NBER WORKING PAPER SERIES

THE DETERMINANTS OF INCOME SEGREGATION AND INTERGENERATIONAL MOBILITY: USING TEST SCORES TO MEASURE UNDERMATCHING

Raj Chetty
John N. Friedman
Emmanuel Saez
Nicholas Turner
Danny Yagan

Working Paper 26748
http://www.nber.org/papers/w26748

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2020

This work was conducted under IRS contract TIRNO- 16-E-00013 and reviewed by the Office of Tax Analysis at the U.S. Treasury. We thank Joseph Altonji, David Deming, Eric Hanushek, Jess Howell, Michael Hurwitz, Lawrence Katz, David Lee, Richard Levin, Sean Reardon, anonymous referees, and numerous seminar participants for helpful comments; Trevor Bakker, Kaveh Danesh, Katie Donnelly Moran, Niklas Flamang, Robert Fluegge, Jamie Fogel, Benjamin Goldman, Clancy Green, Sam Karlin, Carl McPherson, Daniel Reuter, Benjamin Scuderi, Priyanka Shende, Jesse Silbert, Mandie Wahlers, and our other pre-doctoral fellows for outstanding research assistance; and especially Adam Looney for supporting this project. Chetty, Friedman, Saez, and Yagan acknowledge funding from the Russell Sage Foundation, the Bill & Melinda Gates Foundation, the Robert Wood Johnson Foundation, the Schmidt Futures Foundation, the Center for Equitable Growth at UC-Berkeley, the Washington Center for Equitable Growth, the UC Davis Center for Poverty Research, the Alfred P. Sloan Foundation, the Laura and John Arnold Foundation, the Chan-Zuckerberg Initiative, the Overdeck Foundation, and Bloomberg Philanthropies. The opinions expressed in this paper are those of the authors alone and do not necessarily reflect the views of the Internal Revenue Service, the U.S. Treasury Department, the Federal Reserve Board of Governors, the College Board, or the National Bureau of Economic Research. The results in this paper build on those reported in an earlier NBER working paper entitled "Mobility Report Cards: The Role of Colleges in Intergenerational Mobility." A manuscript that combines the results from both papers is available at https://opportunityinsights.org/wp-content/uploads/2020/02/coll_mrc_qje_paper.pdf and is forthcoming in the Quarterly Journal of Economics.

The Determinants of Income Segregation and Intergenerational Mobility:
Using Test Scores to Measure Undermatching
Raj Chetty, John N. Friedman, Emmanuel Saez, Nicholas Turner, and Danny Yagan
NBER Working Paper No. 26748
February 2020
JEL No. J0

ABSTRACT

We analyze how changes in the allocation of students to colleges would affect segregation by parental income across colleges and intergenerational mobility in the United States. We do so by linking data from tax records on parents' incomes and students' earnings outcomes for each college to data on students' SAT and ACT scores. We find that equalizing application, admission, and matriculation rates across parental income groups conditional on test scores would reduce segregation substantially, primarily by increasing the representation of middle-class students at more selective colleges. However, it would have little impact on the fraction of low-income students at elite private colleges because there are relatively few students from low-income families with sufficiently high SAT/ACT scores. Differences in parental income distributions across colleges could be eliminated by giving low and middle-income students a sliding-scale preference in the application and admissions process similar to that implicitly given to legacy students at elite private colleges. Assuming that 80% of observational differences in students' earnings conditional on test scores, race, and parental income are due to colleges' causal effects— a strong assumption, but one consistent with prior work — such changes could reduce intergenerational income persistence among college students by about 25%. We conclude that changing how students are allocated to colleges could substantially reduce segregation and increase intergenerational mobility, even without changing colleges' educational programs.

Raj Chetty
Department of Economics
Harvard University
Littauer 321
Cambridge, MA 02138
and NBER
chetty@fas.harvard.edu

John N. Friedman
Department of Economis
Robinson Hall
Brown University
Providence, RI 02912
and NBER
john_friedman@brown.edu

Emmanuel Saez
Department of Economics
University of California,
Berkeley 530 Evans Hall #3880
Berkeley, CA 94720
and NBER
saez@econ.berkeley.edu

Nicholas Turner
Federal Reserve Board
Nick.Turner@frb.gov

Danny Yagan
Department of Economics
University of California, Berkeley
530 Evans Hall, #3880
Berkeley, CA 94720
and NBER
yagan@berkeley.edu

Publicly available college-level data is available at
https://opportunityinsights.org/data/?geographic_level=0&topic=105&paper_id=536#resource-listing

I Introduction

What role do colleges play in shaping intergenerational mobility in America? In a paper released in 2017 (Chetty et al. 2017), we took a step toward answering this question by publishing statistics on the distribution of students' earnings in their thirties and their parents' incomes for each college in the U.S. using anonymized data from tax records. We then documented three facts that emerged from those statistics. First, the degree of segregation by parental income is very high across colleges, similar to levels of segregation across neighborhoods in the average American city. Second, children from low- and high-income families who attend the same college go on to have relatively similar levels of earnings in adulthood. Third, colleges with high levels of student earnings (e.g., Ivy League colleges) typically have few students from poorer backgrounds, limiting their scope to serve as ladders for upward mobility.

Building on these earlier findings, in this paper we study how much of the difference in the types of colleges that children from low vs. high-income families attend is explained by differences in their qualifications when they apply to college. We then analyze the extent to which changes in the college application and admission process could reduce segregation by parental income across colleges and increase intergenerational income mobility. We study these questions by linking tax data to information on SAT and ACT scores to construct a comprehensive dataset that contains information on the colleges that students attend, their earnings in their early thirties, their parents' household incomes, and their SAT/ACT scores. In our baseline analysis, we focus on children born between 1980 and 1982 – the oldest children whom we can reliably link to parents – and assign children to colleges based on the college they attend most frequently between the ages of 19 and 22.[1]

We begin by evaluating the extent to which differences in parental income distributions across colleges can be explained by differences in academic preparation before students apply to college, as proxied for by SAT or ACT scores.[2] We find that at any given level of SAT/ACT scores, children from higher-income families attend more selective colleges, suggesting that low- and middle-income students "undermatch" to colleges (Bowen, Chingos and McPherson 2009). To quantify the

[1]We measure children's earnings between the ages of 32 and 34; we show that children's percentile ranks in the earnings distribution stabilize by age 32 at all types of colleges.

[2]We follow a large body of prior work in using standardized test scores as a widely available measure of end-of-high-school academic preparation (e.g., James et al. 1989, Dale and Krueger 2002) that is highly predictive of long-term outcomes such as earnings. We confirm and extend these results by showing that SAT scores are strong predictors of later earnings even conditional on parental income, race, and the high school or college a child attends in Online Appendix L. Of course, other measures may also be helpful in assessing academic preparation and qualifications. Our analysis does not speak to the relative merits of test scores vs. other proxies to assess pre-college qualifications.

degree of undermatching, we construct an "income-neutral" student allocation process, in which we fill each college's slot for a current student who has test score s with a *random* draw from the population of college students with test score s who come from the same state and are of the same race. In this scenario, colleges continue to enroll students based on both academic and non-academic credentials but eliminate variation in enrollment rates by parental income – whether due to differences in application, admissions, or matriculation – among students with comparable academic credentials, preserving the racial and geographic composition and the total size of each college. This counterfactual thus provides a natural benchmark to gauge the extent to which student bodies are representative of the underlying population of academically qualified students.[3]

Income segregation across colleges would fall significantly if students enrolled at colleges in an income-neutral manner conditional on their test scores. The degree of under-representation of students from the bottom parental income quintile at selective (Barron's Tier 6 or higher) colleges would fall by 38% relative to a benchmark in which all colleges have the same fraction of bottom-quintile students as in the current population of college-goers. This is because top-quintile students are currently 34% more likely to attend selective colleges than their bottom-quintile peers with the same test scores. The income-neutral allocation would also increase the representation of middle-income students (the second, third, and fourth quintiles) at selective colleges substantially.

The picture is somewhat different at the most selective elite private (Ivy-Plus) colleges. There, the fraction of students from the middle class (the second, third, and fourth quintiles) would rise substantially, from 28% to 38%, under income-neutral allocations. But, there would be little absolute change (from 3.8% to 4.4%) in the fraction of students from the bottom income quintile, reducing under-representation relative to the benchmark in which all colleges have the same fraction of bottom-quintile students by only 9%. These findings show that it is in fact *middle-income* students who attend Ivy-plus colleges at the lowest rates, conditional on test scores – what many have referred to as the "missing middle" at elite private colleges.[4] Our results imply much less undermatching of high-achieving *low-income* students at such colleges than found by Hoxby and Avery (2013) because there are few children from low-income families who have sufficiently high SAT/ACT scores. For instance, only 3.7% of children who score above a 1300 on the SAT come

[3]This counterfactual exercise differs from the approach of simply admitting students with the highest test scores considered by Bastedo and Jaquette (2011) and Carnevale et al. (2019). Since colleges do place significant weight on factors unrelated to test scores in practice, we believe this counterfactual provides a more plausible benchmark for understanding the extent to which differences in test scores can explain income segregation across colleges.

[4]The term "missing middle" has been used to describe the relative under-representation of middle-class students at elite private institutions since at least Todd (1976). More recently, Caroline Hoxby and Sarah Turner document results consistent with these findings, as reported in Rampell (2019).

from families in the bottom income quintile.[5] High-scoring students from low-income families are scarce in substantial part because of disparities in schools, neighborhoods, and other environmental factors that cumulate since birth (Heckman and Krueger 2005, Fryer and Levitt 2013, Chetty and Hendren 2018, Reardon 2019). These pre-college disparities limit the scope to increase the number of students from the lowest-income families at elite colleges purely by recruiting more applications.

Further increasing the fraction of low-income students at selective colleges would require policies that induce low-income students to attend highly selective colleges at higher rates than higher-income students with currently comparable SAT scores. If low-income (bottom quintile) students attended colleges comparable to high-income (top quintile) students with 160 point higher SAT scores, the higher education system would be fully desegregated, in the sense that parental income distributions would be very similar across all colleges.[6] To benchmark the magnitude of this change, a 160-point SAT increment would be equivalent to increasing Ivy-plus attendance rates from 7.3% to 25.8% for low-income students with an SAT score of 1400. This increment is very similar in magnitude to the implicit preference in admissions given to various preferred groups, such as legacy students, recruited athletes, and underrepresented minorities, at elite colleges, who are admitted at substantially higher rates than other students with similar qualifications (Espenshade, Chung and Walling 2004, Arcidiacono, Kinsler and Ransom 2019).[7]

How would such changes in segregation affect intergenerational mobility? To answer this question, we need an estimate of the fraction of the earnings premium at each college (conditional on parental income, race, and SAT/ACT scores) that is due to the causal effect of attending that college. Naturally, our simulated impacts on intergenerational mobility are highly sensitive to this parameter: if differences in earnings across colleges are driven purely by selection rather than causal effects, reallocating students across colleges would have no impact on mobility. To gauge what fraction of the difference in earnings across colleges is due to causal effects, we regress stu-

[5] We find many fewer high-achieving students from low-income families than that estimated by Hoxby and Avery. This difference arises because we measure parental income at the individual level rather than using geographic imputations and because of differences in the thresholds used to define quantiles of the income distribution; see Section V.A for details.

[6] Phasing out this increment roughly linearly from 160 SAT points in the bottom quintile down to 0 for the students in the top quintile leads to equal representation of students from all parental income levels across colleges. Note that we use the SAT here simply as a convenient metric to quantify the degree of need-affirmative preference needed to desegregate colleges; in practice, one could implement such policies using a variety of other metrics and approaches.

[7] Our results do not speak to the debate about whether standardized tests provide comparable measures of aptitude for students from low vs. high income families. We simply use test scores to quantify the gap between students from low vs. high-income families in end-of-high-school academic qualifications. Whether that gap can be closed through changes in K-12 education, test design or preparation, or the college application or admissions process is a question left for future work.

dents' earnings on our estimates of mean earnings premia (conditional on race, parental income, and test scores), controlling for other observable characteristics such as gender, high-school GPA, and high-school fixed effects. We then follow Dale and Krueger (2002) and additionally control for the set of colleges to which a student applied to capture selection on unobservables. Including such controls yields a coefficient between 0.8-1, suggesting that at least 80% of the difference in earnings premia across colleges (conditional on parental income, race, and test scores) reflects causal effects. We therefore assume that 80% of the earnings premium at each college is driven by a causal effect in our baseline analysis. We also assume that student reallocations do not change colleges' causal effects, even though the composition of the student body might change substantially.

We measure intergenerational mobility as the difference in the chance that college students from low vs. high income families reach the top earnings quintile, a simple measure of relative mobility (Chetty et al. 2014). Empirically, this difference is 22 percentage points for children in the 1980-82 birth cohorts. The income-neutral benchmark would narrow the gap by 15%, while need-affirmative admissions would narrow the gap by 27%. These are substantial effects given that children's outcomes in adulthood are shaped by a cumulation of environmental factors from birth until the point they enter the labor market (Chetty and Hendren 2018) and most people spend at most 25% of their pre-labor-market years in college. The precise magnitudes that result from these simulations must of course be interpreted with caution because they hinge on strong assumptions, namely about the causal effect of colleges. Nevertheless, they suggest that changing which colleges students attend – i.e., reducing segregation without making any efforts to increase colleges' value-added or reduce disparities that emerge before students apply to college – could increase economic mobility substantially.

Our analysis builds on and reconciles various conflicting findings in prior work. First, several papers have studied income segregation in higher education by selectivity tier or at selected colleges (e.g., Avery et al. 2006, Goodman 2008, Deming and Dynarski 2010, Hoxby and Turner 2013, Marx and Turner 2015, Andrews, Imberman and Lovenheim 2016, Manoli and Turner 2018). These studies find a wide range of estimates using small samples; for instance, the estimated fraction of students from bottom-quartile families at elite colleges ranges from 3% (Carnevale and Strohl 2010) to 11% (Bowen, Kurzweil and Tobin 2006, Chapter 7) across studies. Our statistics (reported in Chetty et al. (2017)) provide more definitive estimates of the degree of segregation across college tiers, shed light on segregation across colleges within selectivity tiers, and offer the first statistics on top-income shares by college.

Second, a smaller literature has measured the returns to attending certain colleges using quasi-experimental methods (e.g., Black and Smith 2004, Hoekstra 2009, Hastings, Neilson and Zimmerman 2013, Zimmerman 2014, Kirkeboen, Leuven and Mogstad 2016, Cellini and Turner 2019). Our analysis complements these studies by providing information on earnings distributions for all colleges. These data allow us to characterize how students' earnings distributions vary with parental income within each college and identify "outlier" colleges in terms of students' outcomes whose admissions policies or educational practices could be studied in future quasi-experimental work.

Finally, our counterfactual analysis follows prior work examining how alternative admissions rules would affect the composition of colleges by selectivity tier (e.g., Arcidiacono 2005, Bowen, Kurzweil and Tobin 2006, Epple, Romano and Sieg 2006, Krueger, Rothstein and Turner 2006, Howell 2010). This work has again reached conflicting conclusions on the degree of undermatching and the consequences of alternative admissions regimes (Carnevale and Rose 2004, Hill and Winston 2006, Carnevale and Strohl 2010, Bastedo and Jaquette 2011, Hoxby and Avery 2013). In addition to reconciling these findings, we contribute to this literature by (1) analyzing counterfactuals across all colleges rather than by college tier, which proves to be quantitatively important and (2) showing impacts not just on the composition of the student body but on rates of intergenerational mobility.

The paper is organized as follows. Section II describes the data. Section III examines the relationship between SAT/ACT scores and parent income (undermatching). Section IV discusses the counterfactual simulations. Section V concludes. College-level statistics and replication code can be downloaded from the project website.

II Data

In this section, we describe how we construct our analysis sample, define the key variables we use in our analysis, and present summary statistics. The data we use to measure students' and parents' incomes and college attendance are identical to those in Chetty et al. (2017) and hence much of this section overlaps with that paper.

II.A Sample Definition

Our primary sample of children consists of all individuals in the U.S. who (1) have a valid Social Security Number (SSN) or Individual Taxpayer Identification Number (ITIN), (2) were born between 1980-1991, and (3) can be linked to parents with non-negative income in the tax data (see Appendix A for more details).[8] There are approximately 48.1 million people in this sample.

We identify a child's parents as the most recent tax filers to claim the child as a child dependent during the period when the child is 12-17 years old. If the child is claimed by a single filer, the child is defined as having a single parent. We assign each child a parent (or parents) permanently using this algorithm, regardless of any changes in parents' marital status or dependent claiming. Children who are never claimed as dependents on a tax return cannot be linked to their parents and are excluded from our analysis. However, almost all parents file a tax return at some point when their child is between ages 12-17, either because their incomes lie above the filing threshold or because they are eligible for a tax refund (Cilke 1998). Thus, the number of children for whom we identify parents exceeds 98% of children born in the U.S. between 1980 and 1991 (Appendix Table I).[9]

II.B College Attendance

Data Sources. We obtain information on college attendance from two administrative data sources: federal tax records and Department of Education records spanning 1999-2013.[10] We identify students attending each college in the administrative records primarily using Form 1098-T, an information return filed by colleges on behalf of each of their students to report tuition payments. All institutions qualifying for federal financial aid under Title IV of the Higher Education Act of 1965 must file a 1098-T form in each calendar year for any student that pays tuition. Because the 1098-T data do not always cover students who pay no tuition—who are typically low-income students receiving financial aid—we supplement the 1098-T data with Pell grant records from the Department of Education's National Student Loan Data System (NSLDS). See Appendix B for details on these two data sources and how we assign students to colleges.

Because neither of our data sources relies on voluntary reporting or tax filing, our data provide a near-complete roster of college attendance at all Title IV accredited institutions of higher edu-

[8]Because we limit the sample to children who can be linked to parents in the U.S. (based on dependent claiming on tax returns), our sample excludes college students from foreign countries. We limit the sample to parents with non-negative income (averaged over five years as described below in Section II.C) because parents with negative income typically have large business losses, which are a proxy for having significant wealth despite the negative reported income. The non-negative income restriction excludes 0.95% of children.

[9]The fraction of children linked to parents drops sharply prior to the 1980 birth cohort because our data begins in 1996 and many children begin to the leave the household starting at age 17 (Chetty et al. 2014). Hence, the 1980 birth cohort is the earliest cohort we analyze.

[10]Information on college attendance is not available in tax records prior to 1999, and the latest complete information on attendance available from the Department of Education at the point of this analysis was for 2013.

cation in the U.S. Aggregate college enrollment counts in our data are well aligned with aggregate enrollments from the Current Population Survey and college-specific enrollment counts from IPEDS (Appendix Table I, Appendix B).[11]

Definition of College Attendance. Our goal is to construct statistics for the set of degree-seeking undergraduate students at each college. Since we cannot directly separate degree seekers from other students (summer school students, extension school students, etc.) in our data, we proceed in two steps in our baseline definition of college attendance. First, we define a student as attending a given college in a given calendar year if she appears in either the 1098-T or NSLDS data. We then assign each student the college she attends for the most years over the four calendar years in which she turns 19, 20, 21, and 22. If a student attends two or more colleges for the same number of years (which occurs for 9% of children), we define the student's college as the first college she attended.[12] Since we do not observe degree completion, students who do not graduate are included in all of the statistics we report.

To evaluate the robustness of our results, we also consider two alternative attendance measures: *age 20 college* (the college a student attends in the calendar year that she turns 20) and *first-attended college* (the college a student attends first between the calendar years in which she turns 19 and 28).

For certain analyses, we report statistics for groups of colleges rather than individual colleges. We classify colleges as "4-year" or "2-year" based on the highest degree they offer using IPEDS data.[13] Following prior work (e.g., Deming et al. 2015), we use data from the Barron's 2009 index (Barron's Educational Series, College Division 2008) to classify 4-year colleges into five tiers based on their selectivity: Ivy-Plus (the Ivy League plus Stanford, MIT, Chicago, and Duke), other elite (Barron's Tier 1 excluding the Ivy-Plus; 65 colleges for the 1980-1991 birth cohorts), highly selective (Barron's Tier 2; 99 colleges), selective (Barron's Tiers 3-5; 1,003 colleges), and non-selective (Barron's Tier 9 and all four-year colleges not included in the Barron's selectivity index; 287 colleges).

[11]Students at some multi-campus systems cannot be assigned to a specific campus and therefore are aggregated into a single cluster. There are 85 such clusters, comprising 17.5% of students and 3.9% of colleges in our data. Separately, 1.8% of student-year observations are assigned to a "colleges with incomplete or insufficient data" category due to incomplete 1098-T data.

[12]If the student attended multiple "most attended" colleges in the first year, which occurs for 1.6% of students, then a college is chosen at random from that set.

[13]Since many colleges offer both 2-year and 4-year programs, many students attending a "4-year" college may be enrolled in a 2-year program.

II.C Incomes

We obtain data on children's and parents' incomes from federal income tax records spanning 1996-2014. We use data from both income tax returns (1040 forms) and third-party information returns (e.g., W-2 forms), which contain information on the earnings of those who do not file tax returns. We measure income in 2015 dollars, adjusting for inflation using the consumer price index (CPI-U).

Parent Income. We measure parent income as total pre-tax income at the household level. In years where a parent files a tax return, we define family income as Adjusted Gross Income (as reported on the 1040 tax return). This income measure includes both labor earnings and capital income. In years where a parent does not file a tax return, we define family income as the sum of wage earnings (reported on form W-2) and unemployment benefits (reported on form 1099-G). In years where parents have no tax return and no information returns, family income is coded as zero. Importantly, the income distribution in the tax data is very similar to that in the American Community Survey (ACS) when one uses the same income definitions (Appendix C, Appendix Table II).

We average parents' family income over the five years when the child is aged 15-19 to smooth transitory fluctuations (Solon 1992) and obtain a measure of resources available at the time when most college attendance decisions are made.[14] We then assign parents income percentiles by ranking them based on this mean income measure relative to all other parents who have children in the same birth cohort.

Child Income. Our primary measure of children's income in adulthood is total pre-tax *individual* earnings. For single filers, individual earnings is defined as the sum of wage earnings and net self-employment income if positive (i.e., net of one-half of the self-employment tax) as reported on Form 1040. For joint filers, it is defined as the sum of the individual's wage earnings reported on his own W-2 forms, the individual's net self-employment income (if positive) reported on Form SE, and half of the additional wage earnings reported on Form 1040 relative to the sum of the spouses' W-2 wage earnings (see Appendix A for details). For non-filers, individual earnings is defined as the sum of wage earnings reported on the individual's W-2 forms.

[14]Following Chetty et al. (2014), we define mean family income as the mother's family income plus the father's family income in each year from 1996 to 2000 divided by 10 (or divided by 5 if we only identify a single parent). For parents who do not change marital status, this is simply mean family income over the 5 year period. For parents who are married initially and then divorce, this measure tracks the mean family incomes of the two divorced parents over time. For parents who are single initially and then get married, this measure tracks individual income prior to marriage and total family income (including the new spouse's income) after marriage. We exclude years in which a parent does not file when computing mean parent income prior to 1999 because information returns are available starting only in 1999.

We measure children's incomes in 2014 – the most recent year in which we observe earnings – to minimize the degree of lifecycle bias that arises from measuring children's earnings at too early an age. We assign children income percentiles by ranking them based on their individual earnings relative to other children in the same birth cohort. We show in Appendix D that the earnings ranks of children in our analysis sample stabilize by 2014.

We also consider two alternative measures of child income in sensitivity analyses: household income, defined in the same way as parents' household income, and household earnings, the sum of individual earnings (defined as above) for the child and his or her spouse. Household income includes capital income, whereas household earnings does not.

II.D Test Scores and Race

We obtained records from the College Board and ACT on standardized college entrance exam scores and race/ethnicity for children in our analysis sample. Our data cover high school graduating cohorts 1996-2004 for SAT and 1995-2007 for ACT.

We focus on individuals' SAT composite score (ranging from 400 to 1600), defined as the mathematics score plus the critical reading score, and the composite ACT score (ranging from 1 to 36). We map ACT scores into equivalent SAT scores using existing concordance tables, we prioritize the SAT if it is available, and we use an individual's maximum composite score if she has taken multiple of the same tests (see Appendix E for details). We use five race/ethnicity categories (referred to hereafter as race): Black, Asian, non-Hispanic white, Hispanic, and other.

