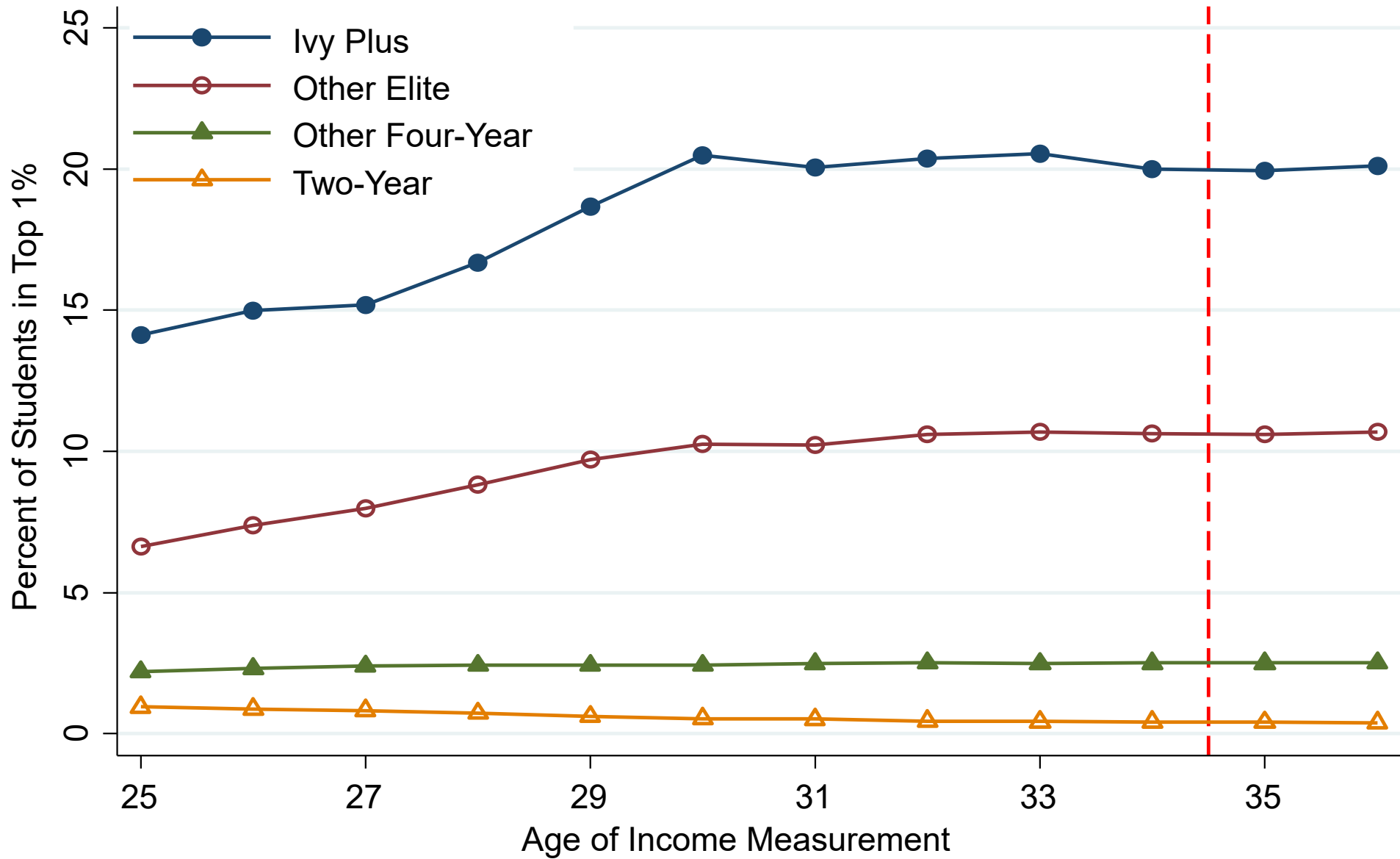
Children's Earnings Ranks by Age of Earnings Measurement