SAT/ACT coverage rates (and therefore race coverage rates) are very high at selective colleges where standardized tests are typically required for admission; for instance, we observe a score for 98.5% of Ivy-Plus attendees. We use SAT/ACT scores and race primarily in our counterfactual analysis in Section IV.[15] In that section, we describe and validate a procedure to impute SAT/ACT scores and race for the 26.2% of students for whom we do not observe a test score and race.

II.E Summary Statistics

Table I reports summary statistics for children in our analysis sample. Overall, 62% of the 10.8 million children in 1980-82 birth cohorts attend college at some point between the ages of 19 and 22. Another 12% attend college at some point by age 28; and 27% of children do not attend college

[15] Due to confidentiality restrictions governing the test score data, we are unable to disclose statistics that make use of test score data and/or race data by college and hence cannot report estimates of earnings conditional on test scores, race, or other related measures in this study.

at all before age 28. The median parental household income of children born between 1980-82 is $59,100. The 20th percentile of the parent income distribution is $24,600 and the 80th percentile is $111,100. The children in these cohorts have median individual earnings of $26,900 in 2014 (at ages 32-34). The 20th percentile of the child earnings distribution is $900 and the 80th percentile is $55,800. Approximately 18.5% of children have $0 in individual earnings in 2014. See Table II and Appendix Table III for additional summary statistics.

III Undermatching: SAT/ACT Scores by Parent Income

The relationship between test scores on college entrance exams and parental income is important for understanding the types of policies that could mitigate segregation in higher education. If a large fraction of high-achieving (high-scoring) students come from low- and middle-income families relative to their representation at highly selective colleges, one could potentially reduce segregation at elite colleges by recruiting and admitting high-achieving, low-income applicants at higher rates. If in contrast low-income students have much lower SAT/ACT scores than high-income students, one may require other approaches such as need-affirmative admissions to reduce segregation across colleges.

Several studies in the literature on "undermatching" have analyzed how SAT/ACT scores vary with parental income, but they have reached conflicting conclusions. Some studies (e.g., Carnevale and Strohl 2010, Hoxby and Avery 2013) find that there are many high-achieving, low-income students, but others (e.g, Carnevale and Rose 2004, Hill and Winston 2006, Bastedo and Jaquette 2011) find relatively few such students.

Our data permit a more precise analysis of the degree of undermatching than prior work by combining administrative data on parental income, college attendance, and SAT/ACT scores. However, like many prior studies, we do not observe test scores for a significant share (26.2%) of college students, presumably because they were not required to take a standardized entrance exam by the college they attended. We impute an SAT score to these students using the SAT/ACT score of the college student from the same parent income quintile, state, and college selectivity tier who has the closest level of earnings in adulthood.[16]

This imputation methodology relies on the assumption that the joint distribution of college, parent income quintile, state, and imputed test scores matches what one would observe if all stu-

[16] All students missing a test score are also missing race, since we obtain race information from the SAT/ACT data. We impute race to these students using exactly the same procedure as for test scores.

dents were to take the SAT or ACT. This assumption would be violated if the latent scores of non-SAT/ACT-takers differ systematically from SAT/ACT-takers. We evaluate the validity of this assumption using data from five states where the SAT or ACT is administered to nearly all students—Louisiana, Connecticut, Maine, North Dakota, and Tennessee. We run our imputation algorithm in two ways: as above, but ignoring state in the imputation algorithm, and then separately pretending that we do not observe SAT or ACT scores for anyone in these five states. We then compare the distribution of imputed scores to the distribution of actual scores. Both unconditionally and within each college tier by parent income quintile cell, the quantiles of the imputed SAT distribution match the quantiles of the actual SAT/ACT distribution almost exactly, supporting the validity of the imputation procedure (Appendix Figure II).[17]

Figure Ia plots the parental income distribution of college students in our analysis sample who have an SAT/ACT test score above 1300 (the 93rd percentile of the SAT/ACT distribution). Appendix Table IV shows the full joint distribution of test scores and parent income ranks among all college students. We find that students from low-income families have substantially lower test scores on average and that there are very few high-achieving students from low-income families.[18] For example, 3.7% of college goers with an SAT/ACT score of at least 1300 come from families in the bottom quintile, while 53.7% come from the top quintile. If we limit the sample to the 73.8% of college goers whose test scores are not imputed, we find even fewer high-scoring, low-income students – e.g., a 3.1% bottom-quintile share among those with scores above 1300 – because low-income college goers are less likely to take the SAT or ACT (Appendix Table V). As an additional robustness check, we replicate this analysis using data from the National Postsecondary Student Aid Study, which has student-reported family income data. The NPSAS-based estimate of the bottom-quintile share of 1300+ scorers is 4.0% (Appendix Table VI).

Our estimates of the fraction of high-achieving students who come from low-income families are broadly similar to those reported by Carnevale and Rose (2004), Hill and Winston (2006), and Bastedo and Jaquette (2011), but are substantially smaller than those estimated in the influential study of Hoxby and Avery (2013).[19] Hoxby and Avery estimate that 17% of graduating high school

[17] Furthermore, we find that running our imputation procedure purely using SAT scores (pretending we do not have ACT data) yields very similar results.

[18] One should not infer from this result that SAT/ACT scores simply serve as a proxy for parent income: parental income ranks actually explain only 8.6% of the variance in SAT/ACT scores in our analysis sample. Though students from lower-income families have lower SAT/ACT scores on average, there are many students from middle- and high-income families who do not have high SAT/ACT scores.

[19] Carnevale and Rose use the National Educational Longitudinal Study of 1988 to find that 3% of those with an SAT score or ACT equivalent above 1300 have bottom-SES-quartile parents, where SES is the NELS-provided socioeconomic-status composite of parent income, education, and occupation. Hill and Winston use population-level

seniors with an SAT score or ACT equivalent of at least 1300 have parents in the bottom quartile of the income distribution.[20] By contrast, our estimate of this statistic is 5.0%. Similarly, Hoxby and Avery estimate that 39% of students with SAT/ACT scores above 1300 come from families below the median, compared with 16.6% in our data.

One reason for this discrepancy may be that Hoxby and Avery impute family income using Census tract-level means rather than using individual-level measures, a natural approach given that parental income is frequently missing and potentially noisy in their self-reported data. However, we find that higher-income children *within* small geographies tend to have higher SAT/ACT scores using our individual-level data. As a result, using tract-level means overestimates the number of students from low-income families who have high test scores. A second reason may be that Hoxby and Avery define the 25th percentile of the income distribution based on family income data from the American Community Survey (ACS), but measure parental incomes based on information drawn from financial aid forms. Because of differences in household units and income definitions across these sources, it is possible that Hoxby and Avery's approach would classify more than 25% of parents as falling in the bottom 25% of the distribution.[21] By contrast, because we compute percentile thresholds and measure parental incomes using the same data, 25% of parents fall in the bottom 25% in our analysis by construction.

Having established the relationship between test scores and parental income, we now analyze how alternative allocations of students to colleges would affect income segregation and intergenerational mobility.

SAT and ACT data to find that 4.8% of those with at least a 1300 have bottom-quintile parents, based on student-reported incomes and American Community Survey thresholds. Bastedo and Jaquette report means and standard deviations from the Educational Longitudinal Study of 2002 that, under Normality, imply that 4.1% of those with an SAT score or ACT equivalent above 1300 have bottom-SES-quartile parents.

[20] Hoxby and Avery also require a self-reported grade point average of A- or higher, but they note that the GPA restriction matters very little once they apply the SAT/ACT restriction.

[21] Hoxby and Avery classify a child as falling in the bottom quartile if the child's estimated family income lies below $41,472, the 25th percentile of family income in the 2008 American Community Survey. The income data they use in their analysis is based on College Scholarship Service (CSS) Profile family income data reported by the student, which in turn comes from parents' tax returns and supplementary information. In the tax data, however, the 25th percentile of the Adjusted Gross Income distribution is about $25,000, well below the ACS threshold. In Appendix C, we show that the differences between the tax data and the ACS are entirely due to differences in the definition of household units and incomes.

IV How Would Changes in the Allocation of Students to Colleges Affect Segregation and Intergenerational Mobility?

In this section, we use our estimates to simulate how income segregation across colleges and intergenerational mobility would change if students were allocated to colleges differently, using data on SAT and ACT scores as a proxy for students' academic qualifications at the point of application. To do so, we simulate how alternative allocations of students to colleges would change the degree of income segregation across colleges and the rate of intergenerational mobility in the economy.

The reallocations we propose hold constant total national spending on higher education, since we hold the number of seats at each college fixed. However, they would require a change in the allocation of funding across families and colleges, as some colleges would have larger shares of low-income students and thus have lower net tuition revenue given the financial aid packages they currently offer. Hence, the counterfactual allocations we simulate below should not be thought of as policy proposals, but rather as benchmarks that shed light on the drivers of segregation across colleges and the potential impacts of changing which students attend which colleges on economic mobility.

IV.A Income Segregation

We begin by evaluating the extent to which income segregation across colleges can be explained by differences in academic credentials when students apply to college (as proxied for by SAT or ACT scores), holding *fixed* each college's current racial composition and the geographic origins of their students. We impose the geographic and racial constraints to better approximate feasible reallocations, recognizing that many institutions (e.g., public state institutions, local community colleges, or Historically Black Colleges and Universities) effectively face geographic or racial constraints in practice.[22] This analysis provides a natural benchmark to gauge the extent to which colleges' student bodies are representative of the underlying population of academically qualified students from which they seek to draw. For example, are the parental incomes of Ivy League students representative of all students with similar test scores who come from the same states and racial groups?

To conduct this analysis, we first record the actual vector of SAT/ACT test scores at each college-by-state-by-race group $\overrightarrow{s}_g$. We then allocate students by filling each college-state-race's

[22]The impacts of our counterfactuals on aggregate segregation and mobility actually turn out to be quite similar if we permit reallocations without any racial or geographic constraints (Appendix Table IX).

slot for a student with test score s with a *random* draw from the state-race's population of college students with test score s. In this "income-neutral" student allocation regime, colleges continue to enroll students based on both test scores and other credentials (e.g., recommendations, extracurriculars), but eliminate variation in enrollment rates by parental income – whether due to differences in application, admissions, or matriculation – among students with comparable test scores in the same state and racial group.

Figure IIa shows how segregation across colleges would change under this counterfactual. The left side of the figure examines the extent to which students from low-income families are exposed to students from high-income families by plotting the fraction of college peers from the top quintile among college students with parents in the bottom quintile. The right side analogously examines segregation among high-income students by plotting the fraction of top-quintile peers for students from the top quintile (see Appendix Table VIII for additional statistics). In each case, we plot three statistics: the actual rates in the data, the rates under the income-neutral allocation counterfactual, and the rates under need-affirmative student allocations (which we discuss below).

Segregation across colleges would fall substantially if college enrollment were income neutral conditional on test scores: for example, the top-quintile peer share of students from low-income families would rise from 22.5% to 27.8%. Since 30.8% of college students come from the top quintile (shown by the horizontal line on the figure), a random allocation of students to colleges among the current pool of college students would yield a top-quintile peer share of 30.8%. Hence, income-neutral allocations would close 63.9% of the gap between the current degree of exposure that students from low-income families have to high-income students and the exposure they would have if colleges were perfectly integrated by income (conditional on the set of students who currently attend college). Put differently, only 36.1% of the income segregation across colleges can be attributed to differences in students' test scores, racial backgrounds, or geographic origins. The remaining 63.9% is driven by a combination of differences in student application choices, college admissions, and matriculation decisions by parental income conditional on these factors.

Although the income-neutral allocation reduces segregation overall, it largely reshuffles students within selectivity tiers and thus has smaller impacts on parental income distributions at more selective colleges. Figure IIb illustrates this result by plotting the fraction of students from the bottom parental income quintile at Ivy-Plus, selective colleges (top six tiers), and unselective colleges (bottom six tiers) in actuality and under the counterfactual (see Table III for statistics for each tier separately). The bottom-quintile share of students at selective colleges overall rises from 7.3% to

8.6%, closing 38% of the gap in their underrepresentation relative to their 10.7% share of college goers overall. This 38% reduction in segregation at selective colleges is substantial, but it is much smaller than the 64% reduction overall.

Impacts at Ivy-Plus Colleges. The impacts of income-neutral allocations at the most selective colleges differ from those in the broader population. At Ivy-Plus colleges, the fraction of students from the bottom quintile remains essentially unchanged under income-neutral allocations in absolute terms (rising from 3.8% to 4.4%), but the fraction of students from the middle class (the second, third, and fourth income quintiles) rises sharply, from 27.8% to 37.9%, as shown in Table III. Figure Ia shows why we see the biggest impacts on the representation of the middle class by plotting the parental income distribution of high SAT/ACT (>=1300) scorers alongside the parental income distribution of actual Ivy-Plus enrollees. Children from the bottom-quintile are represented at nearly the same rate as one would expect given their test scores; children from the middle-class are under-represented at these colleges; and those from the top quintile are over-represented.

Figure Ib presents a more granular depiction of the degree of over/under-representation by parental income. It plots the share of students with an SAT/ACT score of 1400 – the modal and median test score among actual Ivy-Plus students – who attend an Ivy-Plus college. Rather than a flat line, which would have indicated that 1400-scorers from each parent income bin attend an Ivy-Plus college at the same rate, we observe an asymmetric U-shape, with higher attendance rates in the tails. In particular, 1400-scorers with parents from the top and bottom quintiles attend Ivy-Plus colleges at 2.4 and 1.6 times the rate of middle-quintile children with comparable test scores, respectively. We find similar patterns at other test score levels; see Appendix Table VII.

The upshot of this analysis is that there is a "missing middle" at Ivy-Plus institutions – an under-representation of students with high test scores from middle class families relative to students from low-income and especially high-income families. Changes in application or admissions policies that eliminate existing differences in attendance rates conditional on test scores across parental income groups could therefore significantly increase the representation of the middle class (though not low-income) families at the nation's most selective private colleges.[23]

[23]This conclusion differs from that of Carnevale et al. (2019), who report that high-socioeconomic-status (a composite of parent income, education, and occupation prestige) shares at highly selective colleges would barely change under a system in which students with the highest test scores are admitted to the most selective colleges, without regard to other credentials. This is because the students with the very highest SAT/ACT scores tend to come from the highest-income families. Although Carnevale et al.'s pure test-score-admissions counterfactual also achieves income neutrality conditional on test scores, it increases the selectivity of elite colleges, because elite colleges currently admit many students who have SAT scores well below 1600. Our point is that shifting to a system that is income-neutral conditional on the *current* distribution of test scores at elite colleges (thereby preserving current levels of selectivity) would substantially reduce top income shares.

Of course, test scores are an imperfect proxy for academic credentials, and colleges weigh many factors (e.g., extracurriculars, overall fit) beyond academic qualifications in admissions decisions. Therefore, one cannot interpret the counterfactual estimates as representing income segregation under a "meritocracy." Nevertheless, we view this counterfactual as a natural benchmark to gauge the extent to which student bodies are representative of the underlying population of academically qualified students. If one's objective is to have income-neutral enrollment conditional on merit, deviations from this benchmark can be justified at current selectivity levels only if other non-test-score determinants of merit are correlated with parent income.[24]

Need-Affirmative Student Allocations. Although a system of applications and admissions that is income neutral conditional on academic credentials would reduce income segregation significantly, the fraction of students from the bottom income quintile would remain about 50 percent higher at unselective colleges than selective colleges. We therefore now turn to ask how much of a preference one would need to give children from lower-income backgrounds in the student allocation process – or, equivalently, how much lower-income students' test scores would have to rise – to fully eliminate segregation across colleges.

We simulate need-affirmative student allocations by adding Δs_q points to the SAT/ACT scores of children with parents from income quintile $q < 5$. We vary the values of $\{\Delta s_q\}$, leaving SAT/ACT scores for children from the top quintile unchanged ($\Delta s_5 = 0$), in order to identify a profile of test-score increases that results in a constant parental income distribution across all college selectivity tiers. We then re-norm test scores to match the actual distribution and replicate the income-neutral allocation above with these adjusted scores (see Appendix F for details).

Iterating over linearly-declining profiles of $\{\Delta s_q\}$, we find that that adding 160 SAT points for those from the bottom quintile ($\Delta s_1 = 160$) and $\Delta s_q = (1 - \frac{q-1}{5})160$ for $q = 2, 3, 4$ – i.e., increments of 80%, 60%, and 40% of the bottom-quintile increment – produces roughly equal parental income shares across tiers.[25] To understand the practical implications of such an increment, note that 7.3% of children from the bottom parental income quintile with an SAT score of 1400 attend an

[24]It may be useful to consider an analogy to the principle of "disparate impact" in anti-discrimination law. Any hiring practice (e.g., requiring candidates to excel at squash) that has a disparate (differential) impact by gender or race is prima facie evidence of unlawful discrimination and shifts the burden of proof to the employer to show that the practice is consistent with business necessity and has no practical and more neutral alternative. Disparate impact by parental income is not a legal concern, but would be of analogous interest to those seeking a system of college admissions that is income-neutral conditional on merit.

[25]That is, the following groups are treated identically within state-race groups: (s+160)-scorers with bottom-20% parents, (s+128)-scorers with second-quintile parents, (s+96)-scorers with middle-quintile parents, (s+64)-scorers with fourth-quintile parents, and s-scorers with top-quintile parents. Changes in admission probabilities can change applicant pools (e.g., Yagan 2016); our linear gradient reflects the combined effect of application, admission, and matriculation.

Ivy-plus college in our data. Such students would attend Ivy-plus schools at a rate of 25.8% in our need-affirmative 160 point SAT increment scenario. More generally, among students with SAT scores above 1300, the 160 point increment increases the likelihood of attending an Ivy-plus college for a bottom-income-quintile student conditional on their SAT score by a factor of 3.54 on average.

It is instructive to gauge the magnitude of these increments in SAT scores and attendance rates for low-income students by comparing them to admissions preferences currently granted to other groups. Espenshade, Chung and Walling (2004) use admissions data from three elite private colleges to evaluate the extent to which legacies, athletes, and underrepresented minorities are more likely to be admitted, controlling for their credentials at the point of application. They find that the increase in admissions probability for these groups is roughly equivalent to the effect of a 160 point increase in SAT scores.[26] Similarly, Arcidiacono, Kinsler and Ransom (2019) use data from Harvard to estimate that students who are recruited athletes, legacies, those on the Dean's interest list, or children of faculty and staff (ALDCs) have admissions rates 3.4 times higher than non-ALDC students with otherwise similar characteristics.[27] Hence, one way to implement our need-affirmative counterfactual could be to grant a preference in admissions for lower-income students similar to that currently given to other groups. Another approach may be to increase application or matriculation rates for lower-income students relative to high-income students by an equivalent amount.

Figure IIa shows that this degree of need-affirmative student reallocation essentially desegregates the higher education system fully, with exposure rates to students from different income groups similar to what one would obtain under a random allocation benchmark.[28] Moreover, need-affirmative allocations would essentially eliminate differences in parental income distributions across *all* selectivity tiers. The fraction of students from bottom-quintile families is close to the overall mean across all colleges of 10.7% in all college tiers (Figure IIb, Table IIIc). Indeed, the Ivy-Plus colleges would have a *higher* fraction of children from low-income families than almost all other tiers in this scenario.[29] Each tier still has more students from high-income families than low-income

[26] More precisely, Espenshade, Chung and Walling estimate that legacy status is equivalent to 160 SAT/ACT points, recruited athlete status 200 points, African-American status 230 points, and Hispanic status 185 points. Hurwitz (2011) also finds large observed admissions advantages for legacy applicants.

[27] Table 10 of Arcidiacono, Kinsler and Ransom (2019) reports counterfactual admissions rates for admitted ALDC students, removing the ALDC preferences, separately for students of each race. Averaging these counterfactual admissions rates across racial groups using the number of admitted ALDCs from each race (reported in the same table) yields 29.4%, implying admissions rather that are 1 / 29.4% = 3.4 times higher for ALDCs than otherwise similar non-ALDCs.

[28] We present results with alternative increments to SAT/ACT scores in Appendix Figure IV.

[29] Bowen, Kurzweil and Tobin (2006, Chapter 7) also examine the effects of need-affirmative allocations on parental income distributions at 18 elite colleges. Our findings are qualitatively consistent with their results at these 18

families even with need-affirmative allocations because college attendance rates rise sharply with parental income (Chetty et al. 2017) and our counterfactual does not change who attends college. However, among the current pool of college students, treating those from low-income families like legacy students would make parental income distributions similar across all colleges.

IV.B Intergenerational Mobility

Estimating Colleges' Value-Added. To quantify how changes in the allocation of students to colleges would affect intergenerational mobility, we first need estimates of how children's earnings outcomes would change if they were to attend different colleges (i.e., colleges' causal effects or "value-added"). Directly estimating each college's value-added would require a source of quasi-experimental variation at each college and is outside the scope of this paper. Instead, we build on the prior literature and use estimates that are consistent with that work as an input into our simulations.

We begin from our estimates of children's mean earning ranks conditional on their parental income, race, and SAT/ACT scores estimated above.[30] We then estimate the fraction λ of these conditional earnings differences across colleges that is due to causal effects vs. selection by controlling for observable characteristics and for the set of colleges to which a student applied to capture selection on unobservables, following Dale and Krueger (2002).

Formally, consider the regression model

$$y_{iqc} = \alpha + \beta X_{iqc} + f(S_i) + f(p_q) + \theta_r + \delta_c + \varepsilon_{iqc} \tag{1}$$

where y_{iqc} is the earnings rank of student i from parent income rank q who attended college c; X_{iqc} is a vector of observed student-specific characteristics; $f(S_i)$ is a quintic in the student's SAT or ACT equivalent score, an indicator for taking the SAT, and an indicator for taking the ACT (note that some students took both tests); $f(p_q)$ is a quintic in the student's parent income rank; θ_r is a race fixed effect, and δ_c is a college fixed effect. We can estimate the vector of college fixed effects $\Delta_c = \{\delta_c\}$ using a variety of control vectors X_{iqc}. First consider estimates where X_{iqc} is empty and thus the only controls are SAT/ACT scores, parent income, and race; denote these estimates by $\Delta_c^{S,p,r}$. We can then assess the relationship between these test-score-and-parent-income-and-race controlled estimates of colleges' effects with estimates that include additional controls by running

colleges, although our quantitative results differ because their self-reported parent income measures yield low-income shares at elite colleges that are twice as large as ours.

[30] We do not condition on children's pre-college state because of small samples; in particular, under need-affirmative allocations, cells can be small when counterfactually high or low SAT/ACT scorers are assigned to a given college.

the regression

$$\Delta_c^{S,p,r,X} = \alpha + \lambda \Delta_c^{S,p,r} + \nu_c. \tag{2}$$

The parameter λ gives an estimate of the fraction of the baseline test-score-and-parent-income-and-race-controlled difference between any two colleges that would remain, on average, with the addition of further controls. If latent student quality is not correlated with the college he or she attends conditional on the observed characteristics X, the parameter λ can be interpreted as the fraction of the differences between colleges' earnings estimates $\Delta_c^{S,p,r}$ that reflects their causal effects (value-added).

Table IV reports estimates of λ using a range of control vectors X.[31] Columns 1-3 control successively for the following observable student characteristics: interactions between gender, race, and the test score quintic; high school fixed effects; and high-school fixed effects interacted with race. These specifications all yield estimates of $\lambda > 0.9$, i.e. more that 90% of the baseline earnings variation (conditional on parental income, race, and test scores) reflects a causal effect if these observables capture selection.

To assess whether selection on other unobservable dimensions might confound our estimates, we use the set of colleges to which students apply as controls for their latent ability, as in Dale and Krueger (2002, 2014).[32] In Column 4 of Table IV, we follow Dale and Krueger (2014) and control for mean SAT score of the colleges to which students send their SAT/ACT scores (a proxy for college application) and the total number of colleges to which they send their scores in addition to the observable characteristics used in Column 2. Column 5 adds high-school fixed effects interacted with race to Column 4, while Column 6 limits the sample to students in the bottom quintile of the income distribution.[33] These specifications all yield point estimates of $\lambda \geq 0.85$, with a lower bound on the 95% confidence interval of around 0.82.[34]

[31] We exclude students who do not attend any college and omit students with imputed test scores from these regressions.

[32] Controlling for the set of colleges to which students apply is what Dale and Krueger (2002) call a "self-revelation" approach to adjusting for selection; they show that this approach yields estimates that are very similar to specifications that control for the set of colleges to which students are *admitted*. Dale and Krueger (2014) simply control for the application set rather than the admittance set to maximize power in light of this result, and we follow that approach here (since we do not have data on admissions).

[33] As the estimate in Column 6 indicates, we do not find significant heterogeneity in λ across parental income groups. However, the baseline conditional earnings differences from attending a more selective college are larger for students from low-income families. In particular, we replicate Dale and Krueger's result that the return to attending a college with higher average SAT scores is small on average, but larger for low-income students in Appendix Table X.

[34] In their College and Beyond sample, Dale and Krueger find that controlling for the application set reduces the coefficient on mean SAT scores substantially even after controlling for student's own SAT scores and other observables. We believe our findings differ because we have more precise controls for student background (e.g., a precise measure of

Given these estimates, we assume that $\lambda = 80\%$ of the conditional earnings differences observed across colleges are due to causal effects (value-added) and the remaining 20% is due to selection in our baseline simulations.[35] Importantly, we also assume that these estimated causal effects do not change under our counterfactual student reallocations, in particular ignoring potential changes in value-added that may arise from having a different group of students (peer effects).

Income-Neutral Student Allocations. We construct a counterfactual earnings distribution for children at each college based on the observed distribution of earnings for children in each parent income quintile, SAT/ACT score level, race, and college. Mechanically, children are randomly assigned the earnings of another child who is observed as attending their counterfactually assigned college and who has the same parent income quintile, race, and SAT/ACT score with 80% probability and are assigned their actual earnings with 20% probability (reflecting our 80% causal effect assumption). Because this reallocation changes the aggregate distribution of children's earnings in adulthood, we then recompute quintile earnings thresholds based on the new aggregate earnings distribution when computing mobility rates.[36]

Table V shows how the intergenerational transition matrix for college students would change under this counterfactual. Panel A shows the actual transition matrix. For example, the chance of reaching the top earnings quintile ranges from 18.2% for children with parents in the bottom quintile to 40.2% for children with parents in the top quintile, as shown in the fifth column of Table Va. This difference of 22 percentage points is plotted in the first bar in Figure IIc.

The second bar of Figure IIc shows how this gap would change under the income-neutral allocation counterfactual. The chance of reaching the top quintile now ranges from 19.5% to 38.3% across parent income quintiles (Table Vb), a gap that is 14.6% smaller than the empirically observed gap. The gap in children's chances of reaching the top 1% between children from low-income and high-income families falls from 2.8pp to 2.3pp, a similar reduction in percentage terms (Table V). Likewise, the correlation between parents' and children's income ranks among college students falls by 15% under the counterfactual. In sum, the intergenerational persistence of income would

parental income rather than a proxy) and because students' own SAT scores may be a stronger predictor of outcomes today than for students who attended college in the 1970-80s.

[35] To further validate this approach, we compare our estimates to the regression discontinuity estimates of Zimmerman (2014), who essentially estimates the causal effect of attending Florida International University vs. Miami Dade College. Our estimates based on the approach outlined above are similar to Zimmerman's quasi-experimental estimates.

[36] We take non-college-goers earnings as fixed, ignoring the possibility of equilibrium effects on their earnings. We obtain nearly identical results if we do not recompute the thresholds.

fall by about 15% if students were allocated to colleges based purely on their qualifications at the point of application (as proxied for by SAT/ACT scores).

Need-Affirmative Student Allocations. To compute students' earnings distributions under need-affirmative allocations, we follow the same approach as above, using students' *actual* SAT/ACT scores (rather than their adjusted SAT/ACT scores) in the earnings rank reallocation. This approach means that the test score increment granted in the admissions process does not affect students' earnings outcomes aside from the college that they attend.

Under need-affirmative allocations, the chance of reaching the top quintile ranges from 20.8% to 37.0% across parent income quintiles (Table Vc), 26.5% smaller than the empirically observed gap (Figure IIc). The correlation between parents' and children's income ranks falls by 25%. The gap in children's chances of reaching the top 1% between children from low-income and high-income families falls from 2.8pp to 1.9pp, a 32.6% reduction. The impact on children's chances of reaching the upper tail is particularly large because need-affirmative allocations sharply change the distribution of parental incomes at the most selective private colleges, whose students are especially likely to reach the upper tail, as shown in Chetty et al. (2017).

Need-affirmative reallocation has nearly twice as large an effect on mobility rates as income-neutral reallocation because it enables low-income students to attend the highest value-added colleges. The value-added of the colleges that students from low- vs. high-income families attend is essentially equalized under need-affirmative allocations. The difference in the value-added of the colleges attended by students from the top vs. bottom parent income quintile (estimated as described above) falls by 89% relative to the empirically observed difference of 4.5 percentiles. By contrast, income-neutral allocations reduce the gap in college value-added by parental income much less, by only 47% relative to the empirically observed difference. Intuitively, this is because income-neutral allocations tend to reshuffle low-income students across colleges in the same tier as shown above, whereas need-affirmative allocations enable low-income students to get into higher value-added colleges in higher selectivity tiers.

Alternative Assumptions About Causal Effects. In Appendix Table XI, we vary our assumption about the fraction of the difference in earnings across colleges conditional on parental income, race, and SAT/ACT scores that is due to causal effects from $\theta = 100\%$ (pure causal effects, no selection) to $\theta = 0\%$ (pure selection, no causal effects). At the upper bound ($\theta = 100\%$), need-affirmative allocations would reduce the intergenerational persistence of income by 33%. The simulated impact mechanically decreases to 0% at the lower bound of $\theta = 0\%$. Assuming that $\theta > 50\%$ – roughly the lower bound of the 95% confidence interval implied by comparing Zimmerman's (2014) estimate to ours – one could reduce the intergenerational persistence of income by at least 17% (among children who attend college) purely by changing the allocation of students to colleges, without attempting to change any college's production function.[37] These are substantial effects given that gaps in intergenerational mobility emerge from an accumulation of exposure to different environments and schools throughout childhood (Chetty and Hendren 2018). Since colleges account for less than a quarter of the time most children spend in formal education, one would not expect impacts on mobility much larger than 25% purely from changes in higher education.

V Conclusion

In this paper, we showed that allocating students to colleges in an income-neutral manner conditional on their SAT/ACT scores would increase the representation of students from low- and middle-income families at selective colleges substantially, holding fixed the racial composition and geographic origins of their students. At the most selective (Ivy-Plus) colleges, the fraction of students from the middle class would rise substantially, although there would be little absolute change in the fraction of students from the bottom income quintile because so few of them currently have sufficiently high SAT/ACT scores. Under the assumption that 80% of the difference in earnings premia (conditional on parental income, race, and state) are causal, our simulations imply that income-neutral allocations of students to colleges (conditional on test scores) would itself reduce the intergenerational persistence of income by 15%.

To go further, we simulate the consequences of raising lower-income students' test scores or granting them a preference in the application and admissions process similar to that currently given to legacy or minority students at elite private colleges. Such a change would essentially eliminate income segregation across all college tiers and reduce the intergenerational persistence of income by about 25%. We conclude that feasible changes in the allocation of students to colleges could increase intergenerational mobility substantially without any changes to existing educational programs, suggesting value in further efforts to enable students from low- and middle-income families to attend colleges that offer better earnings outcomes.

[37] An alternative possibility is that the ratio of selection effects vs. causal effects is heterogeneous by parent income, with larger causal effects of attending an elite college for children from lower-income families. In Appendix Table XII, we consider a scenario in which causal effects are 0 for reallocations within selective colleges (the top six tiers) for students with parents in the top four quintiles, 40% for reallocations within selective colleges for students with parents in the bottom quintile, and 80% for all other reallocations. In this scenario, need-affirmative allocations would reduce the intergenerational persistence of income by 21.3%.

References

Andrews, R., S. Imberman, and M. Lovenheim (2016). Recruiting and Supporting Low-Income, High-Achieving Students at Flagship Universities. *NBER Working Paper No. 22260*.

Arcidiacono, P. (2005). Affirmative Action in Higher Education: How do Admission and Financial Aid Rules Affect Future Earnings? *Econometrica 73*(5), 1477–1524.

Arcidiacono, P., J. Kinsler, and T. Ransom (2019). Legacy and Athlete Preferences at Harvard. *NBER Working Paper No. 26316*.

Avery, C., C. Hoxby, C. Jackson, K. Burek, G. Pope, and M. Raman (2006). Cost Should Be No Barrier: An Evaluation of the First Year of Harvard's Financial Aid Initiative. *NBER Working Paper No. 12029*.

Barron's Educational Series, College Division (Ed.) (2008). *Barron's Profiles of American Colleges 2009*. Hauppage, NY: Barron's Educational Series, Inc.

Bastedo, M. N. and O. Jaquette (2011). Running in Place: Low-Income Students and the Dynamics of Higher Education Stratification. *Educational Evaluation and Policy Analysis 33*(3), 318–339.

Black, D. and J. A. Smith (2004). How Robust is the Evidence on the Effects of College Quality? Evidence from Matching. *Journal of Econometrics 121*(1-2), 99–124.

Bowen, W. G., M. M. Chingos, and M. S. McPherson (2009, September). *Crossing the Finish Line: Completing College at America's Public Universities*. Princeton University Press. Google-Books-ID: LIdcVVxZb6gC.

Bowen, W. G., M. A. Kurzweil, and E. M. Tobin (2006). *Equity and Excellence in American Higher Education*. University of Virginia Press.

Carnevale, A. P. and S. J. Rose (2004). *Socioeconomic Status, Race/Ethnicity, and Selective College Admissions*, pp. 101–156. New York: The Century Foundation Press.

Carnevale, A. P. and J. Strohl (2010). How Increasing College Access is Increasing Inequality, and What to do about It. In R. D. Kahlenburg (Ed.), *Rewarding Strivers: Helping Low-Income Students Succeed in College*, Chapter 3, pp. 71–190. New York: The Century Foundation Press.

Carnevale, A. P., J. Strohl, M. Van Der Werf, M. C. Quinn, and K. P. Campbell (2019). SAT-Only Admission: How Would It Change College Campuses. *Georgetown University Center on Education and the Workforce Working Paper*.

Cellini, S. R. and N. Turner (2019). Gainfully Employed? Assessing the Employment and Earnings of For-Profit College Students using Administrative Data. *Journal of Human Resources 54*(2), 342–370.

Chetty, R., J. N. Friedman, N. Turner, and D. Yagan (2017). Mobility Report Cards: The Role of Colleges in Intergenerational Mobility. *NBER Working Paper No. 23618*.

Chetty, R. and N. Hendren (2018). The Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects. *The Quarterly Journal of Economics 133*(3), 1107–1162.

Chetty, R., N. Hendren, P. Kline, and E. Saez (2014). Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States. *Quarterly Journal of Economics 129*(4), 1553–1623.

Cilke, J. (1998). A Profile of Non-Filers. U.S. Department of the Treasury, Office of Tax Analysis Working Paper No. 78.

Dale, S. B. and A. B. Krueger (2002). Estimating the Payoff to Attending a More Selective College. *Quarterly Journal of Economics 117*(4), 1491–1527.

Dale, S. B. and A. B. Krueger (2014). Estimating the Effects of College Characteristics over the Career using Administrative Earnings Data. *Journal of human resources 49*(2), 323–358.

Deming, D. and S. Dynarski (2010). *Into College, Out of Poverty? Policies to Increase the Postsecondary Attainment of the Poor*, pp. 283–302. University of Chicago Press.

Deming, D. J., C. Goldin, L. F. Katz, and N. Yuchtman (2015). Can Online Learning Bend the Higher Education Cost Curve? *American Economic Review Papers and Proceedings 105*(5), 496–501.

Epple, D., R. Romano, and H. Sieg (2006). Admission, Tuition, and Financial Aid Policies in the Market for Higher Education. *Econometrica 74*(4), 885–928.

Espenshade, T. J., C. Y. Chung, and J. L. Walling (2004). Admission Preferences for Minority Students, Athletes, and Legacies at Elite Universities. *Social Science Quarterly 85*(5), 1422–1446.

Fryer, R. G. and S. D. Levitt (2013). Testing for Racial Differences in the Mental Ability of Young Children. *American Economic Review 103*(2), 981–1005.

Goodman, J. (2008). Who Merits Financial Aid? Massachusetts' Adams Scholarship. *Journal of Public Economics 92*(10), 2121–2131.

Haider, S. and G. Solon (2006). Life-Cycle Variation in the Association between Current and Lifetime Earnings. *American Economic Review 96*(4), 1308–1320.

Hastings, J., C. Neilson, and S. Zimmerman (2013). Are Some Degrees Worth More than Others? Evidence from College Admission Cutoffs in Chile. *NBER Working Paper No. 19241*.

Heckman, J. J. and A. B. Krueger (2005). Inequality in America: What Role for Human Capital Policies? *MIT Press Books 1*.

Hill, C. B. and G. C. Winston (2006). How Scarce are High-Ability Low-Income Students? In M. S. McPherson and M. O. Schapiro (Eds.), *College Access: Opportunity or Privilege?*

Hoekstra, M. (2009). The Effect of Attending the Flagship State University on Earnings: a Discontinuity-Based Approach. *Review of Economics and Statistics 91*(4), 717–724.

Howell, J. S. (2010). Assessing the Impact of Eliminating Affirmative Action in Higher Education. *Journal of Labor Economics 28*(1), 113–166.

Hoxby, C. and C. Avery (2013). The Missing "One-Offs": The Hidden Supply of High-Achieving, Low-Income Students. *Brookings Papers on Economic Activity 2013*(1), 1–65.

Hoxby, C. and S. Turner (2013). Expanding College Opportunities for High-Achieving, Low-Income Students. *Stanford Institute for Economic Policy Research Discussion Paper No. 12-014*.

Hoxby, C. M. (2015). Computing the Value-Added of American Postsecondary Institutions. *Internal Revenue Service Statistics of Income Working Paper*.

Hurwitz, M. (2011). The Impact of Legacy Status on Undergraduate Admissions at Elite Colleges and Universities. *Economics of Education Review 30*(3), 480–492.

James, E., N. Alsalam, J. C. Conaty, and D.-L. To (1989). College Quality and Future Earnings: Where should you Send your Child to College? *American Economic Review 79*(2), 247–252.

Kirkeboen, L. J., E. Leuven, and M. Mogstad (2016). Field of Study, Earnings, and Self-Selection. *Quarterly Journal of Economics 131*(3), 1057–1111.

Krueger, A., J. Rothstein, and S. Turner (2006). Race, Income, and College in 25 years: Evaluating Justice O'Connor's conjecture. *American Law and Economics Review 8*(2), 282–311.

Kurlaender, M. and K. Cohen (2019). Predicting College Success: How Do Different High School Assessments Measure Up?. *Policy Analysis for California Education, PACE*.

Manoli, D. and N. Turner (2018). Cash-on-Hand and College Enrollment: Evidence from Population Tax Data and the Earned Income Tax Credit. *American Economic Journal: Economic Policy 10*(2), 242–71.

Marx, B. and L. Turner (2015). Borrowing Trouble? Student Loans, the Cost of Borrowing, and Implications for the Effectiveness of Need-Based Grant Aid. *NBER Working Paper No. 20850*.

Rampell, C. (2019). Colleges have Been under Pressure to Admit Needier Kids. It's Backfiring. *The Washington Post*.

Reardon, S. F. (2019). Educational Opportunity in Early and Middle Childhood: Using Full Population Administrative Data to Study Variation by Place and Age. *RSF: The Russell Sage Foundation Journal of the Social Sciences 5*(2), 40–68.

Rothstein, J. M. (2004). College Performance Predictions and the SAT. *Journal of Econometrics 121*(1-2), 297–317.

Sackett, P. R., N. R. Kuncel, A. S. Beatty, J. L. Rigdon, W. Shen, and T. B. Kiger (2012). The Role of Socioeconomic Status in SAT-Grade Relationships and in College Admissions Decisions. *Psychological science 23*(9), 1000–1007.

Solon, G. (1992). Intergenerational Income Mobility in the United States. *American Economic Review 82*(3), 393–408.

Solon, G. (1999). Intergenerational Mobility in the Labor Market. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics*, Volume 3, pp. 1761–1800.

Summit, E. G. (2016). ACT-SAT concordance tables. *https://mytutor.com/wp-content/uploads/2018/06/sat-act-concordance-chart-2016-v-2018-rev.pdf*.

Todd, R. (1976). The Missing Middle. Accessed February 2, 2020. https://www.nytimes.com/1976/02/01/archives/the-missing-middle-on-campus-students-are-clamoring-to-go-to.html.

Yagan, D. (2016). Supply vs. Demand under an Affirmative Action Ban: Estimates from UC Law Schools. *Journal of Public Economics 137*, 38–50.

Zimmerman, S. D. (2014). The Returns to College Admission for Academically Marginal Students. *Journal of Labor Economics 32*(4), 711–754.

Appendices

A. Sample Construction and Income Definitions

Sample Definition. Our primary sample is very similar to the "extended sample" analyzed in Chetty et al. 2014, and much of this appendix is therefore taken directly from Chetty et al. (2014, Online Appendix A).

We begin with the universe of individuals in the Death Master (also known as the Data Master-1) file produced by the Social Security Administration and housed alongside tax records. This file includes information on year of birth and gender for all persons in the United States with a Social Security Number or Individual Taxpayer Identification Number (ITIN).[38] To construct our sample of children, we begin from the set of individuals born in the 1980-1991 cohorts. We measure parent and child income, college attendance, and all other variables using data from the IRS Databank, a balanced panel covering all individuals in the Death Master file who were not deceased as of 1996.

For each child, we define the parent(s) as the person(s) who claim the child as a dependent on a 1040 tax form in the year the child turns 17. Note that the parent(s) of the child are not necessarily biological parents, as it is possible for custodians (regardless of family status) to claim a child if the child resides with them.[39] If parents are married but filing separately, we assign the child both parents. We identify children's parents at age 17 because our goal is to measure the economic resources of the child's family around the time he or she attends college. We do not match children to parents at later ages (e.g., 18 or 19) because many children leave home after age 17 (at differential rates across income groups), creating scope for selection bias.

If the child is not claimed at age 17 on any 1040 tax form, we go back one year (to the year in which the child turns 16) to identify parents. We repeat this process until we find a year when the child is claimed, up to the year in which the child turns 12. Since the tax data start in 1996, for the 1980 cohort, we only match children up to age 16; for the 1981 cohort, up to age 15, etc. In short, we use up to 6 years (from ages 12-17) to find a parental match. If no such parental match is found, then the child's record is discarded.[40]

Importantly, once we match a child to parent(s), we hold this definition of parents fixed regardless of changes in parents' marital status or who claims the child in other years. For example, a child matched to married parents at age 17 but who had a single parent at age 16 is always matched to the two married parents at age 17. Conversely, a child matched to a single parent at age 17 who had married parents at age 16 will be considered matched to a single parent.

Finally, we discard the small set of children whose parents have negative family income (as defined in Section II.A) on average over the 5 year time window when they are aged 15-19. Negative income is generally due to business losses and denotes high potential earnings ability so that ranking such parents at the very bottom is actually misleading.

[38] ITINs are issued by the IRS to individuals who do not have a social security number, for example because they are undocumented immigrants.

[39] Children can be claimed as a dependent only if they are aged less than 19 at the end of the year (less than 24 if enrolled as a student) or are disabled. A dependent child is a biological child, step child, adopted child, foster child, brother or sister, or a descendant of one of these (for example, a grandchild or nephew). Children must be claimed by their custodial parent, i.e. the parent with whom they live for over half the year. Furthermore, the custodial parent must provide more than 50% of the support to the child. Hence, working children who provide more than 50% of their own support cannot be claimed as dependents. See IRS Publication 501 for further details.

[40] Very few children are unclaimed on tax returns (Chetty et al. 2014) because claiming children yields substantial refundable tax credits. Therefore, the children we exclude are almost all non US-residents when they were aged 12-17.

Details on Income Definitions. As discussed in Section II, in our baseline analysis, we measure children's earnings as the sum of individual wage income and net self-employment income (if positive) for year 2014. Here we provide further details regarding those definitions, which are more complex than our parent income definitions because we must apportion total income reported on the tax return across individuals to measure income at the individual level.

For a child who is a non-filer (neither a primary nor a secondary filer on any 1040 return), individual earnings are defined simply as the sum of wage income from the individual's own W-2 forms. For a child who is a single filer, individual earnings are defined as the sum of wage income on the form 1040 and self-employment income from Schedule SE on the 1040 form.[41] We use wage income as reported on Form 1040 (instead of what is reported on W-2 forms) for filers because wage income on Form 1040 includes wages earned abroad, which can be significant particularly for children at the top of the income distribution. In particular, children who move abroad (but are U.S. citizens) are required to file standard tax returns and report their worldwide income, including any foreign earnings, but those earnings do not show up on W-2s.

For a child who is a married filer, individual earnings are defined as the sum of individual self-employment income from Schedule SE form 1040, and individual wage income defined as W-2 wage income plus one half of non-W-2 wage income from Form 1040.[42] Since we do not restrict the sample to children who are alive at the point at which we measure their income, children who are deceased are assigned zero earnings.

B. College Attendance: Data Sources and Methods

In this appendix, we describe the data sources and methods we use to assign students to colleges. The appendix is divided into five subsections. First, we describe our two sources of college attendance records and the differences in how they define colleges and annual attendance. Second, we describe how we homogenize their college definitions. Third, we discuss how we homogenize their annual attendance definitions and compile annual attendance records from the two data sources. Fourth, we describe how we identify and remove a small set of colleges who have incomplete 1098-T data. Finally, we summarize annual enrollment counts for our college attendance definitions.

Data Sources. We combine two data sources to measure student-level college attendance: Form 1098-T records and National Student Loan Data System (NSLDS) Pell grant recipient records.[43] Note that neither data source relies on the student or the student's family to file a tax return, and neither data source contains information on course of study or degree attainment.

Form 1098-T is an information return that is submitted by colleges to the U.S. Treasury Department. Each calendar year, higher education institutions eligible for federal financial aid (Title IV institutions) are required to file a 1098-T form for every student whose tuition has not been waived by the college (i.e. any student who pays or is billed tuition, or who has any non-governmental third party paying tuition or receiving tuition bills on his or her behalf). The form reports tuition

[41] Self-employment income is the amount for total tentative net earnings from self-employment. It is reported on Form 1040, Schedule SE, Section A or B, Line 3. We recode negative self-employment income as zero because negative self-employment income is generally due to business losses and is thus generally correlated with having a high level of latent income or wealth. We multiply self-employment income by 0.9235 to align treatment with wage earnings (as wage earnings are net of the 7.65% employer social security payroll tax).

[42] It is not possible to attribute to each specific spouse 1040 wage income that is not reported on the W2 forms. Hence, our decision to split such wage income equally across spouses.

[43] The full NSLDS data include data on recipients of Pell grants and federally subsidized loans. We use only the Pell grant data in our main attendance measures because almost all non-Pell students in the NSLDS data already appear in the 1098-T data, and using the non-Pell NSLDS records would likely generate more erroneous assignments due to timing inconsistencies across the two types of data (see below) than it would correct missing data.

payments or scholarships received for the student during the calendar year. Title IV institutions include all colleges and universities as well as many vocational colleges and other postsecondary institutions, all of which we refer to as "colleges." Colleges are indexed in the 1098-T data by the college's Employer Identification Number (EIN) and its ZIP code. We use 1098-T data for students during calendar years 1999-2013.