Fraction of Children in Top Quintile

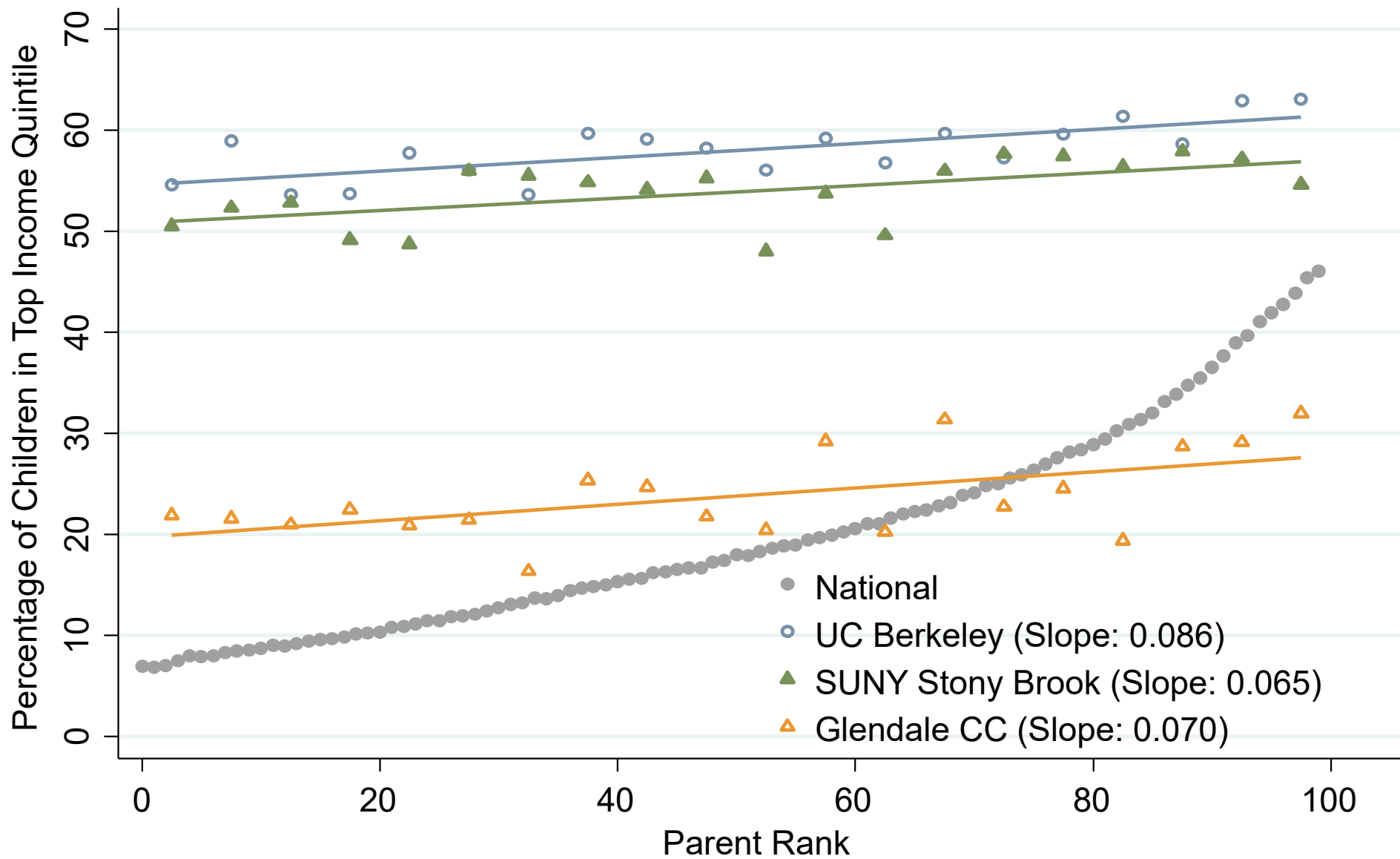


Children's Earnings Ranks by Age of