Most colleges file a 1098-T for every student, regardless of whether the student's tuition has been waived. However, some colleges do not file a 1098-T for students who pay no tuition. Almost all such students with American parents are from low-income families, are eligible for a Pell grant from the federal government, and are required by their colleges to acquire a Pell grant in order to receive their tuition waiver.[44]

We therefore supplement the 1098-T records with records from the administrative NSLDS Pell records. The NSLDS contains information on every Pell grant awarded, including the college receiving the Pell payment (Pell grant payments are remitted directly from the federal government to the college the student attended). The NSLDS Pell data are indexed by award years, defined as the spring of the academic year beginning on July 1. We use NSLDS Pell data for all students in award years 1999-2014, comprising Pell awards for enrollment spells that began between the dates July 1, 1999, and June 30, 2014 (roughly academic years beginning in calendar years 1999-2013). Colleges are indexed in the NSLDS Pell data by the six-digit federal OPEID (Office of Postsecondary Education Identification) identifier.

We use the NSLDS Pell data to impute missing 1098-T data and thereby construct comprehensive student-college-year attendance records 1999-2013.[45] Doing so requires homogeneous college and time-period definitions across the two data sources, but the two data sources differ in these definitions. The next two subsections detail our methods for homogenizing those definitions and constructing comprehensive student-college-year attendance records.

Combining 1098-T and NSLDS Pell Records. Empirical work on higher education is frequently conducted at the level of the six-digit OPEID (hereafter "OPEID"). We therefore use the NSLDS Pell and loan data to construct a crosswalk between EIN-ZIP pairs from the 1098-T data (i.e. the EIN and the ZIP code of the college) and OPEIDs from the NSLDS Pell data. In almost all cases, each EIN-ZIP pair maps to a single OPEID. In the rare cases in which a single EIN-ZIP pair maps to multiple OPEIDs, we cluster the OPEIDs together and conduct our analysis as if the cluster were a single college. We refer to this unit of analysis—either an OPEID or a cluster of OPEIDs—as a "Super OPEID."

Our procedure for mapping EIN-ZIP pairs to OPEIDs relies on the fact that almost all students who receive a federally subsidized loan (and most students who receive a Pell grant) for attending a given college x in academic year t to t+1 will also have a 1098-T from college x in calendar year t or t+1 or both. Thus by merging students in the NSLDS to students in the 1098-T data within narrow time-period bands, we can infer the NSLDS OPEID that corresponds to each 1098-T EIN-ZIP pair.

Specifically, we first merge the full NSLDS data to the 1098-T data at the student level (without using any college identifiers), in order to identify records with both an OPEID (from the NSLDS

[44] The vast majority of students appear in the 1098-T database. When we measure college attendance between the ages of 19 and 22 (as in our baseline analysis), 95.9% of the students in our analysis sample appear in the 1098-T records. A larger share of observations come from the NSLDS Pell records for lower income families (Appendix Figure V), but even in the bottom parent income quintile, 87.1% of students appear in the 1098-T records.

[45] Our approach misses students who attend a college that does not file 1098-T's for all students and who have their tuition entirely waived despite having parental income above the Pell grant eligibility threshold. Such students could include top athletic recruits. We believe that such cases are rare, as shown by the high correlation between the counts of students in our data and total counts from IPEDS.

data) and an EIN-ZIP (from the 1098-T data). We conduct the merge requiring that the NSLDS student's masked taxpayer identification number (TIN, i.e. her masked Social Security Number) equals the 1098-T student's masked TIN, as well as requiring the NSLDS award year equals either the 1098-T calendar year or the 1098-T calendar year plus one. Merging by year and year-plus-one is appropriate given the award year definition (see above subsection on data sources). Only rows that are successfully merged are retained.

The resulting merged dataset contains many correct matches between OPEIDs and EIN-ZIP pairs and some incorrect matches. For example, a student who uses a federally subsidized loan at UC-Berkeley and was billed tuition at both Berkeley (during the school year) and Stanford (for summer school) will have two rows in the merged data: one with Berkeley's OPEID and Berkeley's EIN-ZIP pair and another with Berkeley's OPEID and Stanford's EIN-ZIP pair. In order to correctly map Berkeley's OPEID and EIN-ZIP pair, we rely on the fact that most Berkeley students do not also attend Stanford.

To algorithmically identify the correct link between OPEIDs and EIN-ZIPs, we construct counts by OPEID, EIN-ZIP, and calendar year in the merged dataset. The distribution of counts exhibits very clear mass points and almost always stable across years: nearly all the counts of each OPEID appear in a single OPEID-EIN-ZIP cell, and almost all the counts of each EIN-ZIP appear in a single OPEID-EIN-ZIP cell. Using this algorithm, we construct a mapping of EIN-ZIP pairs to OPEIDs by identifying the OPEID(s) that appear most frequently for each EIN-ZIP pair and thus likely correspond to the same college. In the final step, OPEID-EIN-ZIP triads were confirmed to correspond to the same college via manual comparison of NSLDS college names and 1098-T college names, and the small number of discrepancies were addressed using manual adjustments to the crosswalk.

Finally, we cluster OPEIDs as follows in order to produce our final Super OPEID crosswalk, which maps every OPEID to a single Super OPEID and maps every EIN-ZIP pair to at most one Super OPEID. If an OPEID's matched EIN-ZIP pair(s) matched only to that given OPEID, then we map the OPEID and all of the OPEID's matched EIN-ZIP pairs to a Super OPEID equal to the OPEID.[46] If instead an OPEID's matched EIN-ZIP pair(s) match to multiple OPEIDs, then we map all of the matched OPEIDs and their matched EIN-ZIP pairs to a Super OPEID equal to a unique number that is smaller than the smallest OPEID so that there are no conflicts.[47] OPEIDs that did not credibly match at least one EIN-ZIP pair and EIN-ZIP pairs that did not credibly match to any OPEID are assigned Super OPEID -1 (colleges with insufficient or incomplete data). We treat Super OPEID -1 as a separate "college" and include it in our publicly released statistics, but omit it from most analyses unless otherwise specified.

We use the Super OPEID crosswalk to assign a Super OPEID to every record in the NSLDS data and every record in the 1098-T data. The crosswalk comprises 5,327 Super OPEIDs: 5,208 unaltered OPEIDs (values ranging from 1002 to 42346) and 119 newly created clusters of OPEIDs

[46] For example, Cornell (OPEID 190415) may submit 1098-T forms from the same EIN but from two ZIPs—one ZIP corresponding to its Ithaca campus and another ZIP corresponding to its New York City campus. In this case, we map Cornell's OPEID and its two EIN-ZIP pairs to Super OPEID 190415.

[47] For example, the University of Massachusetts system comprises four undergraduate campuses, each with its own OPEID. However, all University of Massachusetts 1098-Ts are submitted from the same centralized EIN-ZIP. We therefore map all four of University of Massachusetts's OPEIDs and the University of Massachusetts EIN-ZIP to a new Super OPEID value that is smaller than 1000 (125 in the case of the University of Massachusetts). Note that all OPEIDs are larger than 1000.

(positive values below 1002). 2.7% of NSLDS Pell records and 1.1% of 1098-T records from 1999-2013 are assigned Super OPEID -1.[48]

Imputing Calendar Year Attendance Records for Pell Recipients. The vast majority of student-college-year attendance observations appear in the 1098-T data, which measure attendance by calendar year. Therefore, after using our Super OPEID crosswalk to assign a consistent college definition to every NSLDS Pell record and ever 1098-T record, we use information from the NSLDS on dates of attendance to impute missing 1098-T data, thereby yielding comprehensive attendance records by calendar year from 1999-2013.

We map the NSLDS data to calendar years as follows. For every NSLDS Pell student at a Super OPEID x and a Pell award enrollment start date lying in calendar year t, we impute a 1098-T for the student at Super OPEID x in calendar year t. Then, for every NSLDS Pell student at a Super OPEID x and a Pell enrollment start date in the second half of calendar year t and with a Pell grant amount equal to more than 50% the student's maximum eligible Pell amount in the award year, we additionally impute a 1098-T for the student at Super OPEID x in calendar year $t+1$. Finally, we remove duplicate records. The remainder of this subsection explains the logic underlying this imputation strategy further.

The NSLDS Pell data contain the start date of the enrollment period covered by the Pell grant. If the college had submitted 1098-Ts on behalf of a given Pell student whose enrollment period began in calendar year t, the college would likely have submitted a 1098-T for the student in calendar year t (had it been required to do so). Thus, for every NSLDS Pell student with Super OPEID x and an enrollment start date in calendar year t, we impute a 1098-T for the student with Super OPEID x and calendar year t.

If the college had submitted 1098-Ts on behalf of a given Pell student, and if that student's enrollment period straddled a fall and spring term, the college would likely have submitted a 1098-T in calendar year $t+1$ as well as in calendar year t. The NSLDS Pell data do not contain the end date of the enrollment period covered by the Pell grant. However, they do contain the share of the student's maximum eligible Pell amount in the award year that was allocated to the grant. Pell grants for a single semester typically have an amount equal to only half of the student's annual Pell maximum grant amount, even if tuition is very expensive. Hence for every NSLDS Pell student with Super OPEID x who has an enrollment start date between July and December of year t and has strictly greater than 50% of the student's maximum Pell eligibility amount allocated to the grant, we impute a 1098-T for the student with Super OPEID x and calendar year $t+1$.

After these imputations, we drop observations that are duplicates in terms of student, Super OPEID, and calendar year. This allows students to be recorded as having attended any number of Super OPEIDs in a calendar year, but ensures that they are not recorded as having attended any Super OPEID more than once in a calendar year.

There are no public measures of calendar-year Pell attendance that can be used to directly validate the imputation procedure described above. However, indirect validation methods suggest a high degree of fidelity. The share of our students on a Pell grant in the average calendar year is very highly correlated with, and similar in levels to, approximations to annual Pell student shares based on publicly available data. Moreover, at colleges with substantial numbers of students on Pell grants, the imputation algorithm adds almost no net students to 1098-T attendance records—consistent with these colleges issuing 1098-T forms for all students regardless of their tuition billing status and with our algorithm only imputing 1098-Ts in calendar years that the student was in fact enrolled.

[48] The rate of 1098-T assignment to Super OPEID -1 is 11.2% in 1999 and is between 0.04% and 2.2% from 2000-2013. The 1999 1098-T data lack the ZIP code of the college, so in that year only, we assign Super OPEID using the subset of EINs from the Super OPEID crosswalk that map to a single Super OPEID regardless of ZIP code.

Removing College-Years with Incomplete 1098-T Data. A small number of college-year observations have incomplete 1098-T data, either because of errors in administrative records or because of changes in EIN's and reporting procedures.[49] We discard these defective college-years by flagging them using two methods based on counts of total students. The counts described below are constructed using the total counts of forms 1098-T and Pell grants for all children born in 1980-1991, regardless of a successful link to parents and regardless of whether the student attends several institutions.

First, for each college-year, we compare the count of individuals receiving a 1098-T form but excluding Pell grants (what we call the 1098-T-only count) versus the count of individuals receiving either a 1098-T form or a Pell grant (what we call the full count). When the 1098-T-only count is less than 10% of the full count, we conclude that there are too few 1098-T forms for the data to be complete and flag the college-year. In the vast majority of these cases, the 1098-T counts are exactly zero, implying that the college did not report any 1098-T form (most likely because the information was not transmitted correctly to the IRS or because the institution used a different EIN-ZIP in that specific year). We use the 10% threshold as a way to capture rare situations where the 1098-T counts are not exactly zero, but are clearly too small relative to the number of Pell grants to be plausibly complete.

Second, we also flag college-years when full counts are too low (less than 75%) or too high (over 125%) relative to both the preceding and subsequent years. Such abnormal changes in counts likely reflect a data reporting issue rather than true changes in enrollments, which tend to be very smooth across years.

In total, these two flags account for 2.4% of (enrollment-weighted) college-year observations and 21.9% of college-year observations when not weighing by enrollment. The rate is much higher unweighted because the data at very small colleges is much less complete.

We discard college-year records that are flagged as incomplete before assigning students to their "most-attended" college or the first college, in order to ensure that our sample accurately represents attendance at each college. Our baseline measure of a child's most-attended college uses four years of data (the years when the child turns 19, 20, 21, and 22).[50] A college which has defective (and hence discarded) data for more than 1 year out of the 4 relevant years is re-assigned to super_opeid=-1, the pool of colleges with "incomplete or insufficient data." As a result, a college is retained in our cohort-level data only if we have valid data for at least 3 years (out of the 4 years).

Enrollment Counts for Attendance Measures. The steps described above yield a student-college-year level dataset that provides a complete record of college attendance in the U.S. during calendar years 1999-2013 for children born between 1980-1991. This dataset contains 207.6 million observations.

Using this dataset, we construct the three measures of college attendance—the most-attended college (our primary measure), age-20 college, and first-attended college—following the definitions given in Section II.B. In what follows, we document the impact of the restrictions imposed in each

[49] Most of these cases are college-year cells with zero 1098-Ts in the database. For example, in the years when the 1098-T first began to be collected (1999-2002), a number of small universities do not have any records at all in the database. In addition, some universities switch from reporting data separately for each campus to using a single EIN-ZIP for all their campuses, which creates inconsistencies in their data across years.

[50] For example, we measure college attendance using data from 1999 to 2002 for children born in the 1980 cohort. We measure college attendance starting with the year the student turns 19 because the 1098-T data are only available starting in 1999, making 19 the first observed age for the 1980 birth cohort. Omitting the year in which children turn 18 is not consequential because very few children attend college only in the calendar year in which they turn 18; for instance, only 1.6% of the children in the 1982 birth cohort attended college in the year they turn 18 but not between the ages of 19-22.

definition on sample sizes and report the share of observations obtained from the 1098-T vs. NSLDS datasets.

To construct the most-attended definition, we first restrict the full dataset to attendance between ages 19-22, which leaves 114.6 million records. Condensing the student-college-year data to the student level using the most-attended definition (see Section II.B) leaves 33.1 million student-level records. Eliminating students we cannot match to parents or whose parents had negative income leaves 31.0 million records. Finally, restricting to birth cohorts 1980-1982 (as we do in our main analysis) leaves 6.7 million records; including non-college-goers in this sample yields a sample size of 10.8 million children.

As described in Section II.D of Chetty et al. (2017), we impute income statistics and attendance for colleges with missing data for the 1980-82 cohorts using data from the 1983 and 1984 cohorts. Specifically, if a college is missing one or more years of data for the 1980-82 cohorts—either because of incomplete reporting of 1098-T forms or because the college opened more recently—we impute values for the missing cohorts using data from the 1983-84 cohorts. To impute a missing income statistic y_{ct} for college c in cohort t, we first estimate an OLS regression $y_{ct} = \alpha + \beta^y_{1983} y_{c,1983} + \beta^y_{1984} y_{c,1984} + \varepsilon$ using the sample of all colleges with non-missing data in cohort t as well as 1983 and 1984, weighting by enrollment. We then impute values for missing cohorts with the predicted values from this regression, based on each college's actual data in 1983 and 1984 (omitting colleges with missing data for 1983 or 1984). Such imputations account for 9.0% of enrollment-weighted observations in the analysis sample (1980-82 birth cohorts).[51]

We use this imputation procedure to impute data for 596 (27%), 520 (24%), and 406 (18%) colleges in cohorts 1980-1982, respectively, accounting for 570,000 additional students (9.0% of college attendees and 5.0% of all children). For the remaining 264 colleges that are missing data for either the 1983 or 1984 cohorts, we do not impute any values. This leaves us with 11.3 million children in our core sample underlying our main analysis.

9.2% of our annual attendance records for students aged 19-22 were not in the 1098-T data and appeared only in the NSLDS Pell data. Using our most-attended college attendance measure, 4.1% of the students in our analysis sample were not in the 1098-T data and originally appeared only in the NSLDS Pell data. The NSLDS Pell data has a smaller impact at the student level than the student-year level because many students attend a given college for multiple years and receive a 1098-T form in at least one of those years.

We define a child's *age 20 college* as the college the child attends in the calendar year that she turns 20.[52] To construct the age-20 definition, we restrict the full dataset to attendance at age 20, which leaves 30.5 million records. If a student attends multiple colleges at age 20, we weight the student-college-level records such that each student carries a total weight of one, leaving 27.3 million effective records for 26.1 million students. After bringing in non-college-goers under this definition, restricting to birth cohorts 1980-1982, and restricting to students matched to parents

[51] This imputation procedure helps increase the coverage of colleges in the analysis sample because a number of small colleges began reporting 1098-T data only in 2002. However, all of the main findings of the paper hold if we restrict attention to the set of colleges with no imputed data. The imputation leads us to slightly overstate the aggregate college attendance rate in the analysis sample, as some of the students for whom we impute college attendance from later data may already have been assigned to another college that they also attended or to the "colleges with insufficient data" category. Such double-counting turns out to be very small in practice (see later in this appendix for further details).

[52] If a student attends multiple colleges at age 20, we break ties by assigning the college that the student attended in the subsequent year, if any. For observations where ties remain (e.g., because the student attended the same multiple colleges the following year as well), we retain all colleges and weight each student-college observation by the reciprocal of the number of colleges attended (so that the total weight of each student in the analysis remains constant).

with weakly positive income, we have 11.0 million records for 10.8 million children. Finally, we impute income statistics and attendance for colleges with missing data for the 1980-82 cohorts as described above, leaving us with a 11.3 million person sample underlying our age-20 analysis.

We define a child's *first-attended college* as the college that a child attends first between the calendar years in which she turns 19 and 28 (inclusive), breaking ties using the same method as in the age 20 definition. To construct the first-attended definition, we restrict the full dataset to attendance between the ages of 19 and 28, which leaves 175.4 million records. If a student begins multiple "first-attended" colleges in the same year, we assign the student a college based on the method described in Section II.B, leaving 37.0 million records. Bringing in non-college-goers under this definition, restricting to birth cohorts 1980-1982, and restricting to students matched to parents with weakly positive income leaves 10.9 million records for 10.8 million children. Finally, we impute income statistics and attendance for colleges with missing data for the 1980-82 cohorts, leaving us with a 10.8 million person sample underlying our first-attended analysis. The reason that the first-attended definition yields slightly fewer records than the others is that we do not double-count students assigned to Super OPEID -1 (insufficient or incomplete data) in the final imputation step under this definition.

Comparisons to IPEDS Counts. We assess how well our methodology approximates the set of undergraduate degree seekers we seek to identify by comparing the count of students in our data to enrollment data from IPEDS. IPEDS does not have enrollment counts that exactly match our cohort-based definitions and age ranges, making direct comparisons difficult for many colleges, especially those where students enter at various ages. However, at highly selective colleges (defined as 176 colleges in the top two tiers of the Barron's 2009 selectivity index), the vast majority of students enter at age 18 and graduate in four years, making the number of first-time, full-time undergraduate students recorded in IPEDS a good approximation to our definition. Among these colleges, the correlation between our enrollment counts and the number of first-time, full-time undergraduates in IPEDS is 0.99.[53] In addition, IPEDS data show that 98.0% of full-time undergraduate students are degree seekers, suggesting that the number of summer school or extension school students in our sample is likely to be very small.[54]

C. Comparison of Incomes in Tax Data vs. American Community Survey Data

In this appendix, we compare the income distribution of parents in the tax data to incomes reported in the American Community Survey. Most prior work on the income distribution of families in the United States focuses on money income (defined as pre-tax market income plus cash transfers from the government) and uses the household unit (defined as all individuals living in the same dwelling). In contrast, in the tax data, we define income as Adjusted Gross Income (pre-tax income excluding government cash transfers) and define the unit of observation as the tax filing unit (either a single person or a married couple, excluding other household members).

We show that income distributions in the ACS are very similar to those in the tax data if we use the same household unit and income definitions. We focus on the annual family income of children aged 15 in 2000 (Panel A of Appendix Table II) and children aged 15 in 2006 (Panel B). We begin by describing how we define income and family units in the two datasets.

Tax Data. The unit of observation for family income is the tax unit. As described in Appendix A, the tax unit, i.e. the child's parents, are defined as the person(s) who claim the child as a

[53] The IPEDS counts are 3% larger than our counts on average, which likely reflects international students not included in our sample.

[54] Our methodology could be further tested and refined by linking external data on college attendance—for instance, from the National Student Clearinghouse—to the tax records, as in Hoxby (2015).

dependent on a 1040 form in the year the child turns 17. Children are either assigned to married parents or a single parent, and income is defined as Adjusted Gross Income (AGI), which is pre-tax and pre-transfer cash income when the child is 15.[55] AGI is rounded to the nearest $250 for disclosure purposes.

ACS Data. To illustrate how family and income definitions affect the results, in the ACS data, we consider both standard household definitions and construct a concept analogous to tax units. Similarly, we also consider both the traditional total money income measure and a concept analogous to AGI.

To define households, we simply use the household ID that uniquely identifies each household in our ACS sample. We then restrict the sample to children aged 15, excluding a small number of individuals aged 15 who are listed as the head of the household. To create the tax unit claiming the child, we determine who claims the child for tax purposes as follows. A child is assigned to married parents if the child lives in the same household with parents who are married and one of the parents has non-zero income. If both parents' incomes are zero, we use instead the income of the head of household as the head would most likely claim the child for tax purposes in this case. For children not living with married parents, we assign the mother as the parent if she is present in the household and has non-zero income. We assign the father as the parent if the mother is absent or has zero income. Finally, if both father and mother are absent, we define the parent as the head of household.

We obtain total money income for each household member directly from the ACS. We define AGI starting from total money income and subtracting social security income (retirement and disability benefits), veterans benefits, supplementary security income, welfare (public assistance) income, and other non-taxable cash transfers. AGI for the tax unit is defined as the income earned by each of the parent(s), summing across the two parents when the child is assigned to married parents.

Results. Column 1 of Appendix Table II presents statistics on the income distribution in the tax data. Column 2-4 present analogous statistics using the ACS data. Column 2 presents ACS statistics using the tax unit and adjusted gross income, which replicates the concepts we use in the tax data. The statistics in columns 1 and 2 are very similar. Notably, the quantiles of the earnings distribution (P10, P25, P50, P75, P90) are very similar across the two datasets. Mean incomes are higher in the tax data by about 15% because the ACS data are top-coded whereas the tax data are not. The fraction of children with zero tax unit income is slightly higher in the ACS (5.0%) than in the tax data (3.4%), perhaps because survey respondents fail to report very small income amounts.

Column 3 replicates the ACS income statistics at the household level instead of the tax unit level. Because a household can be larger than a tax unit (e.g., a child living with both a parent and a grandparent), adjusted gross incomes at the household level are substantially higher than incomes at the tax unit level. Column 4 then expands the income definition to total money income. This further increases income, particularly in the lower percentiles. The fraction with zero incomes falls to 0.4% and the bottom percentiles (P10, P25, P50) are now substantially higher than in the tax data because of cash transfers.

In sum, a naive comparison between survey data using typical money income and household definitions and tax data using the tax unit and adjusted gross income definitions would mistakenly suggest that tax data incomes are substantially lower than survey income. However, this discrepancy is entirely due to the differences in household unit and income definition. Once one uses the

[55] This is different from our main income concept used throughout the paper, which averages the parent's income when the child is 15-19. We use annual income here for comparison with the ACS data, where we only observe annual income.

same definitions, the distribution of incomes reported in the tax data are well aligned with those in standard surveys.

D. Stability of Children's Earnings Ranks

In our analysis sample comprising the 1980-1982 birth cohorts, we measure earnings at ages 32-34. Measuring children's earnings ranks when they are too young can potentially yield misleading estimates of lifetime earnings ranks because children with high lifetime earnings have steeper earnings profiles (e.g., Haider and Solon 2006; Solon 1999). This issue may be especially acute for analyses of earnings outcomes at elite colleges, where many students go on to pursue advanced degrees. In this appendix, we show that ages 32-34 are sufficiently late in a child's life to obtain a reliable measure of children's ranks at all colleges. Of course, individuals' earnings *levels* continue to rise sharply during their thirties, but a rank-preserving fanning out of the earnings distribution does not affect the rank-based analysis of Section IV.

To evaluate when children's earnings stabilize, we examine how the earnings of children evolve by age at each college. In order to examine the profile of earnings over the broadest range of ages, we go back to the 1978 birth cohort for this analysis. For children born in 1978, we can observe college attendance starting at age 21 in 1999 and earnings up to age 36 in 2014.[56] We assign each child a college based on the college he or she attends most frequently in 1999 and 2000, following the same approach as we use in our baseline college definition described in Section II.B. We assign children percentile ranks at each age by ranking them relative to other children in the 1978 cohort in each calendar year.

Appendix Figure Ia plots the mean earnings ranks of children from ages 25 to 36 for children who attended colleges in four mutually exclusive tiers: Ivy-Plus, Other Elite (Barron's Tier 1 colleges, excluding the Ivy-Plus group), other 4-year colleges, and 2-year colleges. For individuals who attended elite colleges, and especially Ivy-Plus colleges, earnings ranks rise sharply from age 25 to 30. If we were to measure children's earnings at age 25, we would find that children at Ivy-Plus colleges have *lower* income ranks than those who attend less selective 4-year colleges. Mean ranks at elite colleges stabilize at approximately the 80th percentile after age 30, with very little change starting at age 32. In contrast, the age profiles at lower-tier colleges are virtually constant from ages 25 to 36, at approximately the 60th percentile for 2-year colleges and the 70th percentile for non-elite 4-year colleges.

The stabilization of mean earnings ranks once children reach their early thirties holds not just across college tiers, but also across individual colleges. To characterize the college-level patterns, we examine the mean ranks of students who attend each college at each age from 25-36. Appendix Figure Ib plots the (enrollment-weighted) correlation of the mean ranks at each age with the mean ranks at age 36 across colleges. Consistent with the patterns in Appendix Figure Ia, this correlation rises sharply between ages 25 and 30, when it reaches 0.98 and stabilizes. We find analogous stabilization across all quantiles of the distribution by the early 30s, including the probability that children reach the top quintile or the top 1% of their age-specific income distribution (Appendix Figure Ic-d).[57]

[56] We do not use the 1978 cohort for our primary analysis of intergenerational mobility because we cannot link children in the 1978 cohort to their parents based on dependent claiming. However, linking children to their parents is not necessary to analyze the unconditional distribution of children's earnings as we do here.

[57] At the vast majority of colleges, earnings ranks stabilize by age 25, implying that one can reliably analyze earnings outcomes for the 1980-89 cohorts with our publicly available data for most colleges.

E. SAT/ACT Data

For individuals who took either test multiple times, we use the individual's maximum composite score. The mean SAT and ACT scores for children in our sample for whom we observe a score is 989 and 21.8, which are each roughly comparable to the mean scores of 1026 (reported by the College Board) and 20.9 (reported by ACT) for the high school graduating class of 2004.

We combine the SAT and ACT data as follows into a single test score, which lies on the SAT's 400-1600-point scale. For the 47.6% of college students with an SAT score, their SAT/ACT score equals their SAT score – including for the 14.3% of college students with both an SAT score and an ACT score. To facilitate non-parametric matching, we coarsen SAT scores into 20-point bins throughout our analysis. For the 26.2% of college students with an ACT score but not an SAT score, we convert ACT scores to SAT using the 2016 ACT/SAT concordance table (Summit, 2016) in which each ACT score is mapped to a range of SAT scores. For each person with an ACT score, we randomly select a 20-point SAT score bin from the range of possible scores.

F. Methods for Counterfactuals

This appendix provides further details on how we implement the counterfactual allocations of students to colleges and construct counterfactual earnings distributions in Section IV.

For the income-neutral allocations counterfactual, we first rank all college students by their SAT/ACT scores, breaking ties with idiosyncratic noise, and record the ranks of the students actually enrolled at each college c, $\overrightarrow{r}_c$. We then re-rank students by their original test scores, *breaking ties with new idiosyncratic noise*, and allocate ranks $\overrightarrow{r}_c$ to college c. By breaking ties with new idiosyncratic noise, we randomly assign students to each college c within test score bins, including students who were not admitted to or did not apply to college c.

The need-affirmative allocations counterfactual employs the same procedure with one change: students are re-ranked by their original SAT/ACT scores plus the increment corresponding to their parent income quintile, as defined in Section IV.A. By applying each college's actual test score ranks $\overrightarrow{r}_c$ to the post-test-score-bonus distribution, we effectively re-norm the post-test-score-bonus distribution to align with the actual SAT distribution.

We construct a counterfactual earnings distribution for children at each college based on the observed distribution of earnings for children in each parent income quintile, SAT/ACT score level, and college. Children are randomly assigned the earnings of another child who is observed as attending their counterfactually assigned college and who has the same parent income quintile and SAT/ACT score. For example, suppose that Harvard actually enrolled 10 bottom-quintile children with a 1400 SAT/ACT score and that the counterfactual assigned 30 bottom-quintile children with a 1400 SAT/ACT score to Harvard. Each of those 30 students would be randomly assigned the earnings of one of the 10 actual enrollees. In 1.1% of observations in income-neutral allocations and 1.4% in need-affirmative allocations below, a student is allocated to a college in which no student of the same parent income quintile and same SAT/ACT score actually enrolled (e.g., a bottom-quintile 1400 student is allocated to Harvard but Harvard enrolled no bottom-quintile 1400-scorer in actuality). In those cases, we assign those students a earnings rank of an actually enrolled student from the same parent income quintile with the nearest SAT/ACT score, which is on average 31 points away.

G. Replication of Dale and Krueger (2014)

In this appendix, we show that our data yield estimates of the return to attending a highly selective college that are similar to those estimated by Dale and Krueger (2014), and in particular exhibit higher rates of return to attending more selective colleges for students from lower-income families.

We replicate the key specifications estimated by Dale and Krueger in Appendix Table X. We focus on a set of 31 highly selective colleges in the College and Beyond sample and estimate a set of specifications that parallel those in Columns 3 and 4 in Table 2 of Dale and Krueger (2014).[58] The dependent variable is log earnings, excluding observations with earnings below $13,822 in 2007 dollars ($15,800 in 2015 dollars), as in Dale and Krueger's analysis. The key independent variable is the mean SAT score (divided by 100) of the college in 2001, a proxy for the college's selectivity.

Column 1 presents OLS regression estimates controlling only for $f(S_i)$ and $f(p_q)$, the quintics in test scores and parent income rank. The coefficient of 0.016 implies that a 100 point increase in SAT score increases earnings by 1.6% on average. In Column 2, we add fixed effects for the set of colleges to which students sent their test scores (among the 30 colleges in the sample). The coefficient on mean SAT scores remains similar at 1.2% in this specification, with 1.6% lying in the 95% confidence interval. For comparison, Dale and Krueger (2014, Table 2, Column 2) report an estimate of -0.001 (s.e. 0.012) in an analogous specification in their data. Our estimates – both in the baseline specification and the specification that controls for college fixed effects thus lie within one standard error of their point estimate and hence are not statistically distinguishable from it.

This result should not be taken to mean that colleges have no causal effects: as Dale and Krueger emphasize, there are substantial differences in earnings across colleges that are orthogonal to mean SAT scores within the small set of highly selective colleges they study, even after controlling for selection on observables and unobservables. Moreover, we find much larger differences in earnings outcomes when we consider all colleges, looking beyond the highly selective institutions in the College and Beyond sample.

In Columns 3-7 of Appendix Table X, we investigate how the return to attending a more selective college varies with parental income by replicating the specification in Column 1 by parent income quintile. The coefficients on mean SAT scores decline across the columns as we examine higher income groups: attending a college with 100 point higher average SAT scores is associated with a 4% increase in earnings for students from the bottom quintile, but only a 1% increase in earnings for students from the top quintile. Again, we find similar results using specifications that control for the college application set (not shown). These findings match Dale and Krueger's conclusions that the returns to attending a more selective college are larger for students from low-income families.

H. Relationship between ACT/SAT Scores and Earnings

We follow Dale and Krueger (2002), Hoxby and Avery (2013), and many others in using standardized test scores as a proxy for academic credentials. Test scores are widely used as proxies for academic qualifications because they are widely available and because previous work has shown that they are predictive of later outcomes (e.g., Sackett et al. 2012, Kurlaender and Cohen 2019), although the extent to which that predictive power comes simply from correlations with demographics such as parental income and race is debated (Rothstein 2004). In this appendix, we re-evaluate the relationship between SAT scores and earnings using our longitudinal data. We use our baseline

[58]The original College and Beyond sample includes 34 colleges; we do not have data for Morehouse, Tulane, and Williams.

1980-1982 cohorts sample for this analysis, omitting students who do not take either the SAT or ACT, and rescale ACT to SAT scores as discussed in Appendix E.

The series in circles in Appendix Figure V presents a binned scatter plot of the relationship between earnings ranks in adulthood (measured at ages 32-34) and test scores. There is a strong positive, linear relationship across the distribution. Column 1 of Appendix Table XIII presents an OLS regression corresponding to this binned scatter plot. We estimate that a 100-point higher SAT score (out of 1600) is associated with earning $6,744 more annually at agess 32-34, or 2.73 percentile ranks higher in the income distribution.

In subsequent columns of Appendix Table XIII, we assess how this relationship changes as we add additional controls for demographic factors. We find that the relationship between SAT scores and earnings falls by about 20% when we control for parental income, race, gender, and high school. The series in triangles in Appendix Figure VI present a binned scatter plot analogous to the specification in Column 4 by replacing the linear SAT score term with 20 bins for SAT scores. It confirms that there is a strong relationship between SAT scores and later earnings even conditional on demographics throughout the test score distribution.

Columns 5-9 examine this relationship within individual colleges. These coefficients must be interpreted with caution, since the college a student attends is endogenous to their SAT scores. Nevertheless, a 100-point difference in SAT scores between students at the same college, and of the same demographic background, still predicts substantial differences in earnings. This relationship holds within each of the selective tiers of colleges.

We conclude that SAT scores provide an informative proxy for qualifications at the point of college application for our purposes in the sense that it predicts earnings above and beyond demographics. We note, however, that our analysis provides no evidence on how standardize test scores compare to other potential proxies for academic credentials, such as high school grades or other forms of assessment, and hence does not speak to the question of whether standardized tests provide good measures of qualifications more broadly.

TABLE I
Summary Statistics for Analysis Sample

	Sample		
	All Children in 1980-82 cohorts	Analyzed college-goers	Non-Goers in 1980-82 Cohorts
	(1)	(2)	(3)
A. College Attendance Rates			
% Attending College Between Age 19-22	61.83	-	-
% Attending a College in Data Release (based on 80-82 cohorts)	53.07	-	-
% Not Attending any College by Age 28	26.65	-	69.81
B. Parents' Household Income (When Child is Aged 15-19)			
Mean Income ($)	87,335	114,306	50,377
Median Income ($)	59,100	N/A	37,400
20th Percentile Income ($)	24,633		
40th Percentile Income ($)	45,767		
60th Percentile Income ($)	73,500		
80th Percentile Income ($)	111,067		
99th Percentile Income ($)	532,267		
C. Children's Individual Earnings (in 2014, Ages 32-34)			
Mean Earnings ($)	35,526	46,179	20,256
Median Earnings ($)	26,900	N/A	13,600
20th Percentile Earnings ($)	900		
40th Percentile Earnings ($)	18,500		
60th Percentile Earnings ($)	35,200		
80th Percentile Earnings ($)	55,800		
99th Percentile Earnings ($)	182,467		
% Employed	81.68	88.60	70.96
Number of Children	10,757,269	6,244,162	4,106,026
Percentage of College Students Covered	-	93.9%	-

Notes: The table presents summary statistics for the analysis sample defined in Section II. Column 1 includes all children in the 1980-82 birth cohorts. Column 2 limits this sample to students who attend a college (between the ages of 19-22) that is included in the public data release, using imputed data from the 1983-84 birth cohorts for colleges with insufficient data in the 1980-82 birth cohorts (see Appendix B, and Section II for details). This is the sample used for most of our analyses. Column 3 includes children in the 1980-82 birth cohorts who did not attend college between the ages of 19-22. Children are assigned to colleges using the college that they attended for the most years between ages 19 and 22, breaking ties by choosing the college the child attends first. Ivy-Plus colleges are defined as the eight Ivy-League colleges as well as the University of Chicago, Stanford University, MIT, and Duke University. Elite colleges are defined as those in categories 1 or 2 in Barron's Profiles of American Colleges (2009). 4-year Colleges are defined using the highest degree offered by the institution as recorded in IPEDS (2013). Parent income is defined as mean pre-tax Adjusted Gross Income during the five-year period when the child was aged 15-19. Parent income percentiles are constructed by ranking parents relative to other parents with children in the same birth cohort. Children's earnings are measured as the sum of individual wage earnings and self-employment income in the year 2014. At each age, children are assigned percentile ranks based on their rank relative to children born in the same birth cohort. Children are defined as employed if they have positive earnings. In Column 2, the number of children is computed as the average number of children in the cohorts available for a given college multiplied by 3. Medians are not reported in Column 2 because the imputations are implemented at the college rather than individual level. We report dollar values corresponding to other key quantiles in Column 1 because those are the thresholds used to define the income groups we use in our analysis (bottom 20%, top 20%, etc.). All monetary values are measured in 2015 dollars. Statistics in Column 1 are constructed based on Online Data Tables 6 and 9; in Column 2 based on Online Data Table 2; and in Column 3 based on Online Data Table 6, with the exception of median income and earnings, which are constructed directly from the individual-level microdata.

TABLE II

Parent Income Segregation and Children's Earnings Outcomes: Statistics by College Tier

	Share of Parents From			Median Parent Income ($)	Median Child Earnings ($)	Within-College Rank-Rank Slope	Top-Quintile Outcome Rate		Mobility Rate		Num. of Colleges (80-82 cohorts)	Num. of Students (80-82 cohorts)
	Bottom 20% (%)	Bottom 60% (%)	Top 1% (%)				Top 20% (%)	Top 1% (%)	Top 20% (%)	Top 1% (%)		
College Tier:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Ivy-Plus	3.8	18.2	14.5	171,000	82,500	0.086	58.0	12.78	2.18	0.48	12	52,724
Other elite colleges	4.3	21.4	10.0	141,900	65,400	0.060	50.6	5.80	2.20	0.25	62	183,973
Highly selective public	5.5	29.0	2.5	107,300	53,600	0.099	40.7	2.67	2.22	0.15	26	393,548
Highly selective private	4.1	23.9	7.0	124,700	56,500	0.057	42.3	3.33	1.73	0.14	66	134,098
Selective public	8.4	39.8	1.3	87,100	41,600	0.102	23.3	0.70	1.95	0.06	364	1,944,082
Selective private	7.1	37.4	2.4	90,700	44,400	0.080	27.0	1.00	1.91	0.07	446	486,852
Nonsel. 4-year public	17.0	59.5	0.6	61,200	29,800	0.085	13.5	0.19	2.30	0.03	71	257,854
Nonsel. 4-yr. priv. non-prof.	10.7	45.2	2.0	80,500	29,000	0.079	13.6	0.42	1.45	0.04	50	55,947
2-year public and non-prof.	14.6	55.4	0.5	66,900	29,800	0.110	12.3	0.18	1.80	0.03	604	2,021,451
4-year for-profit	21.1	66.8	0.5	51,500	28,900	0.095	12.2	0.15	2.57	0.03	56	126,025
2-year for-profit	20.6	67.3	0.3	51,500	31,300	0.092	13.1	0.17	2.71	0.04	34	42,313
Less than two-year colleges	20.9	65.7	N/A	53,000	18,800	0.096	7.7	0.19	1.60	0.04	13	10,032
All colleges	10.8	45.0	1.7	80,500	38,100	0.090	18.0	0.59	1.95	0.06	1,804	5,708,899

Notes: This table presents statistics on parental income segregation and children's earnings outcomes by college tier; see Section II of this paper and Appendix D of Chetty et al. (2017) for definitions of these tiers. All statistics reported are for children in the 1980-82 birth cohorts. All distributional statistics are enrollment-weighted means of the exact values for each college, except for median parent income and child earnings, which are the mean incomes for the percentile of the overall income or earnings distribution which contains the within-tier median. For example, the median Ivy-Plus parent falls in the 92nd percentile of the overall income distribution and the mean income for Ivy-Plus parents in the 92nd percentile of the overall distribution is $171,000. The exact fraction of students from less than two-year colleges with parents in the Top 1% is not available due to small sample sizes in the publicly available data. The trend statistics are coefficients from enrollment-weighted univariate regressions of the share of parents from the bottom 20% or 60% on student cohort, multiplied by 11; the statistics can therefore be interpreted as the trend change in lower-parent-income shares over the 1980-1991 cohorts. Rank-rank slopes are coefficients from a regression of child income rank on parent income rank with college fixed effects, as in Panels E-G of Table III of Chetty et al. (2017); see notes to that table for further details. Top-quintile outcome rates are the fractions of children who reach the top 20% or 1% conditional on having parents in the bottom quintile. Mobility rates are the fractions of children who have parents in the bottom income quintile and whose own earnings place them in either the top 20% or top 1% of their own age-specific income distribution. Parents' incomes are measured at the household level when children are between the ages of 15 and 19, while children's incomes are measured at the individual level in 2014. See notes to Table I for further details on income definitions and how children are assigned to colleges. Statistics in Columns 1-4, 7-10, and 12 are constructed based on Online Data Table 6; in Column 5 based on Online Data Table 7; in Column 11 based on Online Data Table 3; and in Column 6 directly from the individual-level microdata.

TABLE III
Parental Income Distributions by College Tier Under Counterfactual Student Allocation Rules

	Parent Income Quintile					Share of all college goers
	1 (Bottom 20%)	2	3	4	5 (Top 20%)	
	(1)	(2)	(3)	(4)	(5)	(6)
A. Actual Distributions						
Ivy-Plus	3.8%	5.7%	8.7%	13.4%	68.4%	0.9%
Other elite colleges	4.3%	6.8%	10.2%	15.8%	62.8%	3.3%
Highly selective public	5.5%	9.2%	14.3%	23.4%	47.6%	7.0%
Highly selective private	4.1%	7.6%	12.2%	19.7%	56.5%	2.4%
Selective public	8.4%	12.9%	18.6%	26.1%	34.1%	34.4%
Selective private	7.1%	12.0%	18.2%	25.5%	37.2%	8.6%
Nonselective 4-year public	17.0%	20.4%	22.1%	22.7%	17.7%	4.6%
Nonselective 4-year private non-profit	10.7%	14.7%	19.8%	24.6%	30.2%	1.0%
2-year public and non-profit	14.6%	18.6%	22.2%	24.7%	19.9%	35.5%
4-year for-profit	17.8%	22.3%	22.5%	21.1%	16.3%	1.7%
2-year for-profit	21.5%	23.9%	23.1%	19.5%	12.0%	0.7%
Less than two-year colleges	20.7%	23.2%	21.3%	21.0%	13.8%	0.2%
All colleges	10.7%	14.8%	19.3%	24.4%	30.8%	100.0%
Underrepresentation in selective tiers	31.3%	21.6%	11.4%	-0.6%	-27.9%	
B. Counterfactual Distributions Under Income-Neutral Student Allocations						
Ivy-Plus	4.4%	7.3%	12.1%	18.5%	57.8%	0.9%
Other elite colleges	5.5%	9.2%	14.0%	21.3%	50.0%	3.3%
Highly selective public	6.8%	11.1%	17.0%	24.6%	40.6%	7.0%
Highly selective private	6.1%	10.3%	16.3%	24.1%	43.3%	2.4%
Selective public	9.7%	14.2%	19.6%	25.5%	30.9%	34.4%
Selective private	8.4%	13.0%	18.7%	25.7%	34.3%	8.6%
Nonselective 4-year public	14.5%	18.5%	20.5%	22.8%	23.7%	4.6%
Nonselective 4-year private non-profit	9.1%	13.8%	19.6%	25.9%	31.6%	1.0%
2-year public and non-profit	13.1%	16.8%	20.2%	23.8%	26.1%	35.5%
4-year for-profit	14.9%	18.4%	20.1%	23.0%	23.5%	1.7%
2-year for-profit	17.1%	20.0%	20.9%	21.5%	20.5%	0.7%
Less than two-year colleges	15.0%	18.4%	21.0%	23.3%	22.3%	0.2%
All colleges	10.7%	14.8%	19.3%	24.4%	30.8%	100.0%
Underrepresentation in selective tiers	18.9%	11.7%	3.8%	-2.4%	-12.6%	
C. Counterfactual Distributions Under Need-Affirmative Student Allocations						
Ivy-Plus	11.8%	15.5%	17.2%	21.2%	34.3%	0.9%
Other elite colleges	10.1%	14.1%	17.4%	22.4%	36.0%	3.3%
Highly selective public	9.4%	13.7%	18.5%	24.7%	33.7%	7.0%
Highly selective private	9.0%	13.4%	18.2%	24.3%	35.1%	2.4%
Selective public	10.4%	14.8%	19.8%	25.4%	29.6%	34.4%
Selective private	9.4%	14.0%	19.3%	25.5%	31.9%	8.6%
Nonselective 4-year public	12.7%	16.9%	19.8%	23.0%	27.7%	4.6%
Nonselective 4-year private non-profit	8.7%	12.9%	19.1%	25.6%	33.6%	1.0%
2-year public and non-profit	11.2%	15.0%	19.2%	23.8%	30.8%	35.5%
4-year for-profit	13.1%	17.0%	19.6%	22.9%	27.4%	1.7%
2-year for-profit	13.9%	17.5%	20.0%	22.0%	26.7%	0.7%
Less than two-year colleges	10.9%	16.0%	19.7%	23.1%	30.2%	0.2%
All colleges	10.7%	14.8%	19.3%	24.4%	30.8%	100.0%
Underrepresentation in selective tiers	5.3%	2.5%	-0.1%	-2.3%	-1.2%	

Notes: This table reports parental income distributions by quintile by college tier. Panel A reports actual parent income shares in the analysis sample (1980-1982 birth cohorts). Each cell reports the share of the specified group of colleges that comes from a given parent income quintile. The "Underrepresentation in selective tiers" row reports the percentage difference between the number of students from the relevant parent income quintile pooling the first six tiers and the percentage of students from that quintile pooling all colleges. For example in Panel A, 31.3% = 1-7.3%/10.7%. Panel B repeats Panel A under our income-neutral student allocation counterfactual, allocating students to colleges randomly based on their SAT/ACT scores while holding fixed the distribution of SAT/ACT scores, pre-college states, and race to match the actual distributions at each college. Panel C repeats Panel B after adding 160 points to the SAT/ACT scores of all college goers from the bottom parent income quintile, 128 points to second quintile college goers, 96 points to middle quintile college goers, and 64 points to fourth quintile college goers. See Section IV.A for more details on these counterfactuals. Statistics in this table are constructed directly from the individual-level microdata.

TABLE IV

Fraction of Differences in Earnings Across Colleges Due to Causal Effects

Dep. Var.: College fixed effect, conditional on parent income, race, SAT/ACT, and additional controls

	(1)	(2)	(3)	(4)	(5)	(6)
	Race, Gender, Interacted w/ SAT/ACT	High School FE's	High School FE's Interacted w/ Race	Control for Application Set and HS	Control for Application Set and HS Interacted w/ Race	Bottom Quintile Only
College fixed effect, conditional on parent income, race, and SAT/ACT	1.003	0.907	0.903	0.857	0.850	0.850
	(0.006)	(0.010)	(0.010)	(0.012)	(0.012)	(0.015)
Adj. R-squared	0.968	0.886	0.883	0.889	0.886	0.750
Additional Controls Used to Construct Dependent Variable						
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

Notes: This table reports estimates of the fraction of the differences in mean earnings observed across colleges conditional on parental income, race, and SAT/ACT scores that are due to causal effects, corresponding to the parameter λ in equation (2). The sample comprises all college-goers in our 1980-1982 cohorts who are matched to College Board or ACT data. Each column presents coefficients from univariate OLS regressions run at the college level, weighted by child count, following equation (2). The independent variable in all columns is the college fixed effect obtained from a regression of child earnings rank on college fixed effects, a quintic in parent income percentile, a quintic in SAT/ACT score, an indicator for taking the SAT, an indicator for taking the ACT (as some took both tests), and race/ethnicity indicators, as in equation (1). The dependent variable in each column is the child's college's fixed effect from the same regression, including additional controls. In Column 1, we add a gender indicator, and we fully interact the race, gender, and SAT-quintic. Column 2 adds fixed effects for the child's high school. Column 3 interacts the high school and race indicators. Column 4 replicates Column 2 and controls for the mean SAT score of the colleges to which students sent scores and also the total number of colleges to which the students sent scores, as in Dale and Krueger (2014). Column 5 replicates Column 3, adding the same controls as in Column 4. Column 6 replicates Column 5, restricting attention to children with parents from the bottom quintile. Statistics in this table are constructed directly from the individual-level microdata.