Earnings Measurement

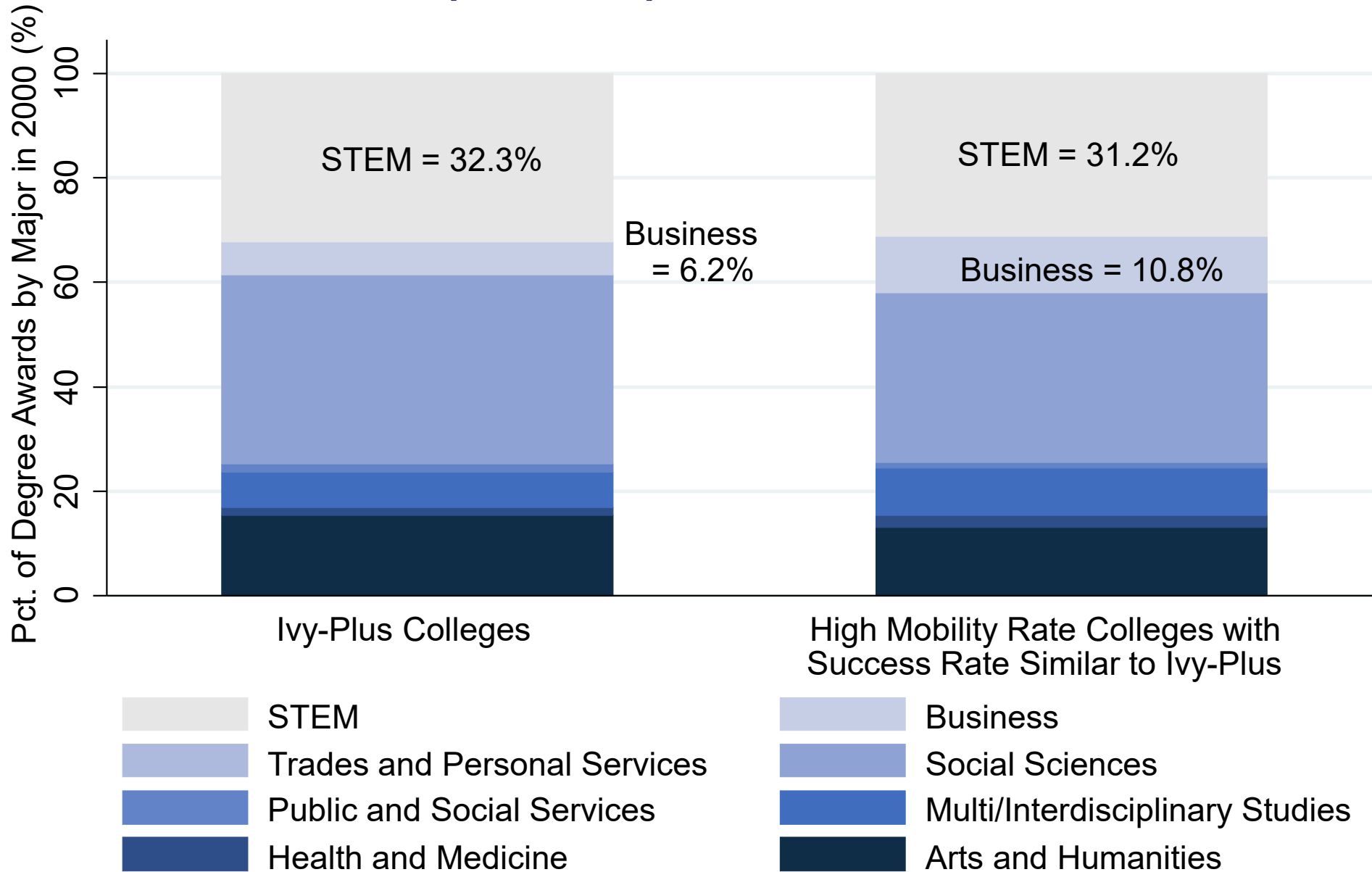
Fraction of Children in Top 1% by Age and College Tier