TABLE V

Actual vs. Counterfactual Intergenerational Transition Matrices

	Fraction of Children with Earnings in Each Group					
	Bottom 20% (1)	Quintile 2 (2)	Quintile 3 (3)	Quintile 4 (4)	Top 20% (5)	Top 1% (6)
A. Actual Outcomes						
..for children with parents from						
Bottom 20%	16.0%	21.1%	23.1%	21.7%	18.2%	0.6%
Quintile 2	14.0%	18.1%	22.4%	24.1%	21.4%	0.7%
Quintile 3	12.8%	15.7%	21.0%	25.6%	24.9%	0.9%
Quintile 4	11.5%	13.7%	18.9%	26.3%	29.6%	1.2%
Top 20%	11.1%	11.6%	14.3%	22.8%	40.2%	3.4%
B. Income-Neutral Student Allocations						
..for children with parents from						
Bottom 20%	15.6%	20.3%	22.7%	21.9%	19.5%	0.7%
Quintile 2	13.7%	17.6%	22.1%	24.2%	22.5%	0.9%
Quintile 3	12.7%	15.4%	20.6%	25.4%	25.8%	1.1%
Quintile 4	11.5%	13.7%	18.8%	26.1%	29.9%	1.3%
Top 20%	11.4%	12.2%	15.0%	23.1%	38.3%	3.1%
Share of rich-poor top-quintile outcome gap narrowed	14.6%					
C. Need-Affirmative Student Allocations						
..for children with parents from						
Bottom 20%	15.2%	19.7%	22.2%	22.1%	20.8%	0.9%
Quintile 2	13.5%	17.2%	21.7%	24.2%	23.3%	1.0%
Quintile 3	12.6%	15.4%	20.4%	25.4%	26.2%	1.2%
Quintile 4	11.5%	13.7%	18.7%	26.1%	30.0%	1.4%
Top 20%	11.6%	12.6%	15.6%	23.2%	37.0%	2.9%
Share of rich-poor top-quintile outcome gap narrowed	26.5%					

Notes: Panel A shows the actual intergenerational income transition matrix for college students in our analysis sample (1980-1982 birth cohorts). Each cell of Panel A reports the percentage of college goers with earnings outcomes in the quintile given by the column conditional on having parents with income in the quintile given by the row for the analysis sample. Panels B and C repeat Panel A under the income-neutral student allocation and need-affirmative student allocation counterfactuals, defined in the notes to Table III. Panels B and C assume that 80% of children's earnings differences across colleges reflect causal effects conditional on SAT/ACT scores, race, and parental income. Mechanically, children are randomly assigned the earnings of another child who is observed as attending their counterfactually assigned college and who has the same parent income quintile, race, and SAT/ACT score. After that counterfactual earnings level is calculated, with 80% probability, children are assigned that randomly assigned earning, and with 20% probability, children are assigned their actual earnings. See Appendix F for details. The share of the rich-poor top-quintile outcome gap narrowed equals ((40.2%-18.2%)-(38.3%-19.5%))/(40.2%-18.2%) = 14.6% in Panel B. The corresponding statistic in Panel C is computed similarly. Statistics in this table are constructed directly from the individual-level microdata.

APPENDIX TABLE I

Counts in Administrative vs. Survey Data by Birth Cohort

Birth Cohort	Size of Birth Cohort Based on Vital Stats.	Number of Citizens in Our Sample	Number of 20 Year Olds in CPS	Number of 20 Year Olds in Our Sample	CPS College Attendees	College Attendees at Age 20 in Our Sample
	(1)	(2)	(3)	(4)	(5)	(6)
1980	3,612	3,189	3,840	3,385	1,839	1,526
1981	3,629	3,403	3,829	3,482	1,845	1,601
1982	3,681	3,493	3,938	3,545	1,998	1,689
1983	3,639	3,470	3,926	3,575	2,009	1,794
1984	3,669	3,664	3,981	3,835	2,030	1,952
1985	3,761	3,776	4,222	3,939	2,187	1,987
1986	3,757	3,764	4,057	3,922	2,022	1,986
1987	3,809	3,836	4,006	4,061	2,078	2,080
1988	3,910	3,960	4,007	4,212	2,147	2,175
1989	4,041	4,103	4,087	4,361	2,254	2,316
1990	4,158	4,227	4,399	4,498	2,389	2,415
1991	4,111	4,178	4,281	4,484	2,433	2,402
1980-1991	45,776	45,062	48,573	47,298	25,231	23,922

Notes: This table compares aggregate counts in our administrative data sample to aggregate counts from the National Vital Statistics System and the Current Population Survey (CPS). All counts are reported in thousands. Column 1 reports the size of the birth cohort according to Vital Statistics in each birth cohort. Column 2 lists the number of citizens in the given birth cohort in our administrative data sample. Values in Column 2 can be larger than values in Column 1 because Column 1 excludes naturalized citizens. Column 3 reports the number of people in the CPS who are age 20 in each birth cohort. Column 4 reports analogous counts of 20 year olds in our sample of children linked to parents in the tax data. Column 5 reports the number of people enrolled in college at age 20 in each cohort. Column 6 reports analogous counts in our sample. Statistics in Columns 2, 4, and 6 of this table are constructed directly from the individual-level microdata.

APPENDIX TABLE II

Income Distributions in the Tax Data vs. the American Community Survey

	(1)	(2)	(3)	(4)
Data source	Tax Data	ACS	ACS	ACS
Family Unit	Tax unit	Tax unit	Household	Household
Income definition	Adjusted gross income	Adjusted gross income	Adjusted gross income	Total money income
Panel A. Families of Children aged 15 in 2000				
Count	4,006,698	4,083,218	4,083,218	4,083,218
Mean	93,041	78,456	86,989	91,254
Fraction with zero income	3.4%	5.0%	3.4%	0.4%
10th Percentile	\$12,000	\$9,634	\$15,140	\$21,265
25th Percentile	\$26,500	\$28,903	\$35,785	\$40,602
Median	\$56,750	\$59,183	\$68,817	\$72,809
75th Percentile	\$100,500	\$99,097	\$108,731	\$112,860
90th Percentile	\$156,750	\$151,398	\$163,372	\$165,161
Fraction with married parents	64.0%	64.9%	64.9%	64.9%
Panel B. Families of Children aged 15 in 2006				
Count	4,531,577	4,347,184	4,347,184	4,347,184
Mean	92,635	80,219	86,662	90,779
Fraction with zero income	4.7%	6.3%	4.7%	1.1%
10th Percentile	\$9,500	\$7,055	\$11,758	\$18,225
25th Percentile	\$23,250	\$27,161	\$32,923	\$37,626
Median	\$50,000	\$58,791	\$66,434	\$70,549
75th Percentile	\$96,750	\$104,130	\$111,702	\$114,642
90th Percentile	\$158,250	\$160,499	\$168,141	\$170,493
Fraction with married parents	59.5%	62.5%	62.5%	62.5%

Notes: This table compares the parental income distributions of children aged 15 in 2000 (Panel A) and in 2006 (Panel B) in the tax data vs. American Community Survey (ACS) data. Column 1 presents statistics from the tax data, where the unit of observation for family income is the tax unit (married parents or single parent) and income is defined as Adjusted Gross Income, which is pre-tax and pre-transfer cash income. Column 2 replicates this in the ACS data, excluding the few children who are heads of household or spouses of heads of household. For children living with married parents, we sum the income of the two parents (if both parents' incomes are zero, we use instead the income of the head of household as the head would most likely claim the child for tax purposes in this case). For children not living with two married parents, we take the income of the mother if present and non-zero and father if the mother's income is zero or the mother is absent. If both father and mother have zero income or are absent, we define the child's parent as the head of household. Column 3 considers adjusted gross income summed across all household members aged 15 or older (instead of just parents). Column 4 considers total household money income (instead of adjusted gross income). Total money income is the standard income definition used in the ACS and is broader than adjusted gross income, as it includes cash government transfers, retirement and disability benefits. All dollar values are expressed in 2015 dollars, adjusting for inflation using the CPI-U. The counts in the first row are actual counts for the tax data and implied population counts corresponding to the ACS sample based on the sampling weights. Statistics in Column 1 of this table are constructed directly from the individual-level microdata.

APPENDIX TABLE III
Additional Summary Statistics for Analysis Sample

	Sample		
		College-Goers in Data Release	
	All Children in 1980-82 cohorts	80-82 cohorts only	Including data imputed from 83-84 cohorts
	(1)	(2)	(3)
A. College Attendance Rates			
% Attending College Between Age 19-22	61.83	-	-
% Attending a College in Data Release (based on 80-82 cohorts)	53.07	-	-
% Attending an Ivy-Plus College	0.49	0.95	0.84
% Attending an Other Elite College	1.71	3.31	3.02
% Attending an Other 4-year College	31.59	59.63	58.08
% Attending a 2-Year or Less College	19.28	36.11	38.06
% Not Attending any College by Age 28	26.65	-	-
B. Parents' Household Income (When Child is Aged 15-19)			
Mean Income ($)	87,335	117,080	114,306
Median Income ($)	59,100	77,100	N/A
% with Parents in Bottom 20%	20.00	10.63	11.12
% with Parents in Top 20%	20.00	30.93	29.92
% with Parents in Top 1%	1.00	1.70	1.62
C. Children's Individual Earnings (in 2014, Ages 32-34)			
Mean Earnings ($)	35,526	47,048	46,179
Median Earnings ($)	26,900	35,800	N/A
% Employed	81.68	88.72	88.60
% in Top 20%	20.00	29.66	28.87
% in Top 1%	1.00	1.73	1.63
% in Top 20% | Parents in Bottom 20%	8.65	18.33	17.44
% in Top 1% | Parents in Bottom 20%	0.22	1.00	0.92
% in Top 20% and Parents in Bottom 20%	1.73	1.95	1.94
% in Top 1% and Parents in Bottom 20%	0.04	0.07	0.06
Number of Children	10,757,269	5,535,694	6,244,162
Percentage of College Students Covered	-	83.2%	93.9%

Notes: The table presents additional summary statistics. Column 1 includes all children in the 1980-82 birth cohorts and replicates Column 1 of Table I. Column 2 limits this sample to students who attend a college (between the ages of 19-22) that is included in the public data release using data purely from the 1980-82 birth cohorts. This is the set of colleges for which we observe a sufficient number of students and have complete attendance records for the 1980-82 cohorts, as described in Section II and Appendix B. Column 3 adds imputed data from the 1983-84 birth cohorts for colleges with insufficient data in the 1980-82 birth cohorts (see Section II.D of Chetty et al. (2017) for details), replicating Column 2 of Table I. This is the sample used for most of our analyses. See notes to Table I for definitions. Statistics in Column 1 are constructed based on Online Data Table 6 and statistics in Columns 2 and 3 are based on Online Data Table 2, with the exception of median income and earnings, which are constructed directly from the individual-level microdata.

APPENDIX TABLE IV
Parent Income Distributions by SAT/ACT Score for College Students

	Parent Income Quintile					Share of all college goers
	1	2	3	4	5	
	(1)	(2)	(3)	(4)	(5)	(6)
A. Share of SAT/ACT Bin from Each Parent Income Quintile						
1500-1600	2.5%	4.7%	8.8%	16.9%	67.2%	0.6%
1400-1490	3.2%	6.4%	12.4%	21.1%	56.8%	2.3%
1300-1390	4.0%	7.7%	14.2%	23.3%	50.9%	5.3%
1200-1290	5.0%	9.4%	16.3%	25.1%	44.2%	10.4%
1100-1190	6.4%	11.4%	18.5%	26.4%	37.4%	16.7%
1000-1090	8.5%	13.4%	19.9%	26.6%	31.7%	19.8%
900-990	11.3%	16.1%	21.0%	25.7%	25.8%	19.2%
800-890	15.4%	19.3%	21.6%	23.4%	20.2%	13.7%
700-790	21.0%	23.1%	21.1%	20.0%	14.8%	7.7%
600-690	25.8%	26.1%	20.3%	16.8%	11.1%	3.1%
500-590	30.8%	28.0%	18.9%	13.9%	8.5%	1.1%
400-490	34.5%	28.3%	18.7%	11.1%	7.4%	0.2%
B. Share of Cumulative SAT/ACT Bin from Each Parent Income Quintile						
≥1500	2.5%	4.7%	8.8%	16.9%	67.2%	0.6%
≥1400	3.1%	6.1%	11.7%	20.3%	58.9%	2.9%
≥1300	3.7%	7.1%	13.3%	22.2%	53.7%	8.2%
≥1200	4.4%	8.4%	15.0%	23.8%	48.4%	18.5%
≥1100	5.3%	9.8%	16.6%	25.0%	43.2%	35.3%
≥1000	6.5%	11.1%	17.8%	25.6%	39.0%	55.0%
≥900	7.7%	12.4%	18.6%	25.6%	35.6%	74.2%
≥800	8.9%	13.5%	19.1%	25.3%	33.2%	87.9%
≥700	9.9%	14.3%	19.3%	24.8%	31.7%	95.6%
≥600	10.4%	14.6%	19.3%	24.6%	31.1%	98.7%
≥500	10.6%	14.8%	19.3%	24.5%	30.8%	99.8%
≥400	10.7%	14.8%	19.3%	24.4%	30.8%	100.0%

Notes: Panel A reports the parent income distribution by SAT/ACT score bin among college goers in our analysis sample. SAT scores for 47.6% of college goers are obtained directly from the College Board; ACT scores for another 26.2% of college goers are obtained from ACT and converted to an SAT score. We impute an SAT/ACT score for the other 26.2% of college-goers using the SAT/ACT score of the student from the same parent income quintile, same pre-college state, and same college tier with the nearest child earnings and non-missing race. Each cell of Columns 1-5 reports the share of students in a given SAT/ACT bin who have parents in the parent income quintile defined in the column heading. The sixth column reports the total share of college goers who fall into the corresponding SAT/ACT bin. Panel B weights the distributional statistics in Columns 1-5 by the overall college-goer shares reported in Column 6 to compute the joint cumulative distribution function of SAT/ACT scores and parent income. For example, Panel B reports that 3.6% of college goers with an SAT/ACT score of at least 1300 have bottom-quintile parents. Appendix Table V shows that similar results are obtained when excluding college goers with an imputed SAT/ACT score, while Appendix Table VI shows that similar results are obtained using data from the National Postsecondary Student Aid Study. Statistics in this table are constructed directly from the individual-level microdata.

APPENDIX TABLE V

Parent Income Distributions by SAT/ACT Score, Excluding Students with Imputed SAT/ACT Scores

	Parent Income Quintile					Share of all college goers
	1	2	3	4	5	
	(1)	(2)	(3)	(4)	(5)	(6)
A. Share of SAT/ACT Bin from Each Parent Income Quintile						
1500-1600	2.2%	4.4%	8.4%	16.8%	68.1%	0.8%
1400-1490	2.8%	5.9%	11.8%	20.9%	58.6%	2.9%
1300-1390	3.3%	6.9%	13.5%	23.1%	53.2%	6.5%
1200-1290	4.0%	8.2%	15.3%	25.1%	47.4%	12.1%
1100-1190	5.0%	9.9%	17.3%	26.7%	41.1%	18.4%
1000-1090	6.6%	11.6%	18.8%	27.2%	35.8%	20.3%
900-990	8.8%	14.1%	20.2%	26.9%	30.0%	18.2%
800-890	12.2%	17.2%	21.1%	25.2%	24.2%	11.8%
700-790	16.8%	21.1%	21.3%	22.3%	18.6%	6.0%
600-690	20.7%	24.2%	20.9%	19.5%	14.8%	2.2%
500-590	25.2%	26.3%	20.0%	16.8%	11.8%	0.7%
400-490	27.7%	27.8%	19.8%	13.9%	10.8%	0.1%
B. Share of Cumulative SAT/ACT Bin from Each Parent Income Quintile						
≥1500	2.2%	4.4%	8.4%	16.8%	68.1%	0.8%
≥1400	2.7%	5.6%	11.1%	20.1%	60.5%	3.6%
≥1300	3.1%	6.4%	12.6%	22.0%	55.9%	10.2%
≥1200	3.6%	7.4%	14.1%	23.7%	51.2%	22.3%
≥1100	4.2%	8.5%	15.5%	25.1%	46.6%	40.6%
≥1000	5.0%	9.6%	16.6%	25.8%	43.0%	60.9%
≥900	5.9%	10.6%	17.4%	26.0%	40.0%	79.2%
≥800	6.7%	11.5%	17.9%	25.9%	38.0%	91.0%
≥700	7.3%	12.1%	18.1%	25.7%	36.8%	97.0%
≥600	7.6%	12.3%	18.2%	25.6%	36.3%	99.2%
≥500	7.8%	12.4%	18.2%	25.5%	36.1%	99.9%
≥400	7.8%	12.4%	18.2%	25.5%	36.1%	100.0%

Notes: The table replicates Appendix Table IV, omitting college goers with an imputed SAT/ACT score. See the notes to that table for details. Statistics in this table are constructed directly from the individual-level microdata.

APPENDIX TABLE VI

Parent Income Distributions by SAT/ACT Score Using NPSAS Data

	Parent Income Quintile					Share of all college goers
	1	2	3	4	5	
	(1)	(2)	(3)	(4)	(5)	(6)
A. Share of SAT/ACT Bin from Each Parent Income Quintile						
1500-1600	2.6%	6.8%	16.1%	20.4%	54.0%	0.9%
1400-1490	5.1%	7.2%	12.9%	27.0%	47.9%	3.1%
1300-1390	3.7%	7.1%	14.5%	25.2%	49.4%	7.2%
1200-1290	5.9%	11.1%	16.1%	25.7%	41.2%	12.2%
1100-1190	7.1%	13.1%	18.1%	26.1%	35.6%	18.4%
1000-1090	8.3%	13.1%	22.0%	26.4%	30.2%	20.3%
900-990	15.8%	19.6%	20.2%	21.9%	22.4%	19.8%
800-890	18.0%	19.8%	22.8%	23.1%	16.3%	11.5%
700-790	20.1%	22.6%	19.6%	19.5%	18.1%	5.2%
600-690	18.8%	18.7%	23.2%	20.2%	19.0%	1.2%
500-590	22.4%	21.7%	11.7%	24.2%	20.0%	0.3%
400-490	44.8%	17.4%	4.7%	33.1%	0.0%	0.0%
B. Share of Cumulative SAT/ACT Bin from Each Parent Income Quintile						
≥1500	2.6%	6.8%	16.1%	20.4%	54.0%	0.9%
≥1400	4.5%	7.1%	13.6%	25.6%	49.2%	4.0%
≥1300	4.0%	7.1%	14.2%	25.4%	49.4%	11.2%
≥1200	5.0%	9.2%	15.2%	25.5%	45.1%	23.4%
≥1100	5.9%	10.9%	16.5%	25.8%	40.9%	41.8%
≥1000	6.7%	11.6%	18.3%	26.0%	37.4%	62.1%
≥900	8.9%	13.6%	18.7%	25.0%	33.8%	81.9%
≥800	10.0%	14.3%	19.2%	24.8%	31.7%	93.4%
≥700	10.6%	14.8%	19.3%	24.5%	30.9%	98.5%
≥600	10.7%	14.8%	19.3%	24.4%	30.8%	99.7%
≥500	10.7%	14.8%	19.3%	24.4%	30.8%	100.0%
≥400	10.7%	14.8%	19.3%	24.4%	30.8%	100.0%

Notes: This table replicates Appendix Table IV using data from the National Postsecondary Student Aid Study (NPSAS) instead of the tax data. The NPSAS contains coarse information on college-goers' parent income and SAT or ACT score. To overcome this problem, we norm NPSAS parent income to match the true distribution of college-goers' parent income quintiles from our analysis sample in the tax data, convert ACT scores to SAT scores, and use parent income and tier to impute missing SAT/ACT scores. Specifically, we use information gleaned from FAFSA AGI and survey questions to generate an observed distribution of parent income within tier and SAT/ACT quartile or missing SAT/ACT score. We then randomly assign incomes to students with unobserved parent income to match the observed distribution within these cells. Next, we assign parental income quintiles to this NPSAS income variable such that the quintile distribution matches that from our main analysis sample. Finally, we impute missing SAT/ACT scores such that the distribution of observed SAT/ACT scores within parent income quintile and tier is preserved. See the notes to Appendix Table IV for details on the statistics reported in this table. Statistics in this table are constructed directly from the individual-level microdata.

APPENDIX TABLE VII

Ivy-Plus Attendance Rate by Parent Income Group and SAT/ACT Score

A. Share attending an Ivy-Plus college with SAT/ACT score…					
	1200	1300	1400	1500	1600
Bottom 20%	0.8%	2.8%	7.3%	21.4%	60.0%
Quintile 2	0.7%	2.2%	4.7%	19.3%	43.9%
Quintile 3	0.7%	2.0%	4.5%	14.5%	54.3%
Quintile 4	0.5%	1.7%	4.4%	15.6%	41.7%
Top 20%	1.0%	3.8%	10.8%	30.9%	60.5%
P80-P90	0.6%	2.0%	5.5%	17.0%	46.0%
P90-P95	0.9%	2.9%	8.4%	24.9%	50.0%
P95-P99	1.4%	5.0%	14.6%	38.8%	71.0%
P99-P100	2.6%	10.0%	26.2%	53.8%	76.7%
B. Share attending an Ivy-Plus college within SAT/ACT score range…					
	1200-1290	1300-1390	1400-1490	1500-1590	1600
Bottom 20%	1.3%	3.9%	11.3%	29.7%	60.0%
Quintile 2	1.1%	3.2%	8.6%	27.3%	43.9%
Quintile 3	1.0%	2.8%	7.6%	22.1%	54.3%
Quintile 4	0.8%	2.6%	7.2%	23.5%	41.7%
Top 20%	1.7%	5.9%	16.4%	38.8%	60.5%
P80-P90	0.9%	3.1%	8.7%	24.6%	46.0%
P90-P95	1.3%	4.5%	12.6%	32.4%	50.0%
P95-P99	2.4%	8.0%	21.5%	45.3%	71.0%
P99-P100	4.8%	14.6%	34.4%	60.7%	76.7%
C. Share attending an Ivy-Plus college with at least SAT/ACT score…					
	1200+	1300+	1400+	1500+	1600
Bottom 20%	3.4%	7.0%	14.5%	31.0%	60.0%
Quintile 2	2.9%	5.7%	11.7%	27.9%	43.9%
Quintile 3	2.6%	5.0%	10.0%	23.5%	54.3%
Quintile 4	2.5%	5.0%	10.1%	24.4%	41.7%
Top 20%	6.7%	12.0%	21.9%	40.0%	60.5%
P80-P90	3.3%	6.2%	12.0%	25.7%	46.0%
P90-P95	5.2%	9.3%	17.3%	33.4%	50.0%
P95-P99	9.8%	16.3%	28.0%	46.8%	71.0%
P99-P100	17.4%	26.8%	42.0%	61.6%	76.7%

Notes: This table shows the Ivy-Plus attendance rate for college-goers in our analysis sample by parent income group and SAT/ACT score. In addition to parent income quintiles, the top 20% is broken down further into 80-90th percentiles, 90-95th percentiles, 95-99th percentiles, and the top 1%. Panel A reports the Ivy-Plus attendance rate for individual SAT/ACT scores. For example, of all college-goers with parents in the bottom 20% and with a 1600 SAT/ACT score, 60.0% attended an Ivy-Plus college. Panel B repeats Panel A for 100-point SAT/ACT score ranges. Panel C reports the Ivy-Plus attendance rate for those with at least a certain SAT/ACT score. See notes to Appendix Table IV for details about SAT/ACT scores and imputation. See Table I for a definition of Ivy-Plus colleges. Statistics in this table are constructed directly from the individual-level microdata.

APPENDIX TABLE VIII

Counterfactual Income Segregation across Colleges for College-Goers

A. Actual Income Segregation across Colleges

	(1)	(2)	(3)	(4)	(5)
	Fraction of college peers from each income group..				
	Bottom 20%	Quintile 2	Quintile 3	Quintile 4	Top 20%
..for children from					
Bottom 20%	15.7%	18.6%	20.6%	22.6%	22.5%
Quintile 2	13.5%	17.3%	20.6%	23.8%	24.8%
Quintile 3	11.5%	15.9%	20.5%	24.9%	27.2%
Quintile 4	10.0%	14.5%	19.7%	25.6%	30.2%
Top 20%	7.9%	12.1%	17.1%	24.0%	38.8%

B. Income-Neutral Student Allocations

	Fraction of college peers from each income group..				
	Bottom 20%	Quintile 2	Quintile 3	Quintile 4	Top 20%
..for children from					
Bottom 20%	12.8%	16.5%	19.5%	23.3%	27.8%
Quintile 2	11.9%	15.8%	19.6%	23.9%	28.8%
Quintile 3	10.8%	15.0%	19.6%	24.6%	30.0%
Quintile 4	10.2%	14.5%	19.4%	25.0%	30.9%
Top 20%	9.6%	13.9%	18.8%	24.6%	33.2%

C. Need-Affirmative Student Allocations

	Fraction of college peers from each income group..				
	Bottom 20%	Quintile 2	Quintile 3	Quintile 4	Top 20%
..for children from					
Bottom 20%	12.2%	15.9%	19.2%	23.3%	29.3%
Quintile 2	11.4%	15.4%	19.4%	23.9%	29.8%
Quintile 3	10.6%	14.9%	19.5%	24.6%	30.4%
Quintile 4	10.2%	14.5%	19.4%	24.9%	31.0%
Top 20%	10.1%	14.4%	19.1%	24.6%	31.8%

Notes: Panel A presents parental income segregation measures across colleges. The sample includes all children in our analysis sample (1980-1982 birth cohorts). Each row corresponds to a group of children based on their own parents' incomes. For each row, each column reports the average composition of peers in the 1980-1982 birth cohorts in the same college using the same parent income quintile groups. Peer composition is computed using leave-out means. Panels B and C replicate Panel A under the two counterfactuals discussed in the notes to Table III. Statistics in this table are constructed directly from the individual-level microdata.

APPENDIX TABLE IX
Parental Income Distributions Under Counterfactual Student Allocation Rules without Racial and Geographic Constraints

	Parent Income Quintile					Share of all college goers
	1 (Bottom 20%)	2	3	4	5 (Top 20%)	
	(1)	(2)	(3)	(4)	(5)	(6)
A. Actual Distributions						
Ivy-Plus	3.8%	5.7%	8.7%	13.4%	68.4%	0.9%
Other elite colleges	4.3%	6.8%	10.2%	15.8%	62.8%	3.3%
Highly selective public	5.5%	9.2%	14.3%	23.4%	47.6%	7.0%
Highly selective private	4.1%	7.6%	12.2%	19.7%	56.5%	2.4%
Selective public	8.4%	12.9%	18.6%	26.1%	34.1%	34.4%
Selective private	7.1%	12.0%	18.2%	25.5%	37.2%	8.6%
Nonselective 4-year public	17.0%	20.4%	22.1%	22.7%	17.7%	4.6%
Nonselective 4-year private non-profit	10.7%	14.7%	19.8%	24.6%	30.2%	1.0%
2-year public and non-profit	14.6%	18.6%	22.2%	24.7%	19.9%	35.5%
4-year for-profit	17.8%	22.3%	22.5%	21.1%	16.3%	1.7%
2-year for-profit	21.5%	23.9%	23.1%	19.5%	12.0%	0.7%
Less than two-year colleges	20.7%	23.2%	21.3%	21.0%	13.8%	0.2%
All colleges	10.7%	14.8%	19.3%	24.4%	30.8%	100.0%
Underrepresentation in selective tiers	31.3%	21.6%	11.4%	-0.6%	-27.9%	
B. Counterfactual Distributions Under Income-Neutral Student Allocations						
Ivy-Plus	3.8%	7.1%	12.6%	21.6%	54.9%	0.9%
Other elite colleges	5.1%	9.2%	15.4%	24.2%	46.1%	3.3%
Highly selective public	7.1%	11.4%	17.6%	25.0%	38.9%	7.0%
Highly selective private	6.6%	11.0%	17.7%	25.1%	39.6%	2.4%
Selective public	9.5%	14.0%	19.2%	25.0%	32.3%	34.4%
Selective private	9.2%	13.5%	18.9%	24.9%	33.5%	8.6%
Nonselective 4-year public	13.1%	16.9%	20.4%	23.9%	25.7%	4.6%
Nonselective 4-year private non-profit	10.5%	14.9%	19.4%	24.8%	30.3%	1.0%
2-year public and non-profit	13.2%	17.1%	20.2%	23.9%	25.6%	35.5%
4-year for-profit	13.2%	17.2%	20.0%	23.9%	25.6%	1.7%
2-year for-profit	16.0%	19.1%	20.4%	22.5%	21.9%	0.7%
Less than two-year colleges	14.2%	18.3%	20.2%	23.1%	24.1%	0.2%
All colleges	10.7%	14.8%	19.3%	24.4%	30.8%	100.0%
Underrepresentation in selective tiers	18.5%	11.7%	3.7%	-1.9%	-12.9%	
C. Counterfactual Distributions Under Need-Affirmative Student Allocations						
Ivy-Plus	11.3%	16.8%	19.9%	23.9%	28.0%	0.9%
Other elite colleges	10.3%	14.7%	19.6%	24.0%	31.4%	3.3%
Highly selective public	10.1%	14.5%	19.3%	24.9%	31.2%	7.0%
Highly selective private	9.9%	14.5%	19.3%	24.9%	31.4%	2.4%
Selective public	10.4%	14.6%	19.3%	24.8%	30.8%	34.4%
Selective private	10.4%	14.6%	19.3%	24.7%	31.0%	8.6%
Nonselective 4-year public	11.1%	15.1%	19.2%	24.1%	30.6%	4.6%
Nonselective 4-year private non-profit	10.7%	14.7%	18.8%	24.8%	31.0%	1.0%
2-year public and non-profit	11.0%	15.0%	19.2%	24.1%	30.7%	35.5%
4-year for-profit	10.9%	15.3%	19.0%	24.0%	30.7%	1.7%
2-year for-profit	11.4%	15.3%	19.0%	23.5%	30.8%	0.7%
Less than two-year colleges	11.5%	14.9%	18.8%	24.2%	30.6%	0.2%
All colleges	10.7%	14.8%	19.3%	24.4%	30.8%	100.0%
Underrepresentation in selective tiers	2.7%	1.0%	-0.3%	-1.2%	-0.3%	

Notes: This table replicates Table III, but reallocates students to colleges randomly conditional on their SAT/ACT scores (or adjusted SAT/ACT scores), without holding fixed the racial composition or pre-college-state distribution of the student body. See notes to Table III for details. Statistics in this table are constructed directly from the individual-level microdata.

APPENDIX TABLE X

Replication of Specifications in Dale and Krueger (2014)

Dep. Var.: Log earnings	All Quintiles	All Quintiles, College Application Set FEs	Bottom Quintile Only	Quintile 2 Only	Quintile 3 Only	Quintile 4 Only	Top Quintile Only
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Average SAT of College Attended	0.016	0.012	0.038	0.032	0.025	0.016	0.010
	(0.002)	(0.003)	(0.009)	(0.007)	(0.005)	(0.004)	(0.003)
Adj. R-squared	0.060	0.099	0.079	0.072	0.062	0.052	0.025

Notes: This table replicates specifications in Dale and Krueger (2014) estimating the relationship between students' earnings outcomes and college selectivity, as measured by students' average SAT scores. It uses the sample of students attending one of the 31 colleges for which we have data in the College and Beyond Survey used in Dale and Krueger (2002, 2014), restricting to students with earnings that exceed $13,822 in 2007 dollars ($15,800 in 2015 dollars). Column 1 reports estimates from a regression of log earnings on the average SAT score of college attended, controlling for a quintic in parent income rank and a quintic in SAT. Column 2 adds fixed effects for the exact set of these 31 schools that each student sent scores to. Columns 3-7 replicate Column 1 using only children from the parent income quintile specified in the column title. Robust standard errors are in parentheses. Statistics in this table are constructed directly from the individual-level microdata. See Appendix G for further details.

APPENDIX TABLE XI

Counterfactual Gaps in Mobility Rates: Sensitivity to Assumptions about Colleges' Causal Effects

Assumed causal share of differences in earnings across colleges, conditional on SAT/ACT scores and parent income (%)	Increment to the SAT Scores of Bottom-Quintile College-Goers				
	0	100	160	200	300
100	18.2%	26.7%	33.1%	37.6%	46.2%
90	16.4%	24.1%	29.8%	33.8%	41.7%
80 (baseline)	14.6%	21.4%	26.5%	30.1%	37.1%
70	12.8%	18.8%	23.2%	26.4%	32.5%
60	11.0%	16.1%	20.0%	22.7%	27.9%
50	9.3%	13.5%	16.7%	18.9%	23.3%
40	7.5%	10.9%	13.4%	15.2%	18.7%
30	5.7%	8.2%	10.1%	11.5%	14.1%
20	3.9%	5.6%	6.9%	7.8%	9.5%
10	2.1%	3.0%	3.6%	4.0%	4.9%
0	0.3%	0.3%	0.3%	0.3%	0.3%

Notes: Consider the difference between the fraction of college students with parents in the bottom vs. top quintile who reach the top earnings quintile. Each cell reports the fraction of this gap that is closed under alternative assumptions about colleges' student allocations rules (columns) and colleges' causal effects (rows). The columns vary the increment to the SAT/ACT scores of bottom-quintile college-goers, with smaller proportional increments to second-quintile (80% of the bottom-quintile constant), third-quintile (60%), and fourth-quintile (40%) college-goers. The first column (0 addition) corresponds to the income-neutral student allocations counterfactual reported in Table V; the third column (160 points) corresponds to the need-affirmative student allocations counterfactual. The rows vary the share of the differences in earnings ranks across colleges conditional on SAT/ACT scores and parental income quintile that are assumed to be causal (λ). The estimates in the third row assume that λ=80% and replicate the analysis in Table V. In the remaining rows, for each assumed causal share λ, we report a weighted average with θ weight on the counterfactual earnings distribution that assumes 100%-causality and *1-* λ weight on the actual (0%-causal) distribution of child earnings. We then recompute quintile earnings thresholds so that each child earnings quintile has 20% of children, taking as given the outcomes of non-college-goers. Statistics in this table are constructed directly from the individual-level microdata.

APPENDIX TABLE XII

Actual vs. Counterfactual Intergenerational Transition Matrices with Heterogeneous Causal Effects

	Fraction of Children with Earnings in Each Group					
	Bottom 20% (1)	Quintile 2 (2)	Quintile 3 (3)	Quintile 4 (4)	Top 20% (5)	Top 1% (6)
A. Actual Outcomes						
..for children with parents from						
Bottom 20%	16.0%	21.1%	23.1%	21.7%	18.2%	0.6%
Quintile 2	14.0%	18.1%	22.4%	24.1%	21.4%	0.7%
Quintile 3	12.8%	15.7%	21.0%	25.6%	24.9%	0.9%
Quintile 4	11.5%	13.7%	18.9%	26.3%	29.6%	1.2%
Top 20%	11.1%	11.6%	14.3%	22.8%	40.2%	3.4%
B. Income-Neutral Student Allocations						
..for children with parents from						
Bottom 20%	15.6%	20.4%	22.7%	22.0%	19.4%	0.7%
Quintile 2	13.8%	17.7%	22.2%	24.2%	22.2%	0.8%
Quintile 3	12.7%	15.5%	20.7%	25.6%	25.5%	1.0%
Quintile 4	11.4%	13.7%	18.9%	26.2%	29.8%	1.3%
Top 20%	11.4%	12.1%	14.9%	22.9%	38.8%	3.3%
Share of rich-poor top-quintile outcome gap narrowed	11.7%					
C. Need-Affirmative Student Allocations						
..for children with parents from						
Bottom 20%	15.3%	19.8%	22.3%	22.1%	20.5%	0.9%
Quintile 2	13.6%	17.3%	21.9%	24.3%	22.8%	0.9%
Quintile 3	12.6%	15.5%	20.6%	25.6%	25.7%	1.1%
Quintile 4	11.5%	13.7%	18.8%	26.2%	29.8%	1.3%
Top 20%	11.6%	12.4%	15.3%	22.9%	37.9%	3.1%
Share of rich-poor top-quintile outcome gap narrowed	21.3%					

Notes: This table replicates Table V, but varies the causal effect of colleges on children's earnings based on parental income quintile, tier attended, and counterfactually assigned tier so that children from lower-income families gain more from attending more selective colleges. The counterfactual allocation of students to colleges is the same as in Table V. Children's earnings, however, are assigned differently. Mechanically, children are first randomly assigned the earnings of another child who is observed as attending their counterfactually assigned college and who has the same parent income quintile, race, and SAT/ACT score. After that counterfactual earnings level is calculated, children whose parents are in the bottom 20%, attend a school in one of the six most selective tiers (first six rows of Table II), and are counterfactually assigned to a school in one of the six most selective tiers have a causal share of 40%. This means the earnings outcome is with 40% probability their counterfactually assigned earnings and with 60% probability it is their empirically observed earnings. Children whose parents are not in the bottom 20%, attend a school in one of the six most selective tiers, and are counterfactually assigned to a school in one of the six most selective tiers have a causal share of 0%. This means that those children's counterfactual earnings is the same as their observed earnings. All other children—regardless of parental income quintile and observed and counterfactual tier—have a causal share of 80%. See the notes to Table V and Section IV.B for more details on the counterfactuals and earnings allocation. Statistics in this table are constructed directly from the individual-level microdata.

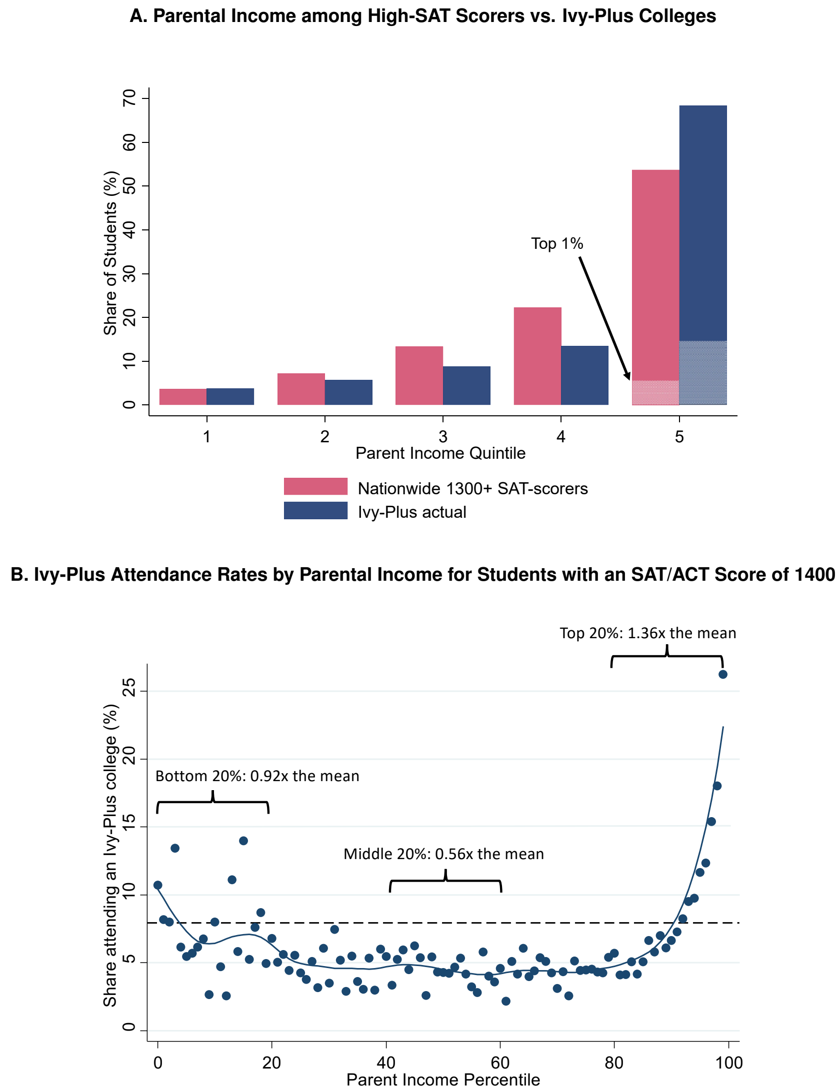
APPENDIX TABLE XIII

Relationship between SAT/ACT Scores and Earnings at Ages 32-34

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
A. Dep. Var.: Individual Earnings in 2014 ($)									
SAT/ACT Score (100 points)	6,744	5,941	5,617	5,307	2,732	5,414	3,944	4,437	3,015
	(21)	(20)	(21)	(21)	(20)	(709)	(214)	(89)	(29)
B. Dep. Var.: Individual Income Rank (Percentiles)									
SAT/ACT Score (100 points)	2.73	2.41	2.24	2.23	1.27	1.26	1.23	1.45	1.44
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.12)	(0.06)	(0.03)	(0.01)
N	4,180,853	4,180,853	4,180,853	4,127,173	4,180,853	51,843	179,723	379,202	1,767,357
Indicator for SAT, ACT, or both taken	X	X	X	X	X	X	X	X	X
Cubic in Parent Income Rank		X	X	X	X	X	X	X	X
Interactions of cubic in Parent Rank, Race, and Gender			X	X	X	X	X	X	X
High School FE's				X					
College FE's					X	X	X	X	X
Restricted to Colleges in Tier						1	2	3	5

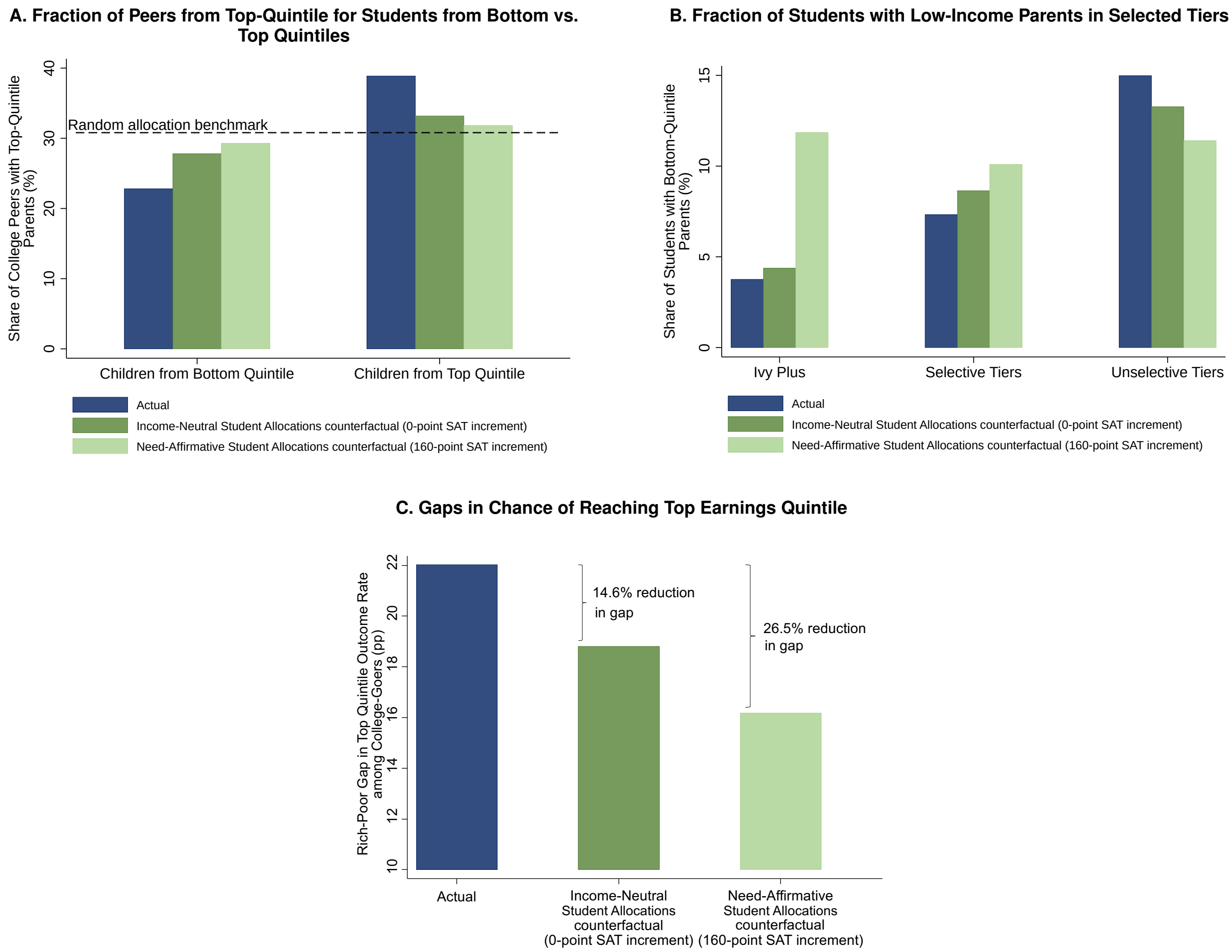
Notes: This table reports estimates from OLS regressions of students' earnings in 2014 on standardized test scores (SAT and ACT). The sample includes all college-goers in our 1980-1982 cohorts for whom we have either SAT or ACT scores. We convert ACT scores to the SAT 1600-point scale. Coefficients reported are multiplied by 100 so that they can be interpreted as the effect of a 100 point increase in the SAT score on the outcome. In Panel A, the left-hand side variable is individual wage earnings (2015 dollars), winsorized at $0 and $1 million; in Panel B, the left-side variable is individual income rank. Each column in each panel reports the coefficient on test scores from a different regression. In Column 1, we regress the outcome on only test scores and an indicator for whether the student took the SAT, the ACT, or both. Column 2 adds a cubic polynomial in parent income rank. Column 3 adds interactions between the parent income cubic, race, and gender. Columns 4 and 5 add high school and college fixed effects respectively. Columns 6, 7, 8, and 9 replicate column 5, restricting the sample to students attending colleges in tiers 1 (Ivy-Plus), 2 (other elite colleges), 3 (highly selective public), and 5 (selective public) respectively. Statistics in this table are constructed directly from the individual-level microdata.

FIGURE I: Ivy-Plus Attendance Rates and SAT Scores by Parental Income



Notes: Panel A plots two series: the parent income distribution of college students nationwide with an SAT/ACT score of at least 1300 (the 93rd percentile), and the parent income distribution of students attending any of the 12 Ivy-Plus colleges, which include the eight Ivy-League colleges as well as the University of Chicago, Stanford University, MIT, and Duke University. See Appendix Table IV for analogous statistics at other SAT/ACT thresholds. See Table III for the parent income distributions of tiers other than the Ivy-Plus. Panel B plots Ivy-Plus college attendance rates by parental income percentile for students with a 1400 SAT/ACT score, the modal and median test score among Ivy-Plus students. The plotted line is an unweighted lowess curve fit through the 100 plotted data points. The dashed horizontal line is the average Ivy-Plus attendance rate for college students with a 1400 SAT/ACT score. See Appendix Table VII and Appendix Figure III for analogous statistics on attendance rates at other test score thresholds. SAT scores for 47.6% of college goers are obtained directly from the College Board; composite test scores for another 26.2% of college goers are obtained from ACT and converted to an SAT score. We impute an SAT/ACT score to the other 26.2% of college-goers using the SAT/ACT score of the student from the same parent income quintile and same college tier with the nearest child earnings. This figure is constructed directly from the individual-level microdata.