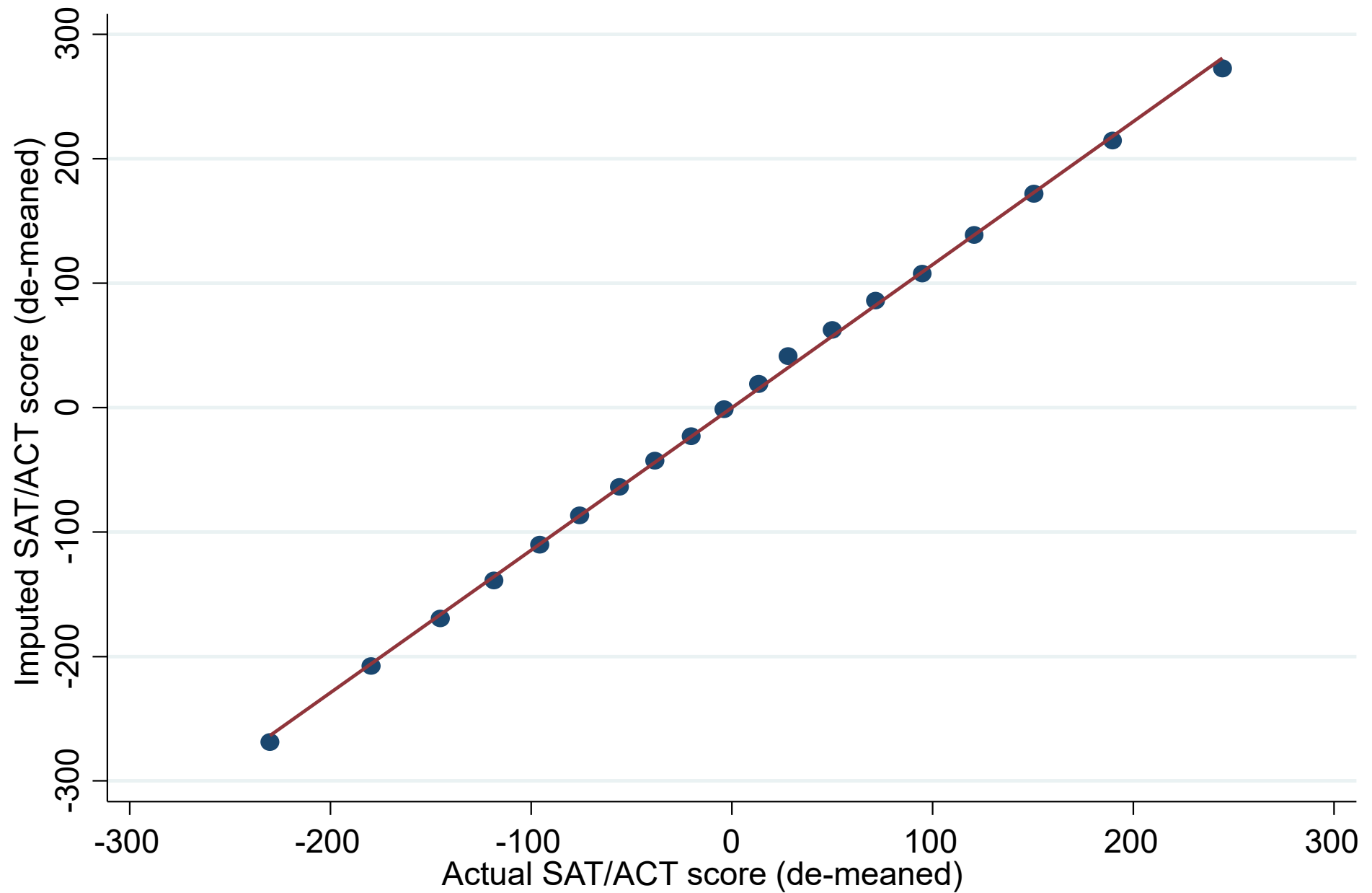
Fraction of Children who Reach Top Quintile by Parent Income Rank