FIGURE II: Impacts of Counterfactuals on Income Segregation and Intergenerational Mobility

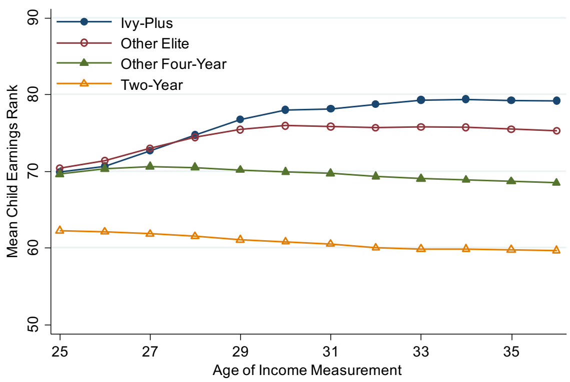


Notes: This figure shows how the income-neutral and need-affirmative student allocation counterfactuals affect income segregation across colleges and intergenerational mobility. The income-neutral counterfactual allocates students to colleges randomly based on their SAT/ACT scores while holding fixed the distribution of SAT/ACT scores, race, and pre-college states to match the empirical distribution at each college. The need-affirmative student allocations counterfactual replicates the income-neutral counterfactual after adding 160 points to the SAT/ACT scores of all college goers from the bottom parent income quintile, 128 points to second quintile college goers, 96 points to third quintile college goers, and 64 points to fourth quintile college goers. See Section IV.A for details on these counterfactuals. Panel A plots the fraction of college peers from the top quintile among college students with parents in the bottom quintile (left triplet of bars) and the top quintile (right triplet of bars) as observed in our data and under the two counterfactuals. These statistics are based on the subset of students who attend college in our analysis sample (i.e., excluding those who do not attend college). The dashed horizontal line shows the fraction of college students who come from the top quintile, which is the fraction of top-quintile peers one would observe if students were randomly allocated to colleges. See Appendix Table VIII for additional statistics on peer exposure across colleges. Panel B plots the fraction of students from the bottom parental income quintile in actuality and under the two counterfactuals at Ivy-Plus colleges, all Selective colleges, and all Unselective colleges. Selective tiers comprise the top six tiers listed in Table II, while Unselective tiers comprise the remaining six tiers. Panel C plots the gap (percentage-point difference) in the fraction of children who reach the top quintile between top-parent-income-quintile college-goers and bottom-parent-income-quintile college-goers in actuality and under the two counterfactuals. Brackets denote the share of the gap narrowed under each counterfactual. The calculations in Panel C assume that 80% of children's earnings differences across colleges reflect causal effects conditional on SAT/ACT scores, parental income, and race; see section IV.B for details. In Appendix Table XI, we report results under alternative assumptions about the causal share. This figure is constructed directly from the individual-level microdata.

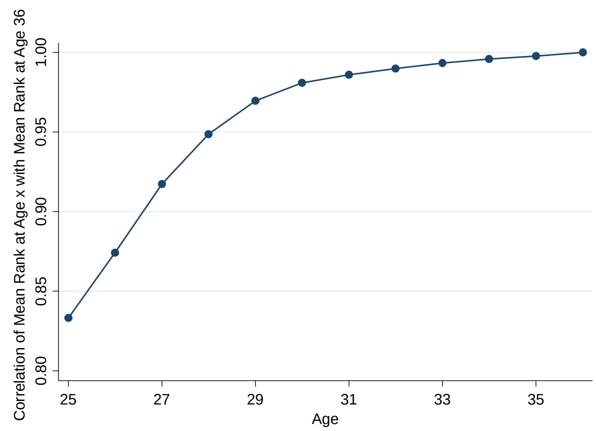
APPENDIX FIGURE I

Children's Earnings Ranks by Age of Earnings Measurement

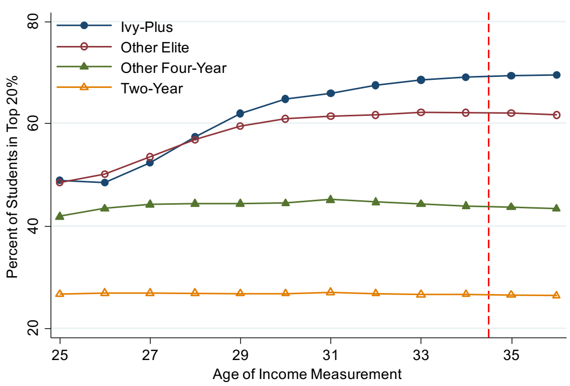
A. Mean Earnings Rank by Age and College Tier



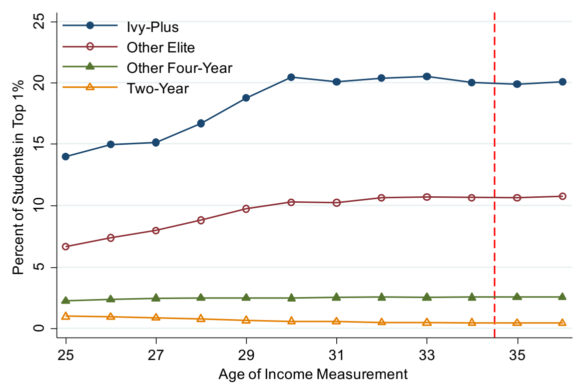
B. Correlation of College Mean Earnings Rank across Ages



C. Fraction of Children in Top Quintile



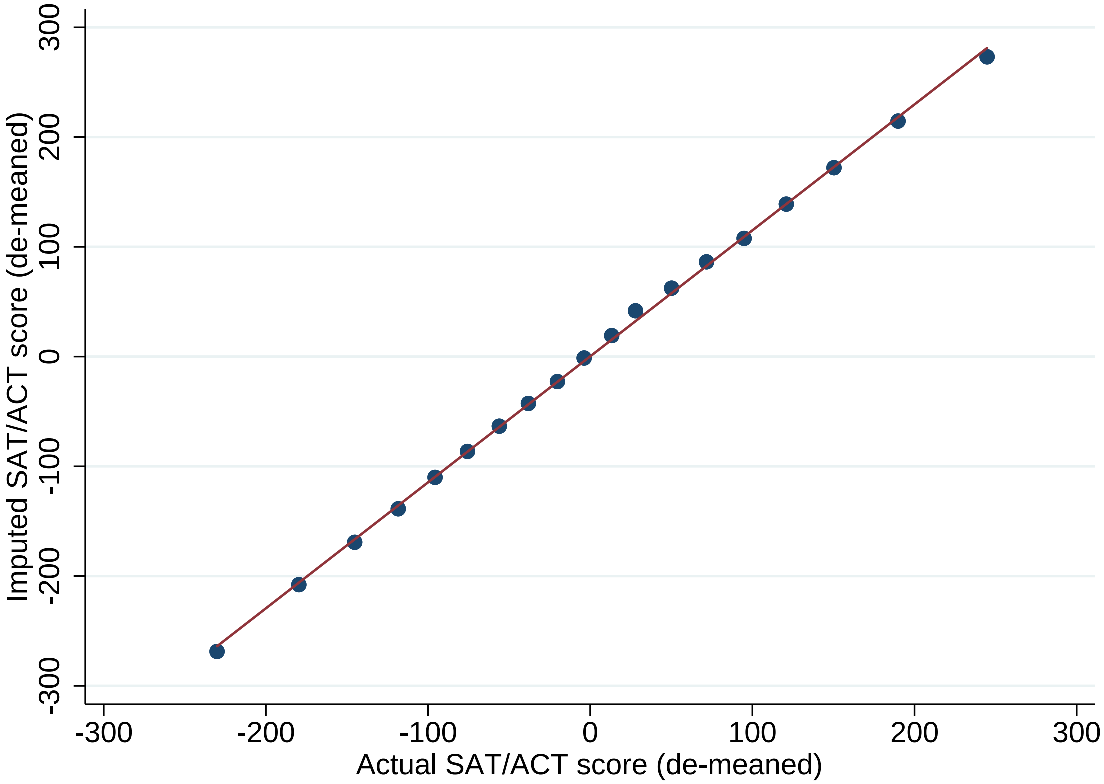
D. Fraction of Children in Top 1% by Age and College Tier



Notes: Panel A plots the mean income rank by age for students who attended colleges in various tiers. Children's incomes are defined as the sum of individual wage earnings and self-employment income. We measure children's incomes at each age from 25 to 36 and assign them percentile ranks at each age based on their positions in the age-specific distribution of incomes for children born in the same birth cohort. Elite colleges are split into Ivy-Plus (the eight Ivy-League colleges as well as the University of Chicago, Stanford University, MIT, and Duke University) and Other Elite (all other elite colleges). Panel B plots the (enrollment-weighted) correlation between the college-level mean income rank of students at age 36 with the college-level mean income rank at ages 25-36. Panels C and D replicate Panel A, changing the outcome variable to the percentage of children who reach the top quintile (Panel C) or top 1% (Panel D) of their age- and cohort-specific earnings distribution. To maximize the age range at which incomes are observed, we use data for children in the 1978 birth cohort in this figure, with individuals assigned to the college they attended at age 22 (in 2000). Because children cannot be linked to parents before the 1980 birth cohort, we use data starting with the 1980 cohort and only observe income up to age 34 in our main analysis. This figure is constructed directly from the individual-level microdata.

APPENDIX FIGURE II

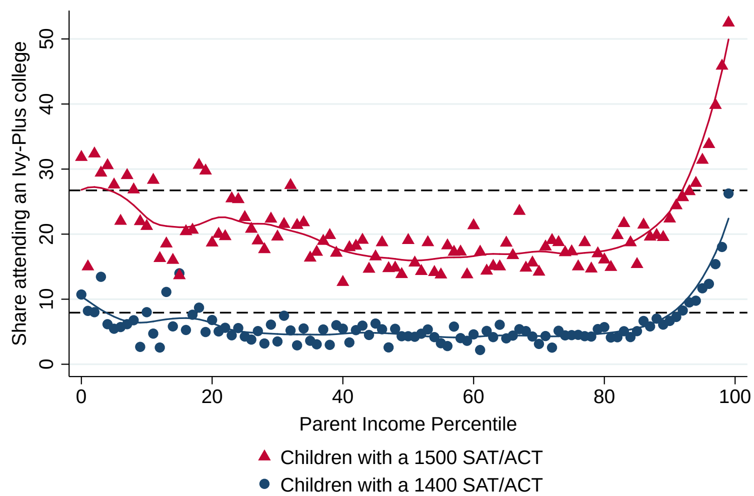
Validation of SAT/ACT Imputation



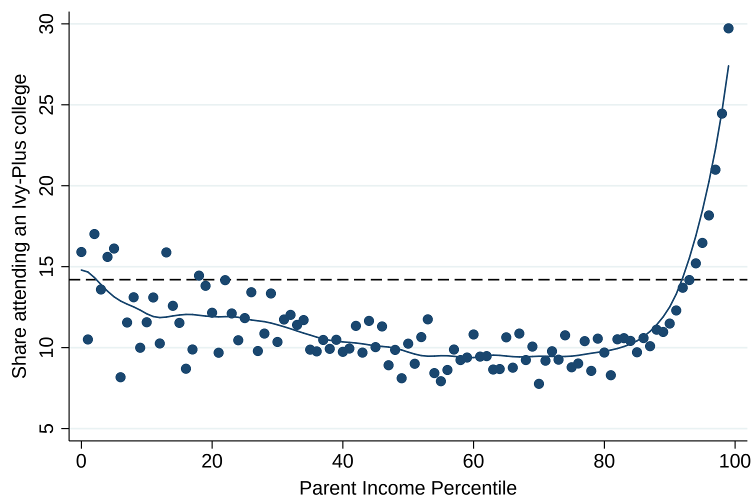
Notes: As noted in Section IV, we impute an SAT/ACT score to the 26.2% of college goers missing an SAT and ACT score using the SAT/ACT score of the college student from the same parent income quintile, same college selectivity tier, and same state who has the closest level of earnings in adulthood. This figure presents a quantile-quantile plot of imputed SAT/ACT score versus actual SAT/ACT score, within college-parent income quintile, using data from five states where the SAT or ACT is administered to essentially all students. We construct the graph as follows. First, we set actual SAT/ACT scores to missing for college goers in the five states (measured using the college-goer's parent ZIP code) with the highest SAT/ACT coverage rates in our data (which range from 89% to 91% of college goers). We then run the imputation procedure described above by parent income quintile and tier (not state). Then, within each college-parent income quintile cell and *restricting to students whose actual SAT/ACT scores were set to missing*, we de-mean imputed SAT/ACT scores and compute ventile thresholds (i.e., the 5th percentile, 10th percentile,..., 95th percentile). We similarly compute de-meaned ventile thresholds for these students' actual SAT/ACT scores. Finally, we restrict attention to the ten colleges with the highest enrollment of these students and plot unweighted mean imputed quantiles versus unweighted mean actual quantiles (e.g., the bottom-left dot is mean imputed 5th percentile versus mean actual 5th percentile). The slope of the best-fit line (1.15) is near one with a constant (0.20) near zero. Hence, within college by parent income quintile cells, the distribution of imputed SAT/ACT scores nearly matches the distribution of actual SAT/ACT scores. Recall that the graph pools across college-parent income quintiles. When repeating the analysis separately for each parent income quintile, the slopes range from 1.12 to 1.20 and constants range from -0.8 to 1.1. When repeating the analysis using the top-10 selective colleges and separately using the top-10 unselective colleges, the slopes range from 1.14 to 1.23 and the constants range from 0.2 to 0.4. This figure is constructed directly from the individual-level microdata.

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

A. Students with Scores of 1400 or 1500 on SAT/ACT



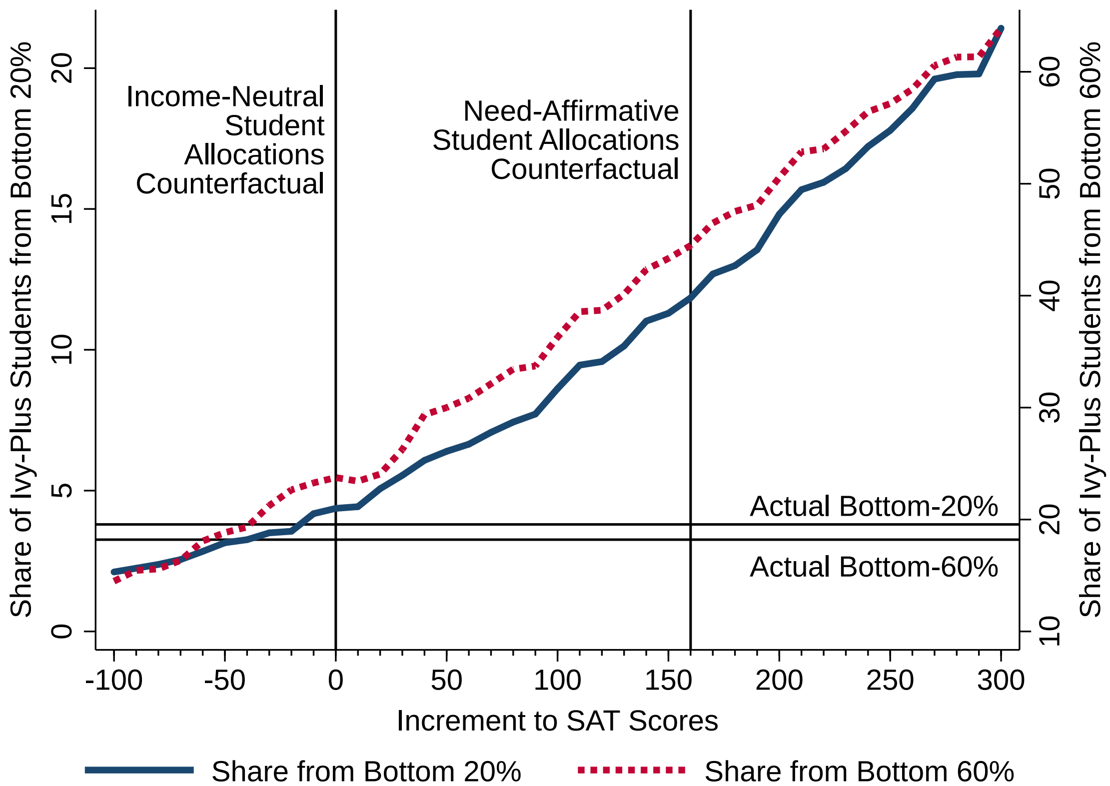
B. All SAT/ACT Scores, Reweighted to Match Ivy-Plus SAT/ACT Score Distribution



Notes: Panel A plots Ivy-Plus college attendance rates by parental income percentile for students with a 1400 or 1500 SAT/ACT score. The 1400 series exactly replicates Figure Ib. Because relatively few low-income students have a 1500 test score, the 1500 series pools students with an SAT/ACT score between 1480-1520. Panel B replicates Figure Ib pooling all SAT/ACT scores, weighted by the SAT/ACT score distribution of actual Ivy-Plus attendees. That is, Panel B plots the share of students who would attend an Ivy-Plus college by parent income percentile if each percentile's test score distribution matched the test score distribution of Ivy-Plus students. See notes to Figure Ib for additional details. This figure is constructed directly from the individual-level microdata.

APPENDIX FIGURE IV

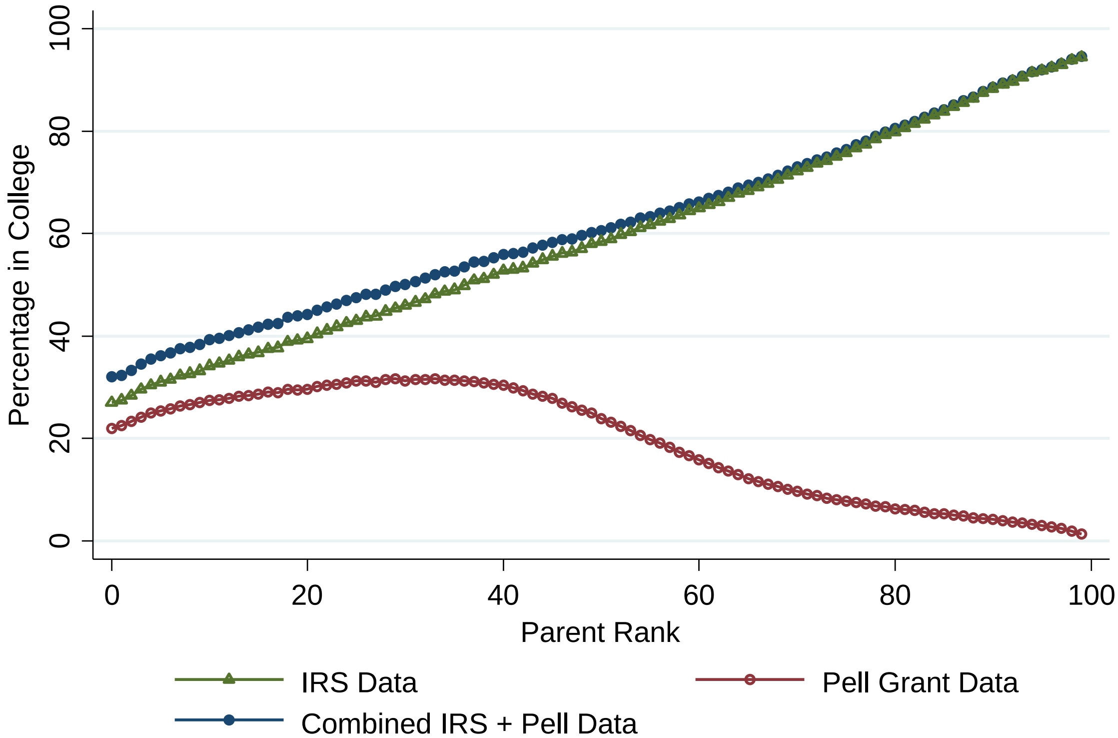
Counterfactual Low-Income Shares at the Ivy-Plus



Notes: This figure plots the share of students from the bottom-20% (left y-axis scale) and bottom-60% (right y-axis scale) in the Ivy-Plus tier, varying the constant added to bottom-20% college-goers' SAT/ACT on the x-axis. Second, third, and fourth quintile college goers' SAT/ACT scores are incremented upward by 80%, 60%, and 40% of the bottom-20%'s upward bonus, respectively. These shares are computed following the method used to construct the need-affirmative student allocations counterfactual described in Section IV.A. The vertical lines show the shares that result from our baseline income-neutral student allocations counterfactual (0 point increment for low-income students) and baseline need-affirmative student allocations counterfactual (160 point increment). The horizontal lines show the actual shares of students from the bottom 20% and bottom 60% at Ivy-Plus colleges in our analysis sample. This figure is constructed directly from the individual-level microdata.

APPENDIX FIGURE V

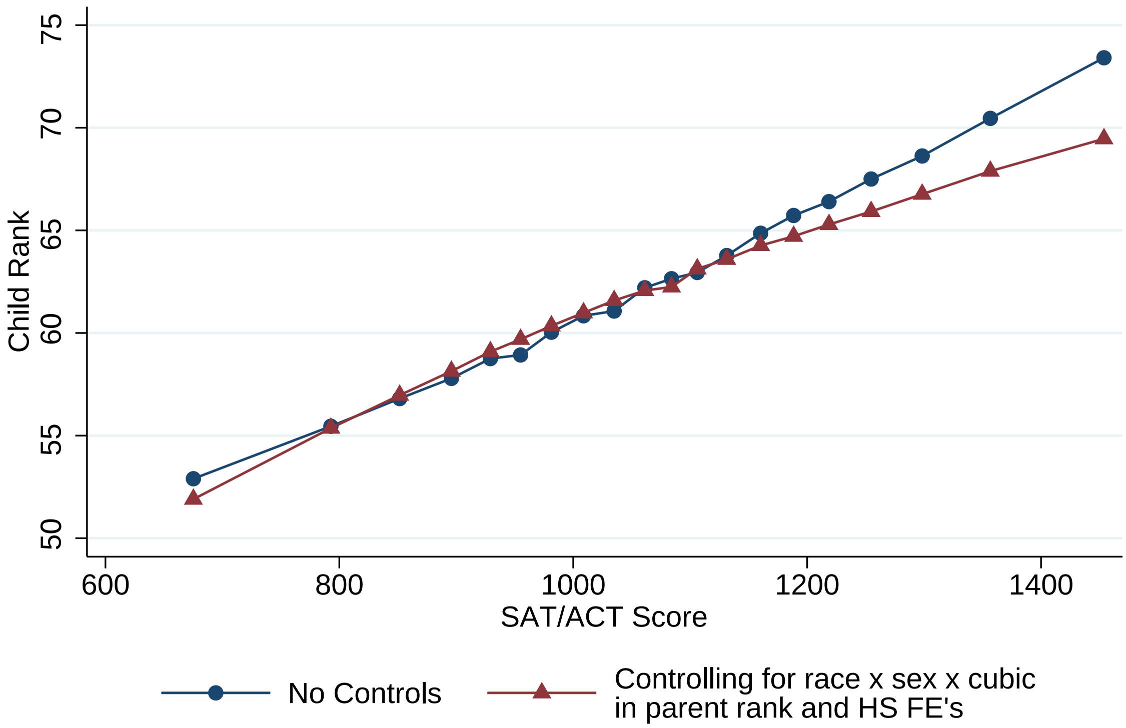
College Attendance Rates in 1098-T and Pell Records by Parent Income



Notes: This figure plots the fraction of students in the 1980-82 birth cohorts in our analysis sample who attend college at any time during the years in which they turn 19-22 by parental income percentile. The series in open circles plots the fraction of students in each parental income percentile with a college attendance record in the NSLDS data only. The series in triangles plots the fraction of students in each parental income percentile with a college attendance record in the 1098-T data only. The series in solid circles plots the fraction who attend college based on the union of the NSLDS and 1098-T data, the measure of attendance we use in our empirical analysis. This figure is constructed directly from the individual-level microdata.

APPENDIX FIGURE VI

Relationship Between SAT Scores and Earnings in Adulthood



Notes: This figure shows the association between students' earnings ranks in 2014 and their standardized test scores (SAT and ACT). The sample includes all college-goers in our 1980-1982 cohorts for whom we have either SAT or ACT scores. We convert ACT scores to the SAT 1600-point scale. We construct binned scatter plots by first regressing children's ranks on twenty indicators (5 percentile point bins) for their test scores and a set of controls. In the series in blue circles, the only control is an indicator for whether the student took the SAT, the ACT, or both tests; in the series in red triangles, we additionally control for a cubic in parental income rank interacted with race and sex as well as high school fixed effects. We then plot the estimated child ranks and mean SAT scores within each of the twenty bins, recentering both variables so that their means match the overall sample means. This figure is constructed directly from the individual-level microdata.