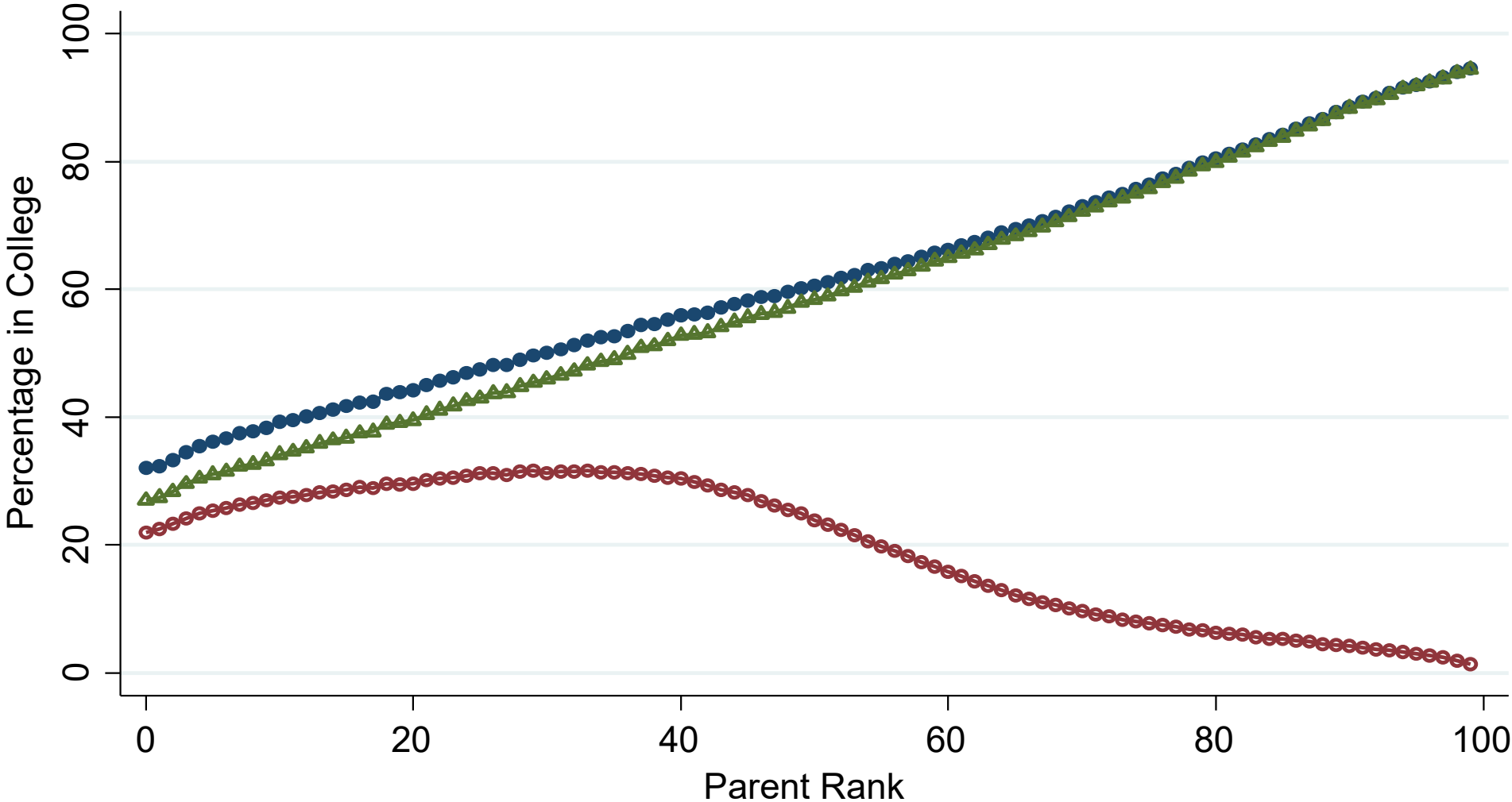
Distribution of Majors: Ivy-Plus Colleges vs. High-Mobility-Rate Colleges with Comparable Top-Quintile Outcome Rates



Validation of SAT/ACT Imputation



College Attendance Rates in 1098-T and Pell Records by Parental Income



- IRS Data
- Pell Grant Data
- Combined IRS + Pell